United States Air Force
Scientific Advisory Board

Report on

Building the Joint Battlespace Infosphere

Volume 2: Interactive Information Technologies

SAB-TR-99-02

December 17, 1999

Cleared for Open Publication – February 2000

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)
01-12-1999

2. REPORT TYPE
Technical

3. DATES COVERED (FROM - TO)
xx-01-1999 to xx-12-1999

4. TITLE AND SUBTITLE
Building the Joint Battlespace Infosphere
Volume 2
Interactive Information Technologies
Unclassified

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McCarthy, J ; Author
Frost, D ; Author
Katz, R ; Author
Sproull, R ; Author
Gavron, V ; Author

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
USAF Scientific Advisory Board
Room 5D982
1180 AF Pentagon
Washington, DC20330-1180

8. PERFORMING ORGANIZATION REPORT NUMBER
SAB-TR-99-02

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
SAF/OS
AF/CC
Pentagon
Washington, DC20330-1670

10. SPONSOR/MONITOR'S ACRONYM(S)
SAF/OS; AF/CC

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
SAB-TR-99-02

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The Joint Battlespace Infosphere (JBI) is a combat information management system that provides individual users with the specific information required for their functional responsibilities during crisis or conflict. The JBI integrates data from a wide variety of sources, aggregates this information, and distributes the information in the appropriate form and level of detail to users at all echelons. The JBI was originally described in the 1998 USAF Scientific Advisory Board (SAB) report Information Management to Support the Warrior. In Chapter 4 of Volume 1 of the report, some interaction technologies were described in the context JBI functions: command, planning, execution, and combat support. In Volume 2, a much wider variety of interaction technologies is examined in greater detail. The goal of Volume 2 is to ensure that the masterpiece that is the JBI technical infrastructure is not partnered with clumsy, outdated user interfaces. Furthermore, the goal of the volume is to make JBI developers plan for future interaction technologies and not simply project current interaction techniques onto the JBI of the future.The volume places interaction techniques into three categories: 1. Capture, which is the input of information to the JBI. 2. Presentation which is concerned with how the users perceive information. 3. Collaboration which focuses on shared workspaces for multiple users.

15. SUBJECT TERMS
Battlespace Infosphere, C4I, command and control, decision support tools, combat support, data fusion, fuselet, battlespace awareness, information management, information access, information grid, information manipulation, collaboration, force templates

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE
Unclassified	Unclassified	Unclassified

17. LIMITATION OF ABSTRACT
Same as Report (SAR)

18. NUMBER OF PAGES
116

19. NAME OF RESPONSIBLE PERSON
Ripperger, Robert
robert.ripperger@pentagon.af.mil

19b. TELEPHONE NUMBER
International Area Code
011
Area Code Telephone Number
703692-5097
DSN
222-5097

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39.18

This report is a product of the U.S. Air Force Scientific Advisory Board Committee on **Building the Joint Battlespace InfoSphere.** *Statements, opinions, recommendations, and conclusions contained in this report are those of the Committee and do not necessarily represent the official position of the U.S. Air Force or the Department of Defense.*

United States Air Force
Scientific Advisory Board

Report on

Building the Joint Battlespace Infosphere

Volume 2: Interactive Information Technologies

SAB-TR-99-02

December 17, 2000

Cleared for Open Publication – February 2000

(This Page Intentionally Blank)

Table of Contents

List of Figures

List of Tables

(This Page Intentionally Blank)

Chapter 1: Introduction to Interactive Information Technologies

The main objective of the Joint Battlespace InfoSphere (JBI) is to provide the right information to the right user at the right time in the right languages and the right media at the right level of detail with the right information analysis tools. Much of the technical infrastructure of the JBI is built around the collection, organization, and aggregation of information. However, if the JBI is to be successful, its technical operation must be tied with interaction mechanisms best supporting users' needs. That is, the JBI developers must pay as much attention to providing information in the "right language," "the right media," and "the right information analysis tools" as they do to using the appropriate object-based middleware infrastructure. Further, much of the JBI's power comes from its pervasiveness (warfighters at all levels use it) and only a fraction of these users should be expected to interact with the JBI via a cathode ray tube display, a keyboard, and a mouse.

In Chapter 4 of Volume 1 of this report, some interaction technologies were described in the context JBI functions: command, planning, execution, and combat support. In this volume, a much wider variety of interaction technologies is examined in greater detail. The goal of this volume is to ensure that the masterpiece that is the JBI technical infrastructure is not partnered with clumsy, outdated user interfaces. Furthermore, the goal of this volume is to make JBI developers plan for future interaction technologies and not simply project current interaction techniques onto the JBI of the future.

This volume places interaction techniques into three categories. The first category is *capture*, which is the input of information to the JBI. Some of the key capture technologies discussed in Chapter 2 are:

- *Conversational query and dialog.* This technology focuses on two-way information transfer between at least two agents, presumably at least one human and one computer.
- *Speech and natural language.* These technologies free the hands of the user and let input occur more naturally.
- *Multimodal interfaces.* These techniques combine technologies with a promising approach using speech with gestures.
- *Drill down.* This technology supports search through vast quantities of data to pull out relevant information.
- *Personal computing devices.* These devices enable warfighters in the field to interact with the JBI using portable devices.
- *Automatic data capture.* These techniques focus on inputting data in an efficient and user-friendly way (for example, scanning barcodes) and making data available to a larger system.

The second category of interaction techniques is *presentation*. Presentation is concerned with how the users perceive information. Some of the key presentation technologies presented in Chapter 3 are:

- *Personal display devices.* These include virtual retinal displays and haptic (that is, force-reflecting) interfaces.

- *Data visualization.* These techniques give users the view of data that provides the needed insight for the tasks being performed.
- *Three-dimensional (3-D) audio.* These techniques help attract the user's attention or focus attention in a particular location.
- *Tailoring.* These technologies match the interface to people and their jobs.

The third group category of interaction techniques is *collaboration*. Collaboration focuses on shared workspaces for multiple users. The challenge in this area is to find the right way to have multiple users share, and perhaps change, information over distance and/or time.

All of these technologies can be applied to the JBI, thereby adding value to the already significant processing done by the JBI's technical infrastructure. This value is added where it matters most: in making warfighters more effective.

Chapter 2: Capture

Capture is the input of information to the JBI. Capture technologies include conversational query and dialog, speech, natural language, multimodal interfaces, annotation, drill down, personal computing devices, and automatic data capture. Each of these is discussed in greater detail below. Communication of information must evolve to be more human-interaction friendly. Capture technologies are critical to the success of the JBI. According to work by Dr. Bonnie John of Carnegie Mellon University (CMU), only about 10 percent of the JBI will be used well with current technologies. One way to enhance the use of the JBI is to develop a model of ideal human use, which could become the basis for training in efficient strategies. Efficient strategies are especially needed to detail, aggregate, manipulate, modify (all or exception), and locate. Efficient strategies use the strengths of the computer for calculating, iterating, and visualizing. John notes that people persistently use inefficient strategies because they have (1) incorrect knowledge, (2) poor interfaces, (3) incomplete experience, and (4) no explicit training. Wizards are good for well-defined, predetermined strategies and require few decision nodes but are useless under other conditions such as users of the JBI will encounter.

2.1 Conversational Query and Dialog

Conversation is the "information transfer between at least two agents in both directions with the agents taking turns in speaking."[1] A major challenge in conversational dialog is the ability to handle mistakes.

Some experimental systems include both speech recognition and synthesis. One example is the Communicator programmer,[2] a complex problem-solving system with personal agent interfaces. Communicator is a travel planner based on observation of two humans performing this task, subsequent Wizard of Oz simulations (in which the experimenter sits behind a curtain and performs the tasks a new technology would perform such as real-time response to spoken queries), data from successive system prototypes, and data available on the web (for example, expedia.com). It uses Microsoft synthesizer, a formant synthesizer (A formant is a pattern of sound waves that make up a vowel utterance.). Concatenative synthesis is being developed for Communicator. This synthesis is based on concatenating very large bodies of short speech utterances. The result sounds more human but requires a lot of memory. Near-term technology could be a simple speech recognition system inside a cellular telephone. The system would then enable wideband speech to be efficiently encoded and transmitted digitally.

Gary Strong, the Defense Advanced Research Projects Agency (DARPA) program manager, described the Communicator programmer. There are three communication problems: (1) text-to-speech synthesis to speak machine output, (2) speech recognition to hear and understand spoken dialog, and (3) voice recognition to verify speaker identity. The goal is to develop dialog interaction that is wireless and mobile, requires no keyboard, and provides context tracking. Querying a human to elicit information will accelerate the mixed initiative. The computer will

[1] Perlis, Purang, and Andersen, 1998, p. 554.
[2] An example of Communicator as applied to making travel arrangements is available at 1-877-268-7526.

initiate contact and all communication will be through spoken language. Significant players are AT&T, CMU, GTE/BBN, Hughes, Lucent, IBM, Lockheed Martin, Hughes, and the Massachusetts Institute of Technology (MIT). Significant advances to date include stochastic modeling of mixed-initiative dialog, hub and module architecture, a stochastic approach to entity discovery in transcribed speech, acoustic cancellation, and an evaluation strategy based on a confusion matrix representation of filled forms. A demonstration will be conducted in June 2000. The innovation is from the shared architecture; a policy group focused on architecture, standards, and content; heavy industrial involvement based on shared architecture; and a technology transition shared with a translingual program.

The Rochester Interactive Planning System[3] (TRIPS) being developed by the University of Rochester is a logistic planner with speech recognition and synthesis, natural language understanding, dialog management, and a scheduler. It supports evaluation of alternatives.

Broadsword[4] is an extensible client-server framework consisting of services and tools to help a user collect intelligence information from heterogeneous and distributed data sources to support information operations. It is divided into three functional areas: Gatekeeper, User Services, and Additional Services. Broadsword provides access to information available from the Secret Internet Protocol Router Network, and the Non-Secure Internet Protocol Router Network. The presentation layer of the system is the point of contact with the user. It contains conventional query services. A query string is converted into keywords used to search free text and databases. Queries are forwarded to data sources through a plug-in Data Interface Agent. The Broadsword team consists of the Air Force Research Laboratory (AFRL/IF), Booz-Allen & Hamilton, Synectics, and State University of New York Institute of Technology.

James Allen and George Ferguson of the University of Rochester are developing a dialog-based approach to mixed-initiative plan management (of which TRIPS is the current prototype). Mixed-initiative plan management (that is, plan construction, evaluation, modification, and execution monitoring) is currently hindered by the lack of a well-developed technology for communicating plans. In most current systems, the human has limited facilities for specifying plans and the machine has limited capabilities for displaying and describing plans. No system can reason about the most efficient way to communicate a plan using different modalities, can support substantial elaboration and clarification subdialogs about plans to a substantial extent, or can support intelligent interactive plan browsing and question answering. Allen and Ferguson propose developing and demonstrating a dialog-based model of plan communication that supports mixed-initiative interactions with integrated graphic display of maps and charts, menus, mouse gestures, and natural language, to enable effective communication about plans. The focus is on the architecture for enabling effective communication for plan management tasks rather than on particular techniques for plan presentation within a modality (such as how to generate better displays).

[3] The acronym is *TRIPS*; the website is http://www.cs.rochester.edu/research/trips/.
[4] See http://www.if.afrl.af.mil/bsword/.

Developers assume an agent-based architecture in which the human-machine interaction is viewed as a dialog. In other words, each new interaction is interpreted in the context of prior interactions, enabling complex plans to be explored, developed, and discussed in the incremental fashion typical of human dialog. Drawing on previous work on dialog modeling, these developers use a three-level model of the interaction. The domain level involves reasoning about the domain—for example, in the Airspace Control Plan (ACP), reasoning about target selection and resource selection. The task level involves reasoning about the problem-solving process itself (for example, in ACP, first determining centers of gravity (COGs) and air objectives, then developing targets). The dialog level involves reasoning about the interaction (for example, in ACP, determining the most effective way to summarize a subplan). All three levels are necessary to produce effective mixed-initiative interaction. These developers are focusing on defining this architecture and defining the task and dialog levels for mixed-initiative plan management systems.

The system will use knowledge of actions at the task level to represent how humans analyze problems and how they interact with the system. It will use knowledge at the discourse level to reason about the effective communication of plans and scenarios. Thus, managing the interaction itself is a planning and execution task, in which the system reasons about the communicative needs and goals of the user and how best to achieve them. A model of obligations and responsibilities, which can be changed interactively to the user’s preferences, drives the system’s behavior. Because the system reasons explicitly about its own goals, obligations, and responsibilities, it will exhibit truly mixed-initiative planning not possible with a traditional planning system. Because of its rich model of the dialog state and communication strategies, it will communicate plans in a way better suited to human needs than was possible in previous systems.

Successful completion of this research will have a significant impact on a wide range of applications, including military planning systems - such as ACP and transportation planning and civilian applications such as crisis management and information retrieval. Effective communication of plans between human and machine is one of the foremost obstacles to true mixed-initiative plan management. Indeed, effective communication of plans is one of the foremost obstacles to all computer use. Getting a computer to do the user’s bidding requires that it know the user’s plans, and knowing what a computer is going to do requires that it summarize and explain the plan that it is following. Better means of expressing goals and plans to computers will eventually revolutionize computer use. This work takes the first key step toward this long-term goal.

Alan Bierman (Duke University) is developing an architecture for voice dialog systems based on Prolog-style theorem proving. A pragmatic architecture for voice dialog machines aimed at the equipment repair problem has been implemented. This architecture exhibits a number of behaviors required for efficient human-machine dialog. These behaviors include (1) problem solving to achieve a target goal; (2) the ability to carry out subdialogs to achieve appropriate subgoals and to pass control arbitrarily from one subdialog to another; (3) using a model to enable useful verbal exchanges and to inhibit unnecessary ones; (4) the ability to change initiative from strongly computer-controlled to strongly user-controlled or to some level in between; and (5) the ability to use context-dependent expectations to correct speech recognition and track user movement to new subdialogs.

Lisa Harper of the MITRE Corporation is developing an architecture for dialog management, context tracking, and pragmatic adaptation in spoken-dialog systems. MITRE is focusing on a software architecture for discourse processing using three component tasks: (1) dialog management, (2) context tracking, and (3) pragmatic adaptation. MITRE defines these tasks and describe their roles in a complex, near-future scenario in which multiple humans interact with each other and with computers in multiple, simultaneous dialog exchanges. A motivation for this work is the use of reusable discourse processing software for integration with nondiscourse modules in spoken-dialog systems.

MITRE is working on an architecture for spoken-dialog systems for both human-computer interaction (HCI) and computer mediation or analysis of human dialog. The architecture shares many components with existing spoken-dialog systems, such as CommandTalk,[5] Galaxy,[6] TRIPS,[7] Verbmobil,[8] and Waxholm.[9] MITRE's architecture is distinguished in its treatment of discourse-level processing. Most architectures contain modules for speech recognition and natural language interpretation (such as morphology, syntax, and sentential semantics). Many include a module for interfacing with the back-end application. If the dialog is two-way, the architectures also include modules for natural language generation and speech synthesis. Architectures differ in how they handle discourse. Some have a single, separate module labeled "discourse processor," "dialogue component," or "contextual interpretation." Others, including earlier versions of the system, bury discourse functions inside other modules, such as natural language interpretation or the back-end interface. Innovations of this work are the compartmentalization of discourse processing into dialog management, context tracking, and pragmatic adaptation and the software control structure for interaction between these and other components of a spoken-dialog system.

Phil Cohen of the Oregon Graduate Institute is a leader in multimodal systems. A new generation of multimodal systems is emerging in which the user will be able to employ natural communication modalities, including voice, hand and pen-based gestures, eye tracking, and body movement,[10,11,12] in addition to the usual graphical user interface (GUI) technologies. To make progress on building such systems, a principled method of modality integration, and a general architecture to support it is needed. Such a framework should provide sufficient flexibility to enable rapid experimentation with different modality integration architectures and applications. This experimentation will enable researchers to discover how each communication modality can best contribute its strengths yet compensate for the weaknesses of the others.

[5] Moore et al., 1997.
[6] Goddeau et al., 1994.
[7] Allen et al.,1995.
[8] Wahlster, 1993.
[9] Carlson, 1996.
[10] Koons et al., 1993.
[11] Oviatt, 1992, 1996.
[12] Waibel et al., 1995.

QuickSet is a collaborative, multimodal system that employs such a distributed, multi-agent architecture to integrate not only the various user interface components, but also a collection of distributed applications. QuickSet provides a new *unification-based* mechanism for fusing representation fragments derived from the input modalities. In so doing, it selects the best *joint* interpretation according to the underlying spoken language and gestural modalities. Unification also supports multimodal discourse. The system is scalable from handheld to wall-size interfaces, and operates across a number of platforms (from personal computers to UNIX workstations). Finally, QuickSet has been applied to a collaborative military training system, in which it is used to control a simulator and a 3-D virtual terrain visualization system.

QuickSet is a collaborative, handheld, multimodal system for interacting with distributed applications. In virtue of its modular, agent-based design, QuickSet has been applied to a number of applications in a relatively short period of time, including

- *Simulation setup and control*. Quickset is used to control LeatherNet,[13] a system employed in training platoon leaders and company commanders at the Marine Corps base at Twentynine Palms, California. LeatherNet simulations are created using the ModSAF simulator[14] and can be visualized in a wall-size computer-assisted virtual environment[15, 16] called CommandVu. A QuickSet user can create entities, give them missions, and control the virtual reality (VR) environment from the handheld personal computer. QuickSet communicates over a wireless local area network via the Open Agent Architecture[17] to ModSAF, and to CommandVu, which have all been made into agents in the architecture.
- *Force laydown*. QuickSet is being used in a second effort called ExInit (Exercise Initialization), which enables users to create large-scale (division- and brigade-size) exercises. Here, QuickSet operates via the agent architecture with a collection of Common Object Request Broker Architecture servers.
- *Medical informatics*. A version of QuickSet is used in selecting healthcare in Portland, Oregon. In this application, QuickSet retrieves data from a database of 2,000 records about doctors, specialties, and clinics.

The objective of BBN's Rough'n'Ready project is to develop a practical system that provides flexible access to information in recorded collaborative events. The system will provide "rough" transcriptions of collaborative events (such as meetings, presentations, and conference calls), which are "ready" for browsing with an appropriate set of tools. A user will be able to access an event from a large remote archive or retrieve a particular part of an event by searching with multivalued queries composed of any combination of topic, proper name, speaker identity, range of dates, or a full-string search.

The approach is based on the integration of several speech and language technologies to produce a structural summary of collaborative events. These technologies include speech recognition, speaker identification, topic and named-entity spotting, and information retrieval. The BBN BYBLOS speech recognition system is being used to produce a rough transcription of the audio

---

[13] Clarkson and Yi, 1996.
[14] Courtmanche and Ceranowicz, 1995.
[15] Cruz-Neira et al., 1993.
[16] Zyda et al., 1992.
[17] Cohen et al., 1994.

track of a recorded event. Names of people, organizations, and locations are found using the BBN statistical named-entity spotting system, IdentiFinder. A new topic classification algorithm recently developed at BBN is used to index the transcriptions by topic. This algorithm enables a story to be classified into several topics out of a set of thousands of possible topics. A new information-retrieval algorithm that is also being developed at BBN will be used to support full-string searches on the automatic transcription. Speaker-identification algorithms developed at BBN are used to locate portions of the audio from each speaker and to label them with the speaker's identity when known. The combined output of these components provides a compact content-based structural summary of large audio archives that will support advanced visualization and navigation capabilities. The summary also forms the basis for highly selective multivalued queries used for retrieval of specific events from the archive.

The difference between data visualization and information visualization is the difference between raw data and the organization of data, their relationships, and their relevance to the task at hand. Good information visualizations reduce perceptual, interpretative, and cognitive burdens by making the visuals relate to specific tasks. Figures 1 and 2 are representative visualizations that Visible Decisions Inc. (VDI) has developed for two domains—logistics and emergency management.

VDI's products range from development tools (for example, In3-D C++, Java, and ActiveX) to end-user tools (for example, SeeIT). VDI's foundation is the In3-D/C++ class library, which, in combination with a builder called In3-D/Studio, enables software developers to construct interactive 2-D and 3-D information visualization applications. The C++ and Java editions are available on Sun, SGI, and Windows. The In3-D class library is VDI's third generation of products since it started in 1992. The core of In3-D is a model, view, controller (MVC) object-oriented library of more than 300 objects and 4,000 methods.

VDI's technology has been built for high performance (using compiled applications, a multi-threaded library, preemptive rendering, levels of detail, real-time data, and multiple concurrent data sources), seamless integration (linking such industry standards as Open Graphics Library (OGL), Microsoft© Foundation Classes (MFC), Motif, and ActiveX to your favorite library), extensibility and scalability (C++ and Model View Controller (MVC)), and ease of deployment (desktop or distributed browser or server applications, low memory overhead, and small disk footprint). Most important, VDI has *an information visualization focus:* the 2-D and 3-D interactive display of data-intensive, dynamic objects and properties.

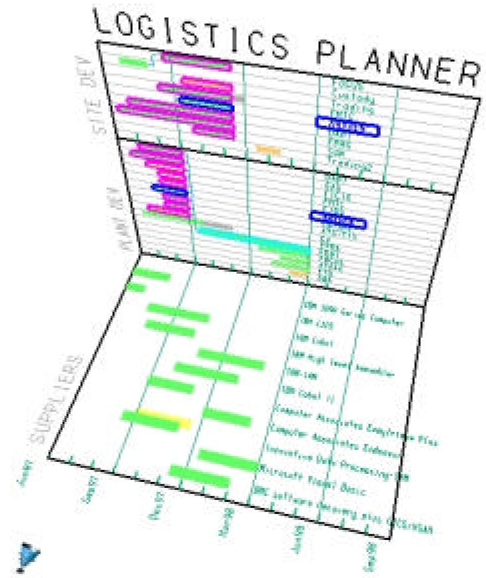

Figure 1. *Logistics planning.*

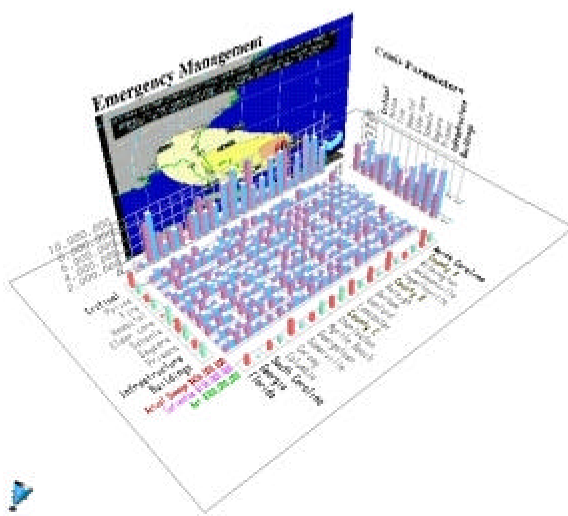

Figure 2. *Emergency management.*

2.2 Speech

The human interface of the future will be characterized by much more natural modes of interaction than are currently possible. While some progress has been made over the years in the area of speech recognition and handwriting analysis, computers are still very far from being able to interact with people in the same way that other people do. Meeting the challenge of creating portable-assistant technologies will consist in large part of enhancing the modes of interaction by which the user is able to input intentions. Speech recognition, eye movement tracking, gesture recognition, and handwriting analysis will all be key components of a naturalistic interface. Successful inference of operator intent from these input modalities will rely heavily on the development of cognitive and perceptual models. In a longer time range, the National Aeronautics and Space Administration (NASA) will explore immersive manipulator control and other advanced concepts for direct-manipulation interfaces.

Many factors can decrease speech recognition. One of these is hyperarticulate speech: "elongation of the speech segment and large relative increases in the number and duration of pauses, … more hyper-clear phonological features, fewer disfluencies, and change in fundamental frequency."[18]

Microsoft is researching speech recognition, synthesis, and personalized voices.[19] DARPA, which focuses on obscure languages, has been a strong supporter of language translation and speech recognition.[20] The National Science Foundation (NSF) focuses on the major world languages.[21] NSF and the European Commission are funding multilingual information access and management research. Jaime Carbonell, director of the Language Technologies Institute at CMU, is the leading person in this area. Victor Zue of MIT is trying another approach. His system can provide weather information for any city in the United States via telephone requests. Tom Landauer of the University of Colorado's Institute of Cognitive Science is trying to train systems. The Text Retrieval Evaluation Conference is run by Donna Whist at the National Institute of Standards and Technology. A compact disc read-only memory (CD-ROM) is provided to conference attendees to answer questions. Speech synthesis is another focus since there is less commercial interest in high-quality, human-like speech. There is a recent report on the state of the art on the NSF website. In addition, NSF supports the University of Pennsylvania on language usage.

A speech-based input or control system provides a convenient hands-free method to interact with computer applications. With proper design, speech input and data manipulation can be faster than more conventional computer interaction methods. Automatic speech recognition (ASR) is the main technology embedded in a speech-based control system. A wide variety of approaches have been developed to extract meaning from an acoustic signal. These methods can be used to facilitate data entry into a JBI and end-user interaction. For example, AFRL has developed paradigms to identify and sort signals of interest to intelligence operators for improved monitoring and reporting of information. Speech control has been demonstrated as an input method in various military aircraft and airborne command and control situations. More recently, the value of speech control has been demonstrated in a Joint Air Operations Center (JAOC) context.

ASR technology has significantly matured in the last several years. Due to dramatic increases in processor speed and memory availability, real-time continuous speech ASR systems are becoming more commonplace and will soon be the preferred human-computer interface technology. Currently available ASR systems generally fall into one of three categories:

- *Dictation systems.* This represents the largest market segment of ASR technology and allows direct speech-to-text dictation for document generation. Companies such as Dragon Systems, IBM, and

[18] S. Oviatt, M. MacEachern, and G. Levow, "Predicting Hyperarticulate Speech During Human-Computer Error Resolution," *Speech Communication*, Vol. 24 (1998), p. 87.
[19] Information-gathering meeting at Microsoft, 16 April 1999.
[20] Information-gathering meeting at NSF, 19 May 1999.
[21] http://www.linglink.lu/hlt/download/mlim.html

Lernout & Hauspie offer products for less than $150. In addition to dictation, these products enable limited navigation of various programs within the Microsoft Windows environment.

- *Computer command and control systems.* These systems allow the user to perform navigation and data entry functions by speech. These systems are being used in a variety of settings including industrial and automotive applications. Many of these systems are speaker independent, meaning that the user does not need to train the application vocabulary prior to use.
- *Computer telephony systems.* This last category is the fastest growing ASR market segment and enables users to interact with computers over standard telephone channels. Applications such as order entry, stock trading, airline reservations, and auto attendant are significantly reducing operating expenses by reducing the need for human operators. These systems typically handle call volumes of 150,000 or more per day and provide a rapid return on investment.

AFRL/HEC has been actively engaged in research to exploit commercial ASR technology in this wide variety of military applications. One example is a recent experiment designed to evaluate the military utility of a speech input system in the production of air tasking orders (ATOs) in a JAOC environment. A prototype speech recognition interface was built into the Theater Air Planning module of Theater Battle Management Core Systems. This interface enabled users to quickly navigate menus, enter mission-planning data, and perform database queries through speech. Nuance, a computer telephony product by Nuance Communications, was chosen as the speech recognition system on the basis of its proven performance and scalability.

Two assessment sessions were performed. The first session consisted of nine subjects from the 505th Command and Control Training and Innovation Center. The second session consisted of eight subjects from several of the Numbered Air Forces, one from the Navy, and one from the Marine Corps. Several of the personnel had no prior experience with ATO production. After familiarization training, each subject participated in six planning exercises, three using speech recognition and three using the conventional mouse-and-keyboard interface.

When speech recognition was used, results showed a reduction of 10 to 20 percent in the time required to complete planning exercises, as well as a reduction in learning curves. Nuance's recognition performance was better than 97 percent for both sessions. Current-generation ASR systems are proving to be a viable alternative to conventional human-computer interface technologies in a growing number of applications. There are some limitations, however, that need to be overcome before ASR technology can gain widespread acceptance:

- *Dialog constraints.* To maintain high performance for command and control applications, current ASR systems require grammar models that indicate which commands can be spoken at any given time in the application. If the user varies significantly from the grammar model, the ASR system either rejects the input or substitutes incorrect commands. Additional research is needed to increase flexibility while maintaining high performance.
- *Noise robustness.* Current ASR systems operate well in low- to mid-noise conditions. High-payoff applications in areas such as flight line maintenance and cockpit command and control require operation in dynamic noise environments. Techniques need to be explored to actively reduce the noise from the speech signal to bring ASR performance up to lower ambient noise performance.
- *Speaker modeling.* Many ASR systems require an extensive enrollment session, forcing the user to provide speech samples for as long as one hour or more. This is typical of the speech-to-text dictation systems on the market today. Additional research is needed to reduce or eliminate the

time required to perform speaker modeling to make the technology more acceptable by the user community.

- *Speaker tracking.* Current demonstrations have been limited to situations where only one person is using speech control with a single application. Additional research is needed to address issues of speaker deconfliction when multiple speakers are interacting with multiple computer application in a common work environment or when a user switches between tasks.
- *Untethered operations.* Freedom of movement and opportunistic use of workstations is an important aspect of the air operations environment. Current speech input systems are very limited in their ability to support this work pattern. New techniques need to be explored to provide high recognition accuracy rates under changes associated with an untethered, mobile user.

2.3 Natural Language

Dr. Alex Rudnicky developed an overview of the core technologies used in speech recognition. Acoustic and lexical modeling is the simplest technology and is statistics based. Today's language modeling is task based (for example, it is specific for command centers or cockpits). More sophisticated systems are understanding systems. These range from simple command and control systems to "open window" systems that do not require users to speak in a set vocabulary and syntax. Dialog management systems model the user, have conversational skills, and can clarify and explain information. Translingual communication enables users with different native languages to communicate.[22]

Natural language is not sufficient for expressing detailed technical information.[23] Pet Net notations can be used in these circumstances to reduce ambiguity. Natural language conventions can also be used—specifically, a prompt consisting of "a leading question, followed by a brief pause, and then a list of key words."[24] Other enhancements include word spotting (the capability to recognize key words) and barge-in (the capability to recognize commands spoken during a prompt).

Microsoft is researching natural language understanding using MindNet, semantics software.[25]

The Command Post of the Future (CPoF) is being developed to increase the speed and quality of command decisions through dynamically applied knowledge.[26] Goals are to increase the speed and quality of command decisions, to enable more effective dissemination of commands, and to enable smaller, more mobile command structures. The bottom line is to shorten the commander's decision cycle to stay ahead of the adversary's ability to react. The decision cycle includes situation assessment, course of action development, detailed planning, and execution. The commander should be more involved in the second step. Detailed planning should be automated. This would lead to recognition-primed decision making. The focus of the program is visualization and HCI. The program will tailor the available information to suit the commander's

[22] During information-gathering trip to Carnegie Mellon University, 17 March 1999.
[23] C. W. Johnson, J. C. McCarthy, and P. C. Wright, "Using a Formal Language to Support Natural Language in Accident Reports," *Ergonomics*, Vol. 38, No. 6 (1995), pp. 1264–1282.
[24] Douglas J. Brems, Michael D. Rabin, and Jill L. Wagget, "Using Natural Language Conventions in the User Interface Design of Automatic Speech Recognition Systems," *Human Factors*, Vol. 37, No. 2 (1995), p. 265.
[25] Information-gathering meeting at Microsoft, 16 April 1999.
[26] Information-gathering meeting at DARPA, 19 May 1999.

situation and decision process. The best of the current systems is the Global Command and Control System that enables only one level of aggregation and requires training and fusion. Visualization complexity is due to static, dynamic, sequenced, and storytelling requirements. There is also contextual complexity from data, device, user preference, user task, and team task. Stoplight technology risk assessments are decision-centered visualization (green), speech-gesture interaction (yellow), automatic generalization of visualization (yellow), and dialog management (red). A series of experiments will be run this year: asymmetric, guerilla, humanitarian for an urban disaster, peacekeeping, and sustained engagement. The contractors are Maya, VDI, Lockheed Martin, and the University of Maryland. The product is a visualization toolkit.

Murray Burke described the High-Performance Knowledge Bases (HPKB).[27] The program will enable the rapid development of large (100,000-axiom) knowledge bases, enabling a new level of intelligence in automated systems. In knowledge-based systems, knowledge is explicit rather than implicit. The typical structure of a large knowledge base comprises an upper ontology, core-theories, middle-level domain theories, and problem-specific theories. Knowledge bases are useful for general reasoning, optimized problem solving, and system integration. Encoding knowledge into logic is difficult. Size and complexity increase the number of interactions. The Cyc Knowledge Library has one million axioms but took 12 years to build. The HPKB will enable theory reuse and manipulation to enhance axiom development. Alphatech and IET are working battlespace reasoning and crisis understanding.

Teams of Science Applications International Corporation (SAIC), Stanford, SRI, ISI, Teknowledge, Cycorp, and Kestrel are working on integration approaches. Technology development includes knowledge servers and editing tools (by Cycorp, Stanford, SRI, University of Southern California [USC], and ISI), advanced knowledge representation and reasoning (by Kestrel, Northwestern, Stanford, and the University of Massachusetts), problem-solving methods (by Stanford, MIT, USC, and Edinburgh), and machine learning and language extraction (by CMU, SRI, MIT, Textwise, and GMU). Some of this work is being funded by the Air Force Office of Scientific Research (AFOSR). The challenge is extracting relationships. The problems are enemy workarounds, course-of-action critique, battlefield movement analysis, and crisis management. The sketch-understanding tool parses a sketch into knowledge. There is also a statement translator that converts structured English paragraphs into knowledge representation. Products are to demonstrate utility, feasibility, reusable library, and component reasoners. Rapid knowledge formation is the follow-on program to create axioms at 400 per hour. The program will include a knowledge entry associate. There is also a commonsense theory testing and conflict detection.

Information extraction (IE) is the ability to identify useful information in text and store it in a structured form like database records. IE capabilities are typically divided into two levels according to the complexity of the information they extract. Shallow extraction refers to extraction of simple information, such as entities (for example, the names of people, facilities, and locations), numerical information (such as monetary values and percentages), and simple events. Deep extraction refers to extraction of much more difficult information such as complex events (that is, scenario characterization). Because deep extraction for multiple topic domains is well beyond the state of the art, current AFRL research has focused on another layer of IE between shallow and deep extraction called “intermediate extraction.” The goal of the work is to produce technology that can reliably extract useful but less complex relationship and event information from free text to aid analysts in understanding unfolding situations. The ability to transform data (free text) into structured information is useful:

- To automate the processing of very large volumes of free-form text, saving time and labor
- To enable persistent storage of the extracted, labeled information in databases
- To enable information to be used by other user support tools (for example, analysis and visualization tools)

There are two basic approaches to developing IE systems: (1) the knowledge engineering (rule-based) approach and (2) the learning (statistical) approach. In the knowledge engineering approach, examining a large corpus of text discovers domain patterns, and rules are constructed by hand. This involves much time, labor, and skill. Skilled computational linguists can develop a good system in a reasonable amount of time, and rule-based systems still perform slightly better *for a single domain*. Furthermore, as new support tools are developed, it is anticipated that the time required to craft a rule-based system will continue to be reduced.

The learning approach to IE system design involves using statistical methods to automatically “learn” rules from annotated training data (text corpora annotated with the correct answers). These systems can be refined according to users’ corrections of system output. Because such systems are data driven, porting them to new domains is quicker and easier and requires less skill. Nevertheless, annotated training data can be difficult and expensive to acquire, and current-generation learning systems still do not perform quite as well as rule-based systems.

IE has obvious implications for data input and information capture for a JBI. It may also provide novel methods for query formation and, when coupled with a GUI, can facilitate understanding through information visualization. AFRL research has resulted in several developments important to maturing this technology. These include:

- A statistical parser to enable domain-independent shallow extraction
- A decision support system to aid text analysis and visualization for situation assessment that contains
  - Domain-independent shallow extraction of entities and simple events
  - Robust processing of diverse text inputs (that is, multiple text types)
  - Extraction of temporal and locative information enabling visualization of timelines and maps
  - A generic intelligence processor toolkit for message processing
  - Identifier automatic learning algorithms that perform shallow extraction of named entities from free text SANDKEY messages with 88 percent accuracy
  - IdentiTagger tool for annotating training data for automatic learning algorithms

---

[27] Information-gathering meeting at DARPA, 19 May 1999.

IE is a relatively young scientific and technology area. Current understanding supports shallow extraction for a single language with a precision of around 90 percent, provided that the problem has been scoped by an expert to a workable level. Precision levels for deep extraction in a single, narrowly defined domain are around 60 percent. Additional research is needed to provide better methods for co-reference resolution (that is, to standardize references to the same entity within the same document—for example, “Clinton,” “the President,” and “he”); to improve extraction precision; to improve the ability to cross-link among references and information domains; and to increase the IE processing speed to meet work demands.

CoGenTex has recently produced a domain-specific machine translation prototype system, TransLex, using several commercial off-the-shelf (COTS) products. A central feature of the system is a lexico-structural transfer method to handle cross-language semantic translation. This method provides a unified syntactic and semantic representation for each lexical item. By including syntactical features, the method can exploit the statistical techniques for analyzing and extracting information from corpora. Furthermore, IE avoids the labor-intensive activity of producing an interlingua. The current system provides English-to-French translation. The system architecture consists of a language parser, the core transfer unit, and a language generator. The transfer module includes an automatic bilingual lexicon extractor. Editing by a linguist, however, is still needed to complete the lexicon. Also, human intervention is required to convert the output of a COTS parser to a suitable format for the transfer module. The system has been demonstrated using a relatively small message set (500 messages). Further development is proceeding under a Phase II Small Business Innovative Research project.

NASA Goddard Space Flight Center has an Agent Technology Group working this area.[28]

2.4 Multimodal Interfaces

“It is widely believed that as the computing, communication, and display technologies progress even further, the existing HCI may become a bottleneck in the effective utilization of the available information flow; thus, in recent years, there has been a tremendous interest in introducing new modalities into HCI that will potentially resolve this interaction bottleneck.”[29] Various HCIs are presented in Figure 3.

[28] The group’s homepage is http://agents.gsfc.nasa.gov. In addition, an “Introduction to Agent Technology” briefing is located at http://groucho.gsfc.nasa.gov/Code_520/Code_522/Tech_Collab/teas/nov21/November.html.

[29] Rajeev Sharma, Vladimir I. Pavlovic, Thomas S. Huang, “Toward Multimodal Human-Computer Interface,” *Proceedings of the IEEE*, Vol. 86, No. 5 (1998), p. 853.

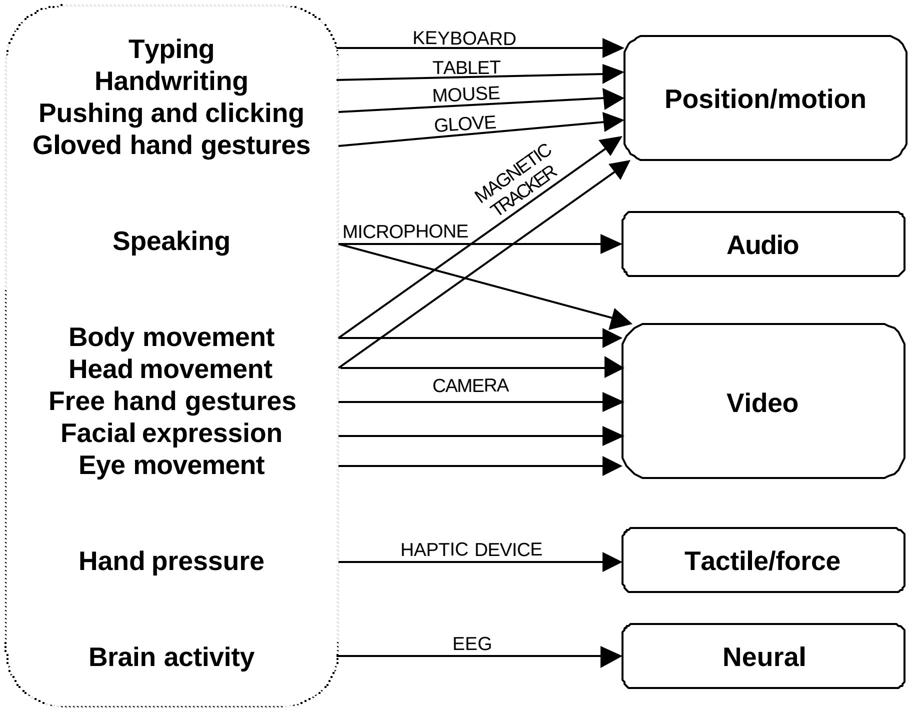

Figure 3. *Existing and near-term HCIs.*[30]

Multimodal interfaces involve the development of software libraries for incorporating multimodal input into HCIs. These libraries combine natural language and artificial intelligence techniques to provide the HCI with an intuitive mix of speech, gesture, gaze, and body motion. Interface designers will be able to use this software for both high- and low-level understanding of multimodal input and generation of the appropriate response.

One specific multimodal technology is the Intelligent Conversational Avatar. The purpose of this project is to develop an Expert System and Natural Language Parsing module to parse emotive expressions from textual input (see Figure 4). Another multimodal technology is GloveGRASP, a set of C++ class libraries that enables developers to add gesture recognition to their SGI applications (see Figure 5). Another technology is the Hand Motion Gesture Recognition System (HMRS). HMRS "is a project to develop a generic software package for hand motion recognition using hidden Markov models, with which user interface designers will be able to build a multimodal input system."[31]

[30] Rajeev Sharma, Vladimir I. Pavlovic, Thomas S. Huang, "Toward Multimodal Human-Computer Interface," *Proceedings of the IEEE*, Vol. 86, No. 5 (1998), p. 857.
[31] http://www.hitl.washington.edu/research/multimodal/.

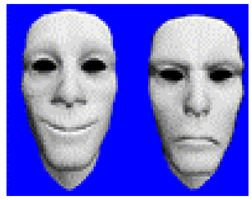

Figure 4. *Intelligent Conversational Avatar.*

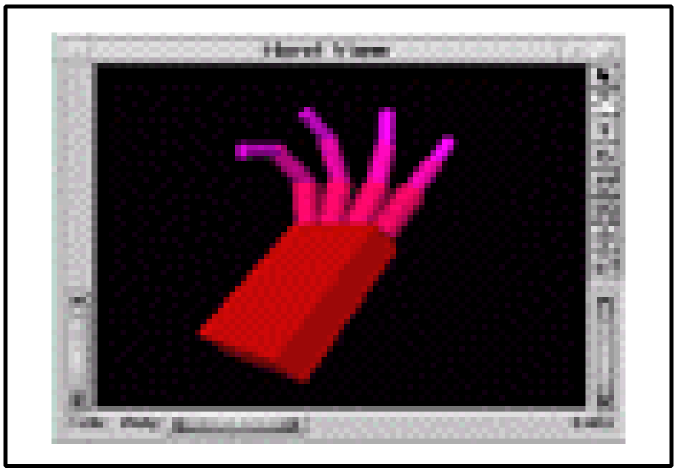

Figure 5. *GloveGRASP.*

Multiple interfaces can be implemented simultaneously and the sensed data fused. There are three levels of fusion. The lowest level, data fusion, is fusion of similar data. The second level, feature fusion, is the fusion of features from closely coupled but dissimilar data (for example, speech and lip movement). The third level, decision-level, is the fusion of decision actions from synchronized data (for example, typing while saying, “Bold”).

Systems that fuse gesture and speech already exist. MDScope, developed at the University of Illinois, is being used for visualizing biomolecular systems in structural biology. QuickSet uses pin gestures and speech to control military simulations in the form of Personal Digital Assistants. Jeanie is a calendar program that fuses pen gesture, speech, and handwriting. Visual Man fuses gesture, speech, and eye gaze to manipulate virtual objects. Finger-Painter fuses gestures and speech to control videos. Virtual-World fuses gestures and speech. The Artificial Life Interactive Video Environment interprets gestures and body movement in entertainment and training applications. Smart Rooms, Smart Desks, Smart Clothes, and Smart Cars fuse gestures and speech to perform butler services for users. Neuro Baby fuses speech and facial expressions to provide companionship to the user. Voice recognition and eye-tracking have also been fused for cockpit applications.[32] Implementation guidelines were developed for eye-voice interaction: (1) facilitate natural interaction, (2) minimize training requirements, (3) use eye point-of-gaze for deictic reference, (4) provide feedback on user commands, (5) provide feedback on visual

[32] F. Hatfield, E. A. Jenkins, and M.W. Jennings, *Eye/Voice Mission Planning Interface* (AL/CF-TR-1995-0204). Wright-Patterson Air Force Base, OH: Armstrong Laboratory, December 1995.

display object selection, and (6) use memory aids for speech input.[33] Additional work is being performed under NSF's Speech, Text, Image, and Multimedia Advanced Technology Effort.

Speech and head movements have also been fused for a hands-free HCI. "Head movement is interpreted to fulfill the positioning task for the mouse cursor while spoken commands serve for clicking, spoken hotkeys, and keyboard evaluation."[34]

To enhance accuracy of gesturing, gestures are rated in four areas[35]:

1. Space
   - Use
   - Indicate
   - Manipulate
   - Describe form
   - Describe function
   - Metaphor
2. Pathetic information
   - Emphasize
   - Maintain discourse
3. Symbols
   - Concept
   - Modifier
4. Emotion
   - Aroused
   - Enthusiastic
   - Happy
   - Relaxed
   - Quiet
   - Dull
   - Unhappy
   - Distressed

Gesturing can be used to enhance videoconferencing by providing mirroring and gesturing to remote sites. "Mirroring enables those at one site to visually coach those at a second site by pointing at locally referenceable objects in the scene reflected back to the second site."[36]

[33] F. Hatfield, E. A. Jenkins, M.W. Jennings, and G. Calhoun, "Principles and Guidelines for the Design of Eye/Voice Interaction Dialogs," IEEE Paper 0-8186-7493-8, 1996.
[34] R. Malkewitz, "Head Pointing and Speech Control as Hands-Free Interface to Desktop Computing," *ACM Conference on Assistive Technologies*, Vol. 3 (1998), p. 182.

Gesturing has also been tested for designating targets in fighter aircraft. The tests were conducted in a fixed-base simulator. Designation was performed using an ultrasonic hand tracker with either a proximity- or contact-cursor-aiding algorithm or a voice recognition system (VRS).[37] The fastest air-to-air target designations were associated with the voice recognition system; the second fastest with the proximity-cursor-aiding algorithm. The proximity-cursor-aiding algorithm, however, was also associated with the highest number of errors.

The JBI should be creating portals or interaction frames for the user to interact with information in the JBI. The interaction frames are likely to be some advanced visualization. The user should be able to directly interact through these visualizations, either through pointing or gesturing. If the commander wants to reposition some assets, he or she could do that using a gesture on the appropriate interaction frame. In the CPoF demonstration, ISX uses gestures to create sentinels—agents that inform the user when a certain situation changes.

Microsoft is researching (1) reasoning and intelligence (Bayesian inference to exploit knowledge bases), (2) natural language understanding (MindNet resolves work semantics), (3) speech (recognition, synthesis, and personalized voices), and (4) vision (gesture recognition, and head, eye, and body tracking). Microsoft is spending $2.9 billion in these areas to reduce the number of service calls.[38] Document abstracting is a critical development and a potential future product. Social interfaces, such as facial gesture recognition, are important for providing online support.

NASA is also developing gesture interpretation as part of the Virtual Spacetime program.[39]

Instead of a single multimodal device, there are complete smart rooms. The notion of a smart room is one that can actually observe the participants and infer various kinds of input by what it sees. For example, if a person makes a gesture across the screen, rather than relying on special display or pointer technology, a camera could recognize the gesture and create the same effect. This would create a totally unencumbered interaction with the user. An example of this type of technology is an MIT project called the Intelligent Room. MIT personnel are embedding computers in ordinary environments so that people can interact with them the way they do with other people—by speech, gesture, movement, affect, and context. Thus command staff members could interact with each other and also with the JBI seamlessly.

---

[35] Hummels and Stappers, "Meaningful Gestures for Human Computer Interaction: Beyond Hand Postures," *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition* (1998), p. 593.

[36] L. Conway and C. J. Cohen, "Video Mirroring and Iconic Gestures: Enhancing Basic Videophones to Provide Visual Coaching and Visual Control," *IEEE Transactions on Consumer Electronics*, Vol. 44, No. 2 (1998), pp. 388.

[37] T. J. Solz, J. M. Reising, T. Barry, and D. C. Hartsock, "Voice and Aided Hand Trackers to Designate Targets in 3-D Space," *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, 2734, 1996, pp. 2–11.

[38] Information-gathering meeting at Microsoft, 16 April 1999.

[39] A report describing their progress to date is available at http://science.nas.nasa.gov/Pubs/TechReports/RNReports/sbryson/RNR-92-009/RNR-spacetime.html.

2.5 Annotation

The Dutch have developed a voice annotation system for use in virtual environments.[40] These researchers define annotation as "generic units of information that are represented by a visual (3-D) marker."[41] The system was written in C using the Simple Virtual Environment library.

2.6 Drill-Down

Jaime Carbonell's information management manifesto demands getting the right information to the right people at the right time in the right languages and the right media at the right level of detail. To meet this manifesto, he has applied a number of drill-down technologies. These technologies include speech recognition, information retrieval (challenges include scaling up, increasing accuracy, retrieving novel information, retrieving information in languages other then the one used to develop the query, and searching different media), fact extraction (including topic detection and tracking), summary (traditionally just to develop an abstract, new research is applying user profiles to extract information that matches the user's interest), fusion (combining information from multiple sources into one document and presenting the reliability of the information), and translation (rapid development translation is used to develop a viable translation in 6 months rather than 6 years). Retrievals are based on relevance and novelty. Carbonell uses a spiral technique to move out from the information identified as the most relevant. Relevance is defined using Bayesian statistics and, therefore, must have previous usage data to generate. He also described a retrieval system being developed for the Army Intelligence community. It is based on segmentation, detection, and tracking of information in the open literature. Finally, he showed scores from a contest of summary schemes. The score was the number of questions answered correctly based on the summary versus reading the entire document (90 percent). CMU had the best summary with about 75 percent correct. There is also a time-series synthesis that generates a timeline summary that presents only novel information. Carbonell warned about model drift—the further away in time the model is from the event, the less accurate the prediction. Models must be always maintained.[42]

Andrew Moore (Center for Automated Learning and Discovery) described data mining work. The purpose of data mining is to develop rules to predict future events. First-generation algorithms include regression, neural nets, and decision trees. The next-generation will (1) learn across fully available mixed-media data, (2) learn across multiple internal databases plus web and news feeds, (3) learn actively by closing the experiment-hypothesis loop, and (4) most important, learn decisions rather than predictions. All dimension trees (ADtrees) is a technology being applied to provide a 100-fold enhancement of the search. There are currently 11 industry partners applying the CMU data mining tools. Moore described the Sloan Sky Survey that will provide a database of 500 million galaxies, each having 500 attributes being recorded. There are clustering technologies that are being applied to these data and enhancing computational speed

---

[40] J. C. Verlinden, J. D. Bolter, and C. van der Mast, *Virtual Annotation: Verbal Communication in Virtual Reality* (PB95215505). Delft, Netherlands: Delft University of Technology, 1993.
[41] Ibid., Section 2, Paragraph 1.
[42] Based on a presentation to the SAB Interact panel at Carnegie Mellon University, 17 March 1999.

for efficient search. Other applications include drug discovery databases, 3M process data, Caterpillar's parts inventory management, and cell chip analysis data streams. Data mining has also been used for anomaly detection across a wide range of applications, including pricing of electrical power, managing heating in a building, applying and drying paint, and detecting vibration changes in engines. Moore believes that the future strength of data mining systems is integration with decision application tools.[43]

Informedia is a system developed to automatically extract information from video to enable full-content search. This requires speech recognition, image extraction, and natural language interpretation. Problems include the inability to scale up the original algorithms; developing paragraphing for image understanding based on face, text, and objects; inaccuracies; and ambiguities. The real advantage of Informedia is the ability to combine technologies: scene changes, camera motion, face detection, text detection, word relevance, and audio level.[44]

Informedia has a search and retrieval capability. It can present results with color coding, indicating the relevance of the video on each of the query words. A text transcription can be presented as well as a video clip and a filmstrip. A map can be automatically brought up on the basis of the speech recognition and indicate dynamics, for example, movements of people and goods over distance. The map can also be used to issue a query. In addition, Informedia can also translate from and into Spanish. A skim, a shortened version of a video, is also available. It is automatically generated to provide the most information in the least amount of time. It is also possible to cut the video for insertion into PowerPoint or Word documents. Finally, a face search is available. There is a commercial product, FACE IT, which does this as well.[45]

Microsoft predicts that computers will be used for understanding, learning, communicating, consuming, and entertaining. Research goals are generating life-like speech from textual data, artificial singing, analyzing language, and developing user agents.[46] These agents monitor events to provide help (Lumiere in Office 97), infer operational needs from browsing (implicit queries), and monitoring mail to autocategorize it (Lookout SpamKiller). Research on user interfaces is on telepresence, speech, vision, graphics, and wizards. One outgrowth is TerraServer to scale up to big databases, including images such as in the Image Pyramid.[47] Another is data mining, defined as finding interesting structure (patterns or relationships) in databases.[48] Tasks in data mining include prediction, segmentation, dependency modeling, summarizing, and trend and change detection and modeling. Scalability is a critical research area since it is assumed that there will be trillions of clients (devices, doors, rooms, cars, etc.).[49]

[43] Based on a presentation to the SAB Interact panel at Carnegie Mellon University, 17 March 1999.
[44] Ibid.
[45] Ibid.
[46] Information-gathering meeting at Microsoft, 16 April 1999.
[47] See http://terraserver.microsoft.com.
[48] See http://research.microsoft.com/~fayyad.
[49] Information-gathering meeting at Microsoft, 16 April 1999.

Speed is a critical issue in data mining. Algorithms have been developed that improve the identification of patterns in the data more 100 times faster than the brute-force method.[50] Other issues include the size of the database, nonsystematic errors, handling of null values, incomplete or redundant data, and changing database contents.[51] One method to enhance speed is to use counterfactuals to generate information. This is done by asking, "What patterns match these data?"[52]

Howard Wactlar (CMU) and Kathy McKeown and Judith Clavins (Columbia) are working on multitext fusion. They are developing summary algorithms for medical records. Scin Chin (University of Arizona) is developing automatic analysis systems to identify the key themes in text and electronic conversations.[53]

Another tool is InfoSleuth, a consortial project carried out by MCC on behalf of Raytheon Company, General Dynamics Information Systems, Inc., SAIC , NCR Corporation, TRW, Inc., Schlumberger Limited, and Rafael. According to its website, "InfoSleuth implements a community of cooperating agents that discovers, integrates and presents information on behalf of a user or application, for which it provides a simple, consistent interface. The information it accesses is distributed and heterogeneous, for example the types of information available through an intranet in a large corporation or on the World Wide Web."[54]

Another form of drill-down is provided by zooming user interfaces (ZUIs) that present information graphically. Users zoom out to gain context and zoom in to focus on the information of interest. Currently available ZUIs include Information Visualizer (Xerox PARC), Perspecta, Merzcom, Tabula Rasa (New York University), and Pad++ (University of New Mexico). Boderson and Meyer listed the following requirements for ZUIs: "maintain and render at least 20,000 objects with smooth interaction,… animate all transitions, use off-the-shelf hardware,… support high quality 2-D graphics,… provide rapid prototyping facility,… support rich dynamics,… support rich navigation metaphors,… support standard GUI widgets,… offer a framework for handling events,… run within existing windowing and operating system."[55]

The Joint Force Air Component Commander[56] focuses on agile (efficient, responsive, timely, broad-spectrum) and stable control of military operations. Formerly it focused on making operational planning effective, efficient, flexible, and responsive. A lesson learned was that managing the dynamics of decision making in a rapidly changing environment requires more than planning. Understanding downstream effects is critical. Further, "bad" dynamics behavior includes failure, slow rise, fast rise, overshoot, undershoot, thrashing, and oscillations in reacting

[50] C. Chang, J. Wang, and R. K. Chang, "Scientific Data Mining: A Case Study," *International Journal of Software Engineering and Knowledge Engineering*, Vol. 8, No. 1 (1998), pp. 77–96.
[51] Vijay V. Raghavan, , Jitender S. Deogun, and Hayri Sever, "Introduction," *Journal of the American Society for Information Science*, Special Topic Issue: Knowledge Discovery and Data Mining, Vol. 49, No. 5 (1998), pp. 397–402.
[52] V. Dhar, "Data Mining in Finance: Using Counterfactuals to Generate Knowledge From Organizational Information Systems," *Information Systems*, Vol. 23, No. 7 (1998), p. 432.
[53] Information-gathering meeting at NSF, 19 May 1999.
[54] See http://www.mcc.com/projects/infosleuth.
[55] Boderson and Meyer, 1998, pp. 1104–1105.
[56] Information-gathering meeting at DARPA, 19 May 1999.

to the enemy. Experiments will be conducted to evaluate the effects of changes in timing, process, and structure on planning, controlling, planting, assessing, and observing. Command and control is difficult because systems are large, distributed, and dynamic, have uncertain and limited information, and require humans to be in the decision loop.

The purpose of the Advanced Intelligence, Surveillance, and Reconnaissance (ISR) Management program[57] is to optimize ISR support to the dynamic battlefield. Currently, information needs are not well linked to ISR activity, due to limited flexibility, stovepiped organizations, and assets optimized for technical performance. The project goals are (1) integration between ISR, operations, and support activities, (2) dynamic in-time response to operational timelines, changing priorities, and environments, (3) end-to-end management of the ISR process, and (4) optimization of assets to maximize satisfaction of information needs. The missing technologies are determination of the actual intelligence needs, optimization of math, and integration of enabling technologies. The approach consists of dynamic processes to integrate operations, logistics, and ISR; global optimization of ISR confederation to provide maximum information support; responsiveness to operational timelines, changing priorities, and environments; and continuous execution supporting reaction and recovery. Proactive ISR operational support is the interpretation of the commander’s vision. Technology challenges include global optimization of ISR with uncertainties (threat location, probability of intercept, intentions, and execution of strategy), timeliness (real-time control of ISR assets, exploitation of opportunistic collections), and situational awareness (interpret the commander’s vision, operational plan, and current situation; correlate ISR support; and predict the future situation to anticipate needs). The architecture includes four components: asset allocation, strategy, information, and workflow. A critical technology being developed is agent-based workflow management (Smart Workflow for ISR Management). Context wrapping is an infrastructure tool that assesses the capabilities of each of the four components—for example, who will use this and what they will do. The goal is information-needs generation in less than one hour. The program was at Joint Expeditionary Force Experiment 99 as a Category 3 initiative and will be completed in FY02. The pieces are available for demonstration at the Technology Integration Center.

The Advanced Research Projects Agency Rome Planning Initiative (ARPI) research and development process is driven by a series of Integrated Feasibility Demonstrations (IFDs) and Technology Integration Experiments (TIEs), which assess technical progress and evaluate its operational impact. A Common Prototyping Environment (CPE) was developed during Phases I & II and an Air Campaign Planning Tool Testbed Environment during Phase III to support demonstration and testing of technologies; experimental system integration and evaluation activities; and re-use of databases, knowledge bases, software modules, and test scenarios. *Tier 1* includes a number of independent research projects that are oriented toward developing operationally focused knowledge-based reasoning technology that addresses critical problems in military planning and scheduling. The exit criteria for technology migration to *Tier 2* include successful demonstration of capabilities in research oriented TIEs. The *Tier 2* effort consists of

[57] Information-gathering meeting at DARPA, 19 May 1999.

the IFDs and CPE-supported activities, which evaluate technical progress, merge the individual developments in *Tier 1* into experimental systems and rapid prototypes, and integrate ARPI-developed technologies with other components to address specific operational problems. An IFD shows the operational communities new planning and scheduling capabilities to obtain constructive feedback on their applicability to critical operational functionality measured against criteria for success defined by end users. *Tier 3* involves the user-guided insertion of ARPI technology and systems into user-supported operational prototypes.

Critical to the success of the program was the definition of functional requirements that were dependent on the ARPI development of "missing technology." Some specific correlations between operational planning requirements and knowledge-based scheduling and planning technology are shown in Table 1.

Table 1. *Operational Planning Requirements and Knowledge-Based Scheduling and Planning Technology*

Artificial Intelligence/ Planning Research World	**Joint Planning and Operations World**
Generative planning	Commander's objectives, concept of operations, force/resource selection and reuse, objectives and task decomposition
Constraint-based planning	Resource constraint analysis, feasibility analysis, time-phasing, etc.
Case-based planning	Force analysis, planning Module library Development, failure analysis, plan revision techniques
Intelligent and object-oriented databases	Distributed, heterogeneous intelligence and situation assessment databases
Interactive graphics and editing of timelines, schedules, resources, maps, graphs, representations, etc.	Manual data analysis, plan refinement, and briefing Production

2.7 Personal Computing Devices

Personal computing devices include two-handed interfaces, laptops and palmtops, 4-D mouse, and wearable interfaces.

2.7.1 Two-Handed User Interfaces

The goal of the field analyzer project is to develop a two-handed user interface to the stereoscopic field analyzer, an interactive 3-D scientific visualization system. The stereoscopic field analyzer displays scalar and vector fields represented as a volume of glyphs, and the user manipulates the graphical representation using two Polhemus Fastrak sensors with three buttons each (see Figure 6). Each hand holds one sensor and has a distinct role to play, with the left hand responsible for context setting of various kinds, and the right hand responsible for picking and fine manipulation.[58]

Figure 6. *Field analyzer.*

2.7.2 Laptops and Palmtops

Small computers range from laptops to palmtops to pen-based computers. Current laptops are quite impressive in terms of computing power and display capability. One pays the price in terms of power and weight, and most still require a large amount of keyboard entry. Palmtops, like the Toshiba Libretto, are certainly a lot less bulky, but they are display disadvantaged. Pen-based computers, like the Fujitsu 2300, are an interesting middle ground, in that they provide better display capability than the palmtops, have the potential for much more natural interaction with the pen and are less bulky than a typical laptop. The biggest problem is that the operating systems are not quite ready for pen-based computing. It is expected that the JBI will have users with all of these classes of computers and that there will need to be facilities to support the wide range of input and output devices.

2.7.3 4-D Mouse

The emergence of 3-D software is creating opportunities for new input devices that offer 3-D specific features and controls and yet are easy to learn and use. Wacom is currently selling graphics tablets into traditional (2-D) graphics market. With the new 4-D Mouse (see Figure 7), it is seeking to capitalize on the rapid expansion of commercial applications in the 3-D graphics market. Commercially available 3-D input devices fail to meet the needs of the rapidly expanding general 3-D design market because they are too expensive, require substantial practice, and are dedicated to only 3-D tasks.

[58] See http://www.hitl.washington.edu/research/sfa/.

Figure 7. *4-D mouse*.[59]

2.7.4 Wearable Interfaces

Wearables typically comprise a belt or backpack personal computer, a see-through or see-around head-mounted display (see Figure 8), wireless communications hardware, and an input device such as a touchpad or a chording keyboard. A crucial question is how to best use these devices to create an intuitive interface for the user.

Figure 8. *See-through head-mounted displays*.

The Human Interface Technology Lab has been experimenting with body-stabilized spatial information displays on a wearable platform and has found that users perform faster with a spatial display and are able to remember more displayed information.[60]

[59] See http://www.hitl.washington.edu/research/4dmouse/.
[60] See http://www.hitl.washington.edu/research/wearint/.

2.8 Automatic Data Capture

Point-of-use data capture is critical to the information support and administrative portion of the JBI. Data are generated in a great many locations, but they are not always captured in a usable form and shared with those who need them. Examples include administrative data such as health care and financial records, maintenance data collected during inspections or generated by embedded information systems in an aircraft, and data that show the expenditure of consumables such as fuel and ammunition. With the emerging availability of wearable computers and wireless technology, many data can be captured at the point where they are generated and automatically become available to the JBI. The challenge is not the technology itself, but to develop the business rules and processes, which need to be automated.

2.8.1 Barcoding

Barcoding technology is widely recognized as a means of maintaining inventory control—that is, ensuring that the right materials are ordered at the right time and in the right quantities. Barcodes aid in inventory control as they facilitate real-time updating of automated systems, enable accurate inventory levels (and hence, accurate reorder points), and enable flexible serial or lot tracking of materials in transit (see Figure 9 for a typical use of bar coding).

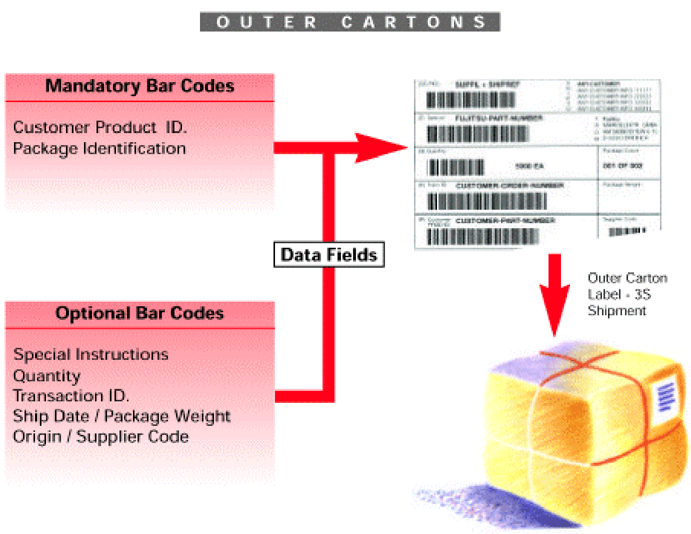

Figure 9. *Example use of barcode technology.*

2.8.2 Smart Tags

A related technology with many possible military adaptations is an application called Automatic Vehicle Identification (AVI) or “tagging.” This refers to the components and processes by which toll-collection equipment can determine ownership of the vehicle for the purpose of charging the toll to the proper customer. AVI technology can be broken into two main categories: laser and radio frequency (RF). Laser systems use a barcoded sticker attached to the vehicle (often on the driver’s side rear window), which is read by a laser scanner as the vehicle passes through the toll lane. Basically, laser systems operate in a manner similar to grocery store checkout scanners. RF systems use a transponder (tag) mounted either on the vehicle's bumper or inside the windshield or roof; the tag is read by a RF reader/antenna. Laser technology has several drawbacks that limit its use in the toll-collection environment, especially in an open-road system. Chief among these drawbacks are ease of forgery and the system's sensitivity to weather and dirt. In addition, the laser scanner is limited in the distance it can be placed from the vehicle. RF technology overcomes these limitations and as such is proving to be the AVI technology of choice for new systems. In addition to toll collection, some types of RF tags are also being used for vehicle-to-roadside communications. This technology allows a tag equipped with some form of readout to inform the driver of traffic conditions. There are three main RF technologies that are either in use today or undergoing extensive trials: RF tags, RF smart tags, and smart cards with RF transponders.[61]

2.8.3 Global Positioning System Locators

The last technology in this section is the Global Positioning System (GPS) locator. This has the ability not only to tell about itself, but also to provide its location. This is another example of a datum that can be supplied automatically rather than burdening the user.

[61] See http://www.ettm.com/avi.html.

Chapter 3: Presentation

Presentation is the medium and format of information that is input to or output from the JBI. Presentation technologies include personal display devices, data visualization, 3-D audio, and tailoring. Each of these technologies is described below.

3.1 Personal Display Devices

Personal display devices include virtual retinal displays, tactile vests, and haptic interfaces.

3.1.1 Virtual Retinal Display

The virtual retinal display team has focused on developing improvements to the current prototype systems and on creating the parts needed for future prototypes. The virtual retinal display, based on the concept of scanning an image directly on the retina of the viewer's eye, was invented at the Human Interface Technology Lab in 1991. The development program began in November 1993 with the goal of producing a virtual display with full color, a wide field of view, high resolution, high brightness, and low cost.

Two prototype systems are being demonstrated. The first is a bench-mounted unit that displays a full-color, video graphics array (VGA) image with 640 × by × 480 resolution updated at 60 Hertz. It operates in an inclusive or a see-through mode. The second system is a portable unit, displaying a monochrome, VGA-resolution image. The portable system is housed in a briefcase, allowing for system demonstrations at remote locations (see Figure 10).

Figure 10. *Virtual retinal display.*

The largest component in the portable system is the commercially purchased vertical (60 Hertz) scanner. A new vertical scanner being designed should significantly decrease the device's size and cost. Once this design is complete, a head-mounted demonstration prototype will be assembled.

Commercial applications of the virtual retinal display are being developed at Microvision Inc.[62]

3.1.2 Tactile Vest

A specific example of a haptic interface is a tactile vest. This could provide another way of cueing the user to look in a particular direction by simulating a tap on the shoulder or another kinesthetic cue.

3.1.3 Haptic Interfaces

Haptic (force-reflecting) interfaces can provide useful kinesthetic information in virtual environments. Several haptic interfaces are already in use. One is a 4–degree-of-freedom controller used to train surgeons.[63] Even more state of the art is a type of haptic interface developed by CMU researchers. The CMU haptic display uses magnetic levitation to physically interact with simulated objects and environments on computer screens. The device is unique because it enables people not only to touch these objects, but to reach in and manipulate them in three dimensions as well. S. Mascaro and H. H. Asada at MIT[64] have developed a fingernail sensor to enable the user to maintain tactile sensitivity while activating virtual switches. The sensor works by detecting color changes in the fingernail. It can control virtual switches that are metallic plates placed anywhere.[65] It uses wireless communication.

Stephen Ellis[66] is studying the phenomenology associated with immersive visual technologies. He demonstrated a stereo helmet-mounted display used to view aircraft landing at the Dallas–Fort Worth airport. The phenomenology includes latency, haptic feedback, and correction of sensor anomalies. Steve stressed that he solved the problems of speed, accuracy, and attraction.

A virtually augmented cockpit (advanced head-down, head-up, and helmet-mounted displays, 3-D audio, and haptic displays) resulted in mission performance in a simulated air combat mission.[67]

Data gloves sense the location, force, and position of each fingertip. The gloves can be used to coordinate human and robot efforts.

---

[62] See http://www.hitl.washington.edu/research/vrd/.

[63] R. Baumann and R. Clavel, "Haptic Interface for Virtual Reality Based Minimally Invasive Surgery Simulation," *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Vol. 1 (1998), pp. 381-386.

[64] S. Mascaro and H. H. Asada, "Hand-in-Glove Human-Machine Interface and Interactive Control: Task Process Modeling Using Dual Petri Nets," *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Vol. 2 (1998), pp. 1289–1295.

[65] See http:www.interact.nsf.gov/cise/html/sfpr?OpenDocument.

[66] Information-gathering meeting, 15 April 1999.

[67] M. W. Haas, S. L. Beyer, L. B. Dennis, B. J. Brickman, M. M. Roe, W. T. Nelson, D. B. Snyder, A. L. Dixon, and R. L. Shaw, *An Evaluation of Advanced Multisensory Display Contents for Use in Future Tactical Aircraft* (AL/CF-TR-1997-0049). Wright-Patterson Air Force Base, OH: Armstrong Laboratory, March 1997.

This coordination is useful when[68]

- [The] human has limited knowledge about the process as well as the functionality of machines
- Human error must be detected and corrected
- High safety standards must be maintained, although humans and machines work closely
- Human actions must be recorded together with the machine's action
- Humans are unable to provide detailed commands to the machines

3.2 Data Visualization

Design guidelines for 3-D visualization displays are being developed at the University of Toronto.[69] Based on performance of a path-tracing task, these authors recommend (1) 3-D rather than 2-D displays and (2) combined rotational and stereoscopic displays enhanced with multiple static viewing displays.

The goal of the Joint Logistics Advanced Concept Technology Demonstration[70] is to develop and integrate web-based logistics joint decision-support tools for the Global Combat Support System. Every 6 months there is a demonstration. The user requirements were (1) force capability assessment, (2) support concept generation and evaluation, (3) distribution, material management, and maintenance analysis, and (4) visualization. Not being addressed due to funding constraints are consumption planning and analysis and reconstitution. Joint decision-support tools were developed to support capabilities that users needed-force browser, ground logistics, joint electronic battlebook, air logistics, data mediation, visualization, and capabilities assessment. The core requirements that they be web-based, have a common look and feel, and support any echelon, collaboration, phase of the campaign, real data, box, visualization, tool integration, and user involvement. The second exercise was a large-deployment scenario: 40,000 line items and actual Time-Phased Force Deployment Data (TPFDD). The following were demonstrated: user interface, account management, common look and feel, dispersed sites, live data, collaboration (MITRE collaboration tool), and mapping (Joint Mapping Toolkit compatible but not integrated). In the next 6 months, Distribution Material Management and Maintenance Analysis sustainment visibility will be added. There will also be TPFDD collaboration, infrastructure for ports and airfields; there will be Transportation Coordinator's Automated Information for Movement System II for actuals on units, people, and things; and there was Foal Eagle 99. There is a follow-on: the Joint Theater Logistics Advanced Concept Technology Demonstration run by the Defense Information Systems Agency as the executive agent. Its purpose is to meld operations and logistics more closely. Current issues are network stability, user accounts, data mediation, Navy and Air Force Joint Total Asset Visibility data, equipment condition codes, and actual movement data at the item level.

---

[68] S. Mascaro and H. H. Asada, "Hand-in-Glove Human-Machine Interface and Interactive Control: Task Process Modeling Using Dual Petri Nets," *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Vol. 2 (1998), p. 1295.

[69] R. L. Sollenberger and P. Milgram, "Effects of Stereoscopic and Rotational Displays in a Three-Dimensional Path-Tracing Task," *Human Factors*, Vol. 35, No. 3 (1993), pp. 483–499.

[70] Information-gathering meeting at DARPA, 19 May 1999.

Autometric's[71] focus is diverse imagery analysis from mapping the moon to making movies to supporting intelligence gathering. Autometric has developed mensuration programs that are being used in missiles. Its Enhanced Geo Data Environment (EDGE) products are used for whole-earth visualization, modeling, and simulation. EDGE is a 4-D, whole-earth visualization with drag-and-drop of weather and environment to look from anywhere to anywhere with a settable clock to move through a model at any speed. Data can be fused from several sources. EDGE is written in C++. Once imported, the data are automatically georeferenced. Two data sets of differing resolution can be overlaid. The highest resolution is always on top. Imagery can be manipulated to control opacity, brightness, contrast, red threshold, green threshold, and blue threshold. The map manager lets users overlay imagery on maps. Outlines of political borders can be overlaid as well as features such as lakes, rivers, roads, and railroads. A folder manager contains all the materials that were used for a geospatial location. Annotations can be added. A spatial query server pulls only the data that are of interest to the user. The operators make a circle or rectangle to mark data and search for vectors such as parcels, blocks, schools, and zoning. Assessed home costs and taxes can be listed for the queried location. Color-coding can be used to indicate the age of the data set. Sullivan then demonstrated a 3-D image. Sites can be identified and the attributes of each-for example, missile battery capability-stored. Elevation data of any resolution can be added. Overhead sensor flight paths and area of regard can be shown. There is an animation capability to show when sensors will be over a particular site. Intervisibility analysis can be performed. ATOs can be visualized and reviewed. Voice can be used with the NT version of EDGE. Future enhancements will add attributes to objects for more advanced querying and performance models for moving items such as aircraft.

The Joint Multi-dimensional Education and Analysis System is used to provide 3-D visualization at the National Defense University to support crisis management training. The system is a web browser tied to EDGE. Pages were generated from the course syllabus and tied to the geospatial portion of EDGE. EDGE can be used with a 3-D light table.

The Mission Familiarization VR Program visualizes regions (typical EDGE), compounds (tied to blueprint-like data and sensor information, such as surveillance cameras), and buildings (using digital pictures to create a coherent image of walking through a building).

An NT version of EDGE is available and is called Merlin. It added 3-D to Raytheon's correlation data. There is a Netscape plug-in (an Internet Explorer version is also available) that will be used on the Central Intelligence Agency's (CIA) World Factbook and the National Imagery and Mapping Agency's Imagery for Citizens websites to enable users to obtain data by selecting from a globe rather than from an alphabetical list of countries.

In Takeo Kanade's work on Virtualized Reality™,[72] television can give a view into another part of the real world, such as a sporting event. This capability is great, but each viewer gets the same view, whether they want it or not, and none of the viewers has any power to control that viewpoint. In contrast, VR immerses viewers in virtual worlds even though their bodies are still

[71] Information-gathering meeting at DARPA, 20 May 1999.
[72] See http://www.cs.cmu.edu/afs/cs/project/VirtualizedR/www/VirtualizedR.html.

in the real world. Each viewer moves independently and freely throughout this world, allowing people to see events from their own viewpoint. VR, though, has focused on creating purely virtual worlds that do not correspond to anything in the real world. In addition, typical virtual worlds look too artificial to convince viewers that they are in another part of the real world.

Takeo Kanade's work combines the technology behind television, VR, and Computer Vision to create virtual models of real-world events—what he calls Virtualized Reality dynamic event models. These models can be used to construct views of the real events from nearly any viewpoint, without interfering with the events. Like VR, Virtualized Reality dynamic event models enable viewers to see whatever they want to, but unlike VR, this "other world" is actually a real event, and the views of this event are photorealistic.

NASA Ames is developing a realistic auditory environment for virtual displays (see Figure 11). Durand R. Begault et al. have defined "*auditory presence* to mean the ability to subjectively convince the user of their presence in an auditory environment. On the other hand, *auditory virtualization* refers to the ability to simulate an acoustic environment such that performance by the listener is indistinguishable from their performance in the real world." Judgments of visual display quality are enhanced with higher-quality auditory displays, according to Russell Storms, a doctoral candidate at the Naval Post Graduate School.

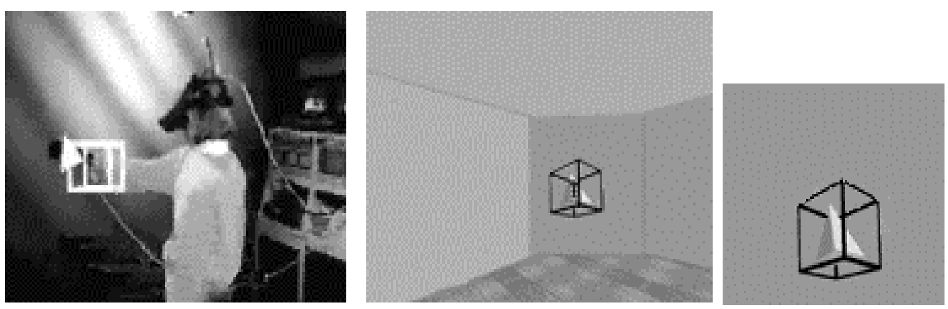

Figure 11. *Left: a subject performing 3-D tracking by attempting to keep the tetrahedron inside the moving cube. Center: the subject's actual view through the head-mounted display is represented by a screen image. Right: a closer view.* [73]

Steve Roth[74] developed an information-centric approach to visualization. He is a member of the CMU Robotics Department and head of MAYA Viz, which produces custom visual interfaces. Roth identified the need to support visualization: integration; for tolerance for unpredictable user need for information; for user control of scope, focus, and level of detail; and for collaboration and coordination. Roth helped develop three systems: the System for Automatic Graphic Expression, an automated visualization design; AutoBrief, an automated multimedia presentation (text and graphics); and Visage, an information-centered user interface approach that makes System for Automatic Graphic Expression graphics dynamics. In Visage, every item is an object that can be manipulated. The tool also supports an interactive slide show as well as the ability to filter based on object attributes.

[73] Durand R. Begault et al., 1998.
[74] Information-gathering meeting at Carnegie Mellon University, 17 March 1999.

MITRE developed the Hyperspace Structure Visualization tool, a navigation mechanism in which the user is able to view a hierarchy of the browse space. The left-hand side of Figure 12 displays the traditional HTML layout of a web page; the right-hand side displays the Hyperspace Structure Visualization tool, a hierarchical, automatically generated, navigable view.

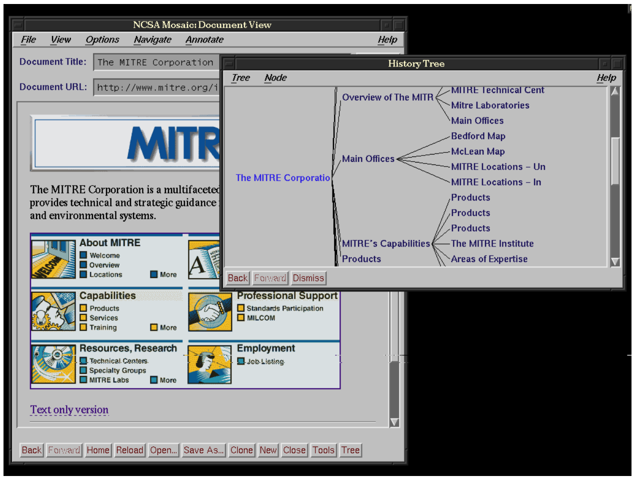

Figure 12. *Hyperspace structure visualization*.[75]

MITRE's "Geospatial News on Demand Environment" is a Geographic Visualization for searching georeferenced data. A sample is presented in Figure 13.

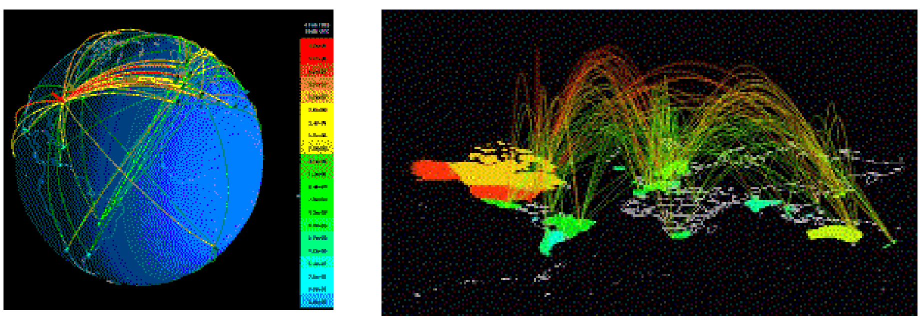

Figure 13. *Visualization of geospatial relationships*.[76]

A third MITRE visualization tool is Multisource Integrated Information Analysis, developed to display sensor and battlefield information. A sample display is presented in Figure 14, in which the *x*–*y* dimension shows the coordinates of a geospatial area, and the *y* coordinate displays time.

---

[75] N. Gershon, "Moving Happily Through the World Wide Web," *IEEE Computer Graphics and Applications*, Vol. 16, No. 2 (1966), pp. 72–75.
[76] Chase et al., 1999.

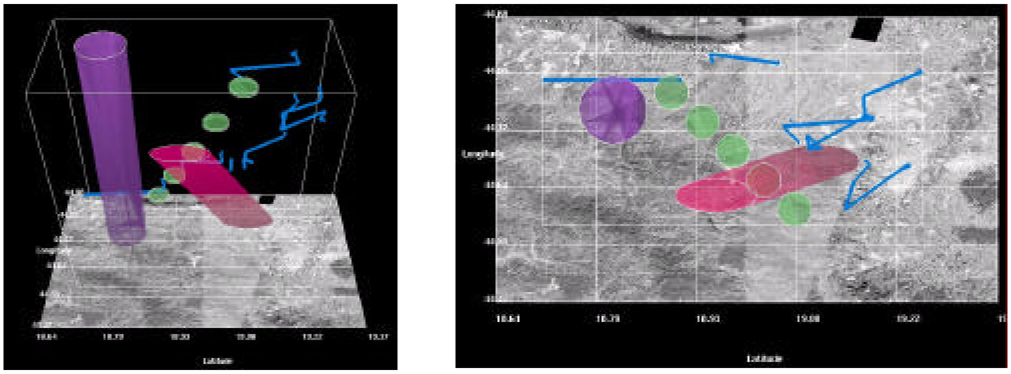

Figure 14. *Sensor coverage visualization.*[77]

A fourth MITRE visualization tool is the Collaborative Omniscient Sandtable Metaphor—a digital sandtable (see Figure 15).[78]

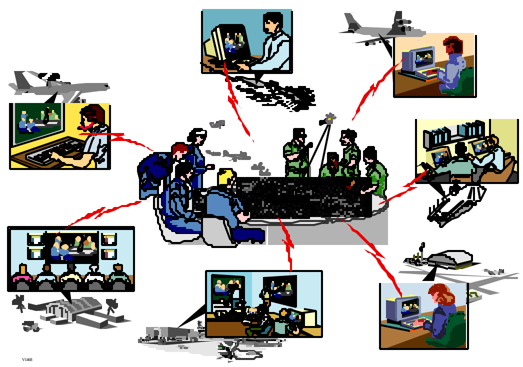

Figure 15. *Conceptual view of the Collaborative Omniscient Sandtable Metaphor.*[79]

3.2.1 Holographic or 3-D Displays

Since much of what the user is going to interact with is geospatial, it makes sense that in particular situations, the most effective way to present the information is in a true 3-D environment. Currently this has the disadvantage of requiring special glasses or environments (see Figure 16), but research is gradually reducing the encumbrances required for this type of interaction.

---

[77] Chase et al., 1999.

[78] An example of an air traffic simulator that includes a head-mounted display in use (see Figure 16) is located at http://duchamp.arc.nasa.gov/research/AOS_currentplan.html. Additional information is available in two papers: Ellis, "Sensor Spatial Distortion, Visual Latency, and Update Rate Effects on 3-D Tracking in Virtual Environments"; and Durand R. Begault, S. R. Ellis, and E. M. Wenzel, "Headphone and Head-Mounted Visual Displays for Virtual Environments," invited paper, 15th International Conference of the Audio Engineering Society (dbegault@mail.arc.nasa.gov or db@eos.arc.nasa.gov).

[79] Chase et al., 1999.

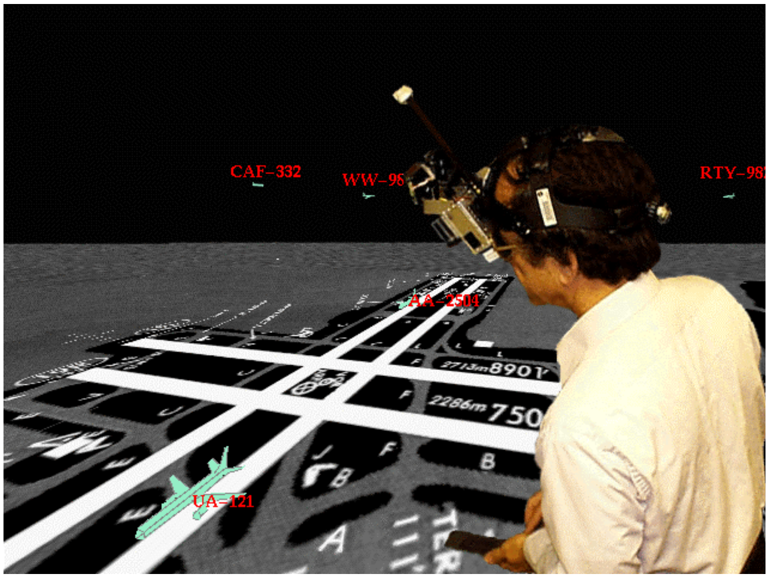

Figure 16. *Air traffic control simulator using a head-mounted display*.

3.2.2 Large-Screen Displays

Battle commanders need to see all relevant information with clarity, speed, interactivity, and organization. To date the display systems have been one of the bottlenecks in the information channel to the user. With large-screen display technology such as Sarnoff's System Technology for Advanced Resolution or the AFRL Data Wall, scalable, interactive display technology that will support very large (20-foot) display surfaces with hundreds of megapixels will soon be available. This will enable multiple people to interact with maps and other information displays very naturally. The displays could be used for small-group problem solving or small or large group briefings. A key part of the notion is that a workstation is not tied to a particular element of the screen. The notion is that displays could be arbitrarily positioned.

3.3 3-D Audio

3-D audio depends on sound localization—"the ability to identify the position of a sound source in space."[80] Position of sound can be determined using binaural cues—specifically, sound reaching the ears at different times (as much as 700 µsec) and intensities—as much as 40 decibels (dB). Sounds 20 or 120 degrees from straight ahead have the greatest intensity difference. Localization of sounds below 2,000 Hz is based primarily on time differences, above 400 Hz on intensity differences. Localization is poorest at 2,000 and 4,000 Hz. Phase cues are also used for localization of periodic sounds but only if successive cycles are at least 1,600 Hz apart. Rise times of 100 milleseconds (ms) or more also aid in sound localization. In addition, head movements provide cues but only for sounds with a duration greater than 250 ms. Monaural cues (sound shadowing by the head and loudness changes of moving sounds) are also used in sound localization. Sound intensity is the primary cue for distance to a sound source.

[80] Kenneth R. Boff and Janet E. Lincoln, eds., *Engineering Data Compendium: Human Perception and Performance.* Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory, 1988, p. 672.

S. T. Pope and L. E. Fahlen used the model in Table 2 to map sound features to spatial cues in a virtual environment.[81]

Table 2. *Mapping Between Sound Features and Spatial Cues*

Feature	**Cue**
Amplitude (loudness)	Distance to source
Inter-aural delay time	Direction to source (Haas Precedence Effect)
Inter-aural balance	Direction to source (in the horizontal plane)
Spectrum (many dimensions)	Distance (low-pass filter), direction (nonlinear filter), characteristics of the space
Reverberation	Distance (direct signal ratio), characteristics of the space (reverberation contour)

Uses for and problems associated with 3-D audio are described in the following sections.

3.3.1 Usage

3-D audio has been used to (1) improve intelligibility, (2) provide navigation cues, (3) warn of threats, (4) support targeting, (5) indicate location of a wingman, (6) give location cues to air traffic controllers, (7) help the blind navigate, and (8) provide hands-free communication. Evaluation of each of these uses is described in the following sections.

3.3.1.1 Improve Intelligibility

In a survey of 76 experienced military pilots, M. D. Lee et al.[82] reported strong preferences for 3-D sound in wingman communication and threat warning. The pilots included 30 Air Force, 43 Navy, 2 Marine, and 1 Army; each completed a survey after listening to prepared topics using stereo headphones. There was a strong preference for 3-D audio from the simulated actual location for threat warnings (62 percent), communication with wingman (67 percent), and intercom communications (54 percent). For system status information, 57 percent of the pilots wanted the audio source to be the relevant visual display. There were no majority responses for the use of 3-D for navigation, malfunction messages, flight configuration, and ground-to-air or other communications. A major concern, however, was overload of the auditory channel.

NASA[83] has consistently shown an improvement of approximately 6 dB in intelligibility through the use of 3-D audio communications. The advantage may be due to the use of head-shadow and binaural interaction (also known as the “cocktail party effect”).

3-D audio was integrated with a helmet-mounted display in a TAV-8B Harrier, and three uses of 3-D audio were evaluated during the flight test: (1) spatially separated communications, (2) threat

[81] S. T. Pope and L. E. Fahlen, “The Use of 3-D Audio in a Synthetic Environment: An Aural Renderer for a Distributed Virtual Reality System,” *IEEE Virtual Reality Annual International Symposium* (1993), pp. 176–182.

[82] M. D. Lee, R. W. Patterson, D. J. Folds, and D. A. Dotson, “Application of Simulated Three-Dimensional Audio Displays to Fighter Cockpits: A User Survey,” *Proceedings of the IEEE 1993 National Aerospace and Electronics Conference*, Vol. 2 (1993), pp. 654–660.

[83] Durand R. Begault, “Head-Up Auditory Display Research at NASA Ames Research Center,” *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, Vol. 1 (1995), pp. 114–118.

location cueing, and (3) target location aiding. 3-D audio increased speech intelligibility (see Figure 17) with the helmet-mounted display; 3-D audio increased the percent of correct detection (see Figure 18) but not the distance (see Figure 19).

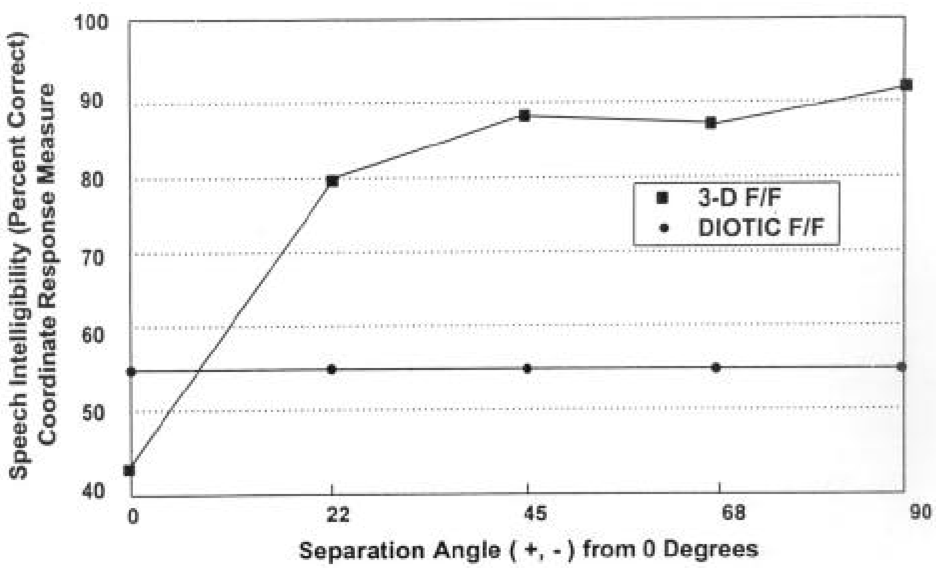

Figure 17. *3-D audio speech intelligibility normal (diotic) vs. spatially separated (3-D).*[84]

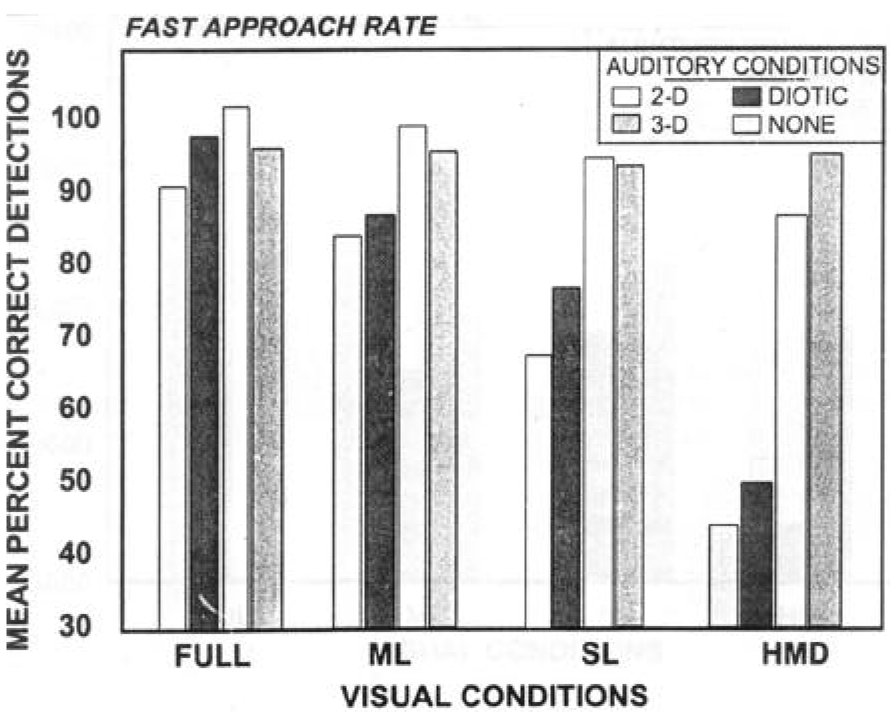

Figure 18. *Mean percent correct detection of targets with and without 3-D audio cueing.*[85]

[84] R. L. McKinley, W. R. D'Angelo, and M. A. Ericson, "Flight Demonstration of an Integrated 3-D Auditory Display for Communication, Threat Warning, and Targeting," *AGARD Conference Proceedings Audio Effectiveness in Aviation*, AGARD-CP-596 (1996), p. 6-8.

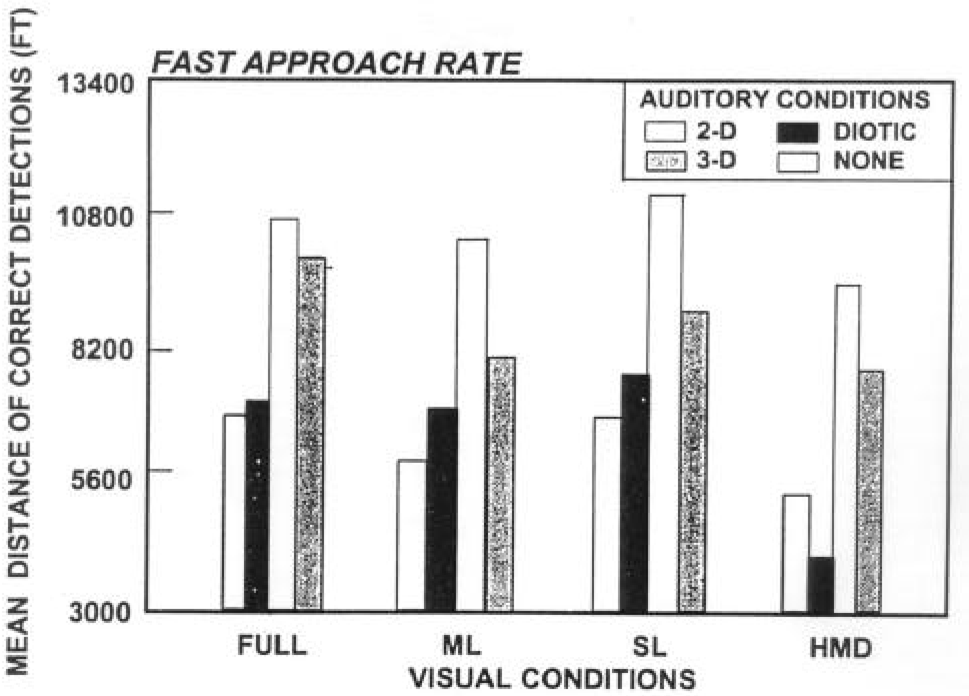

Figure 19. *Mean distance of correct visual detections of targets with and without 3-D audio cueing.*[86]

The U.S. Army[87] compared the number of correct pilot responses (that is, the pilot replied on the target radio channel when a target message was present) in three radio signal presentation modes: diotic, dichotic, and 3-D audio. In the diotic mode, speech messages from three simulated radios were routed to both ears equally; in the dichotic mode, speech messages from two simulated radios were routed to one ear and the third radio to the other ear; in the 3-D mode, the three radios were presented respectively at 90°, 270°, and 315° azimuth. Data were collected in the Army Research Institute Simulation Training Research Advanced Testbed for Aviation simulator. The subjects were 11 U.S. Army helicopter pilots certified in the AH-64 helicopter who performed the radio identification task while engaging in target acquisition and responding to aircraft malfunctions. The results showed significantly better performance (5.0) using the 3-D audio than the diotic displays (2.0) currently used in helicopters (performance for dichotic displays was 3.9).

3.3.1.2 Provide Navigation Cues (Spatial Presentation)

C. Hendrix and W. Barfield[88] asked 16 university students to rate their perceived level of presence in a virtual world with and without 3-D audio. Only two locations were presented—one simulating a radio broadcast, the other simulating change being deposited in a vending machine. 3-D audio significantly increased the ratings of presence but not realism in the virtual environment.

---

[85] Ibid., p. 6-11.
[86] R. L. McKinley, W. R. D'Angelo, and M. A. Ericson (1996), p. 6-8.
[87] E. C. Haas, C. Gainer, D. Wrightman, M. Couch, and R. Shilling, "Enhancing System Safety With 3-D Audio Displays," *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (1997), pp. 868–872.
[88] C. Hendrix and W. Barfield, "Presence in Virtual Environments as a Function of Visual and Auditory Cues," *Proceedings of the Virtual Reality Annual International Symposium* (1995), pp. 74–82.

As Sen M. Kuo and G. H. Canfield stated, “3-D sound could be of great benefit in VR, augmented reality, or remote operator environments to efficiently transfer position information.” They also state that, “the acoustic path from the loudspeakers to the destination will introduce spectral and phase distortion in the signals.”[89] To overcome these problems, these authors developed a dual-stage algorithm and then modified it to use low-level additive noise. The algorithm was tested only with spectrally flat signals.

3.3.1.3 Warn of Threats

The U.S. Marines flight-tested 3-D audio displays in an AV-8B in the Fall of 1991. The displays were those developed by AFRL. The test evaluated the utility of these displays for warning of a missile approach. Results indicated that missiles could be located within 10 degrees.[90]

U.S. Air Force researchers[91] reported that subjects could detect a monochrome silhouette of an SU-27 aircraft with the naked eye as well as with a helmet-mounted display if 3-D audio cueing was used. The rated workload (according to the NASA Task Load Index) was lowest in the 3-D audio condition as compared to no sound or nonlocalized sound conditions.

NASA has shown a 500-ms improvement in acquiring targets since it started using a 3-D audio version of the Traffic Alert and Collision Avoidance System (TCAS).

3.3.1.4 Support Targeting

While pilots participating in the 3-D audio AV-8B display flight tests reported targeting accuracy within 15 degrees azimuth, which they felt was adequate to orient toward a target,[92] elevation cues were less accurate and enabled only rough judgments of low or high. However, M. A. Ericson, R. L. McKinley, M. P. Kibbe, and D. J. Francis reported[93] that in-flight, 3-D audio reduced target acquisition times.

The U.S. Air Force, in a series of laboratory experiments,[94] reported significantly shorter search times for targets using 3-D audio cues. The worst performance occurred at +/-150 degrees azimuth, but even that performance was better with than that without 3-D audio. D. R. Perrot et al. had reported a similar enhancement with 3-D audio in a two-alternative visual search task (see Figure 20).

---

[89] Sen M. Kuo and G. H. Canfield, “Dual-Channel Audio Localization and Cross-Talk Cancellation for 3-D Sound Reproduction,” *IEEE Transactions on Consumer Electronics*, Vol. 43, No. 4 (1997), p. 1189.

[90] Jane’s Information Group, “AV-8B to Test 3-D Audio Displays,” *International Defence Review*, Vol. 24, No. 6 (1 December 1991), p. 176.

[91] R. L. McKinley, W. R. D’Angelo, M. W. Haas, D. R. Perrot, W. T. Nelson, L. J. Hettinger, and B. J. Barickman, “An Initial Study of the Effects of 3-Dimensional Auditory Cueing on Visual Target Detection,” *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 119–123.

[92] M. P. Kibbe and D. J. Francis, “TAV-8B Flight Test Results for the Demonstration of an Airborne 3-D Audio Cuer,” *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (1994), p. 987.

[93] M. A. Ericson, R. L. McKinley, M. P. Kibbe, and D. J. Francis, “Laboratory and In-Flight Experiments to Evaluate 3-D Audio Display Technology,” *Proceedings of the Space Operations, Application, and Research Conference* (1993), pp. 371–377.

[94] D. R. Perrott, J. Cisneros, R. L. McKinley, and W. R. D’Angelo, “Aurally Aided Detection and Identification of Visual Targets,” *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 104–108.

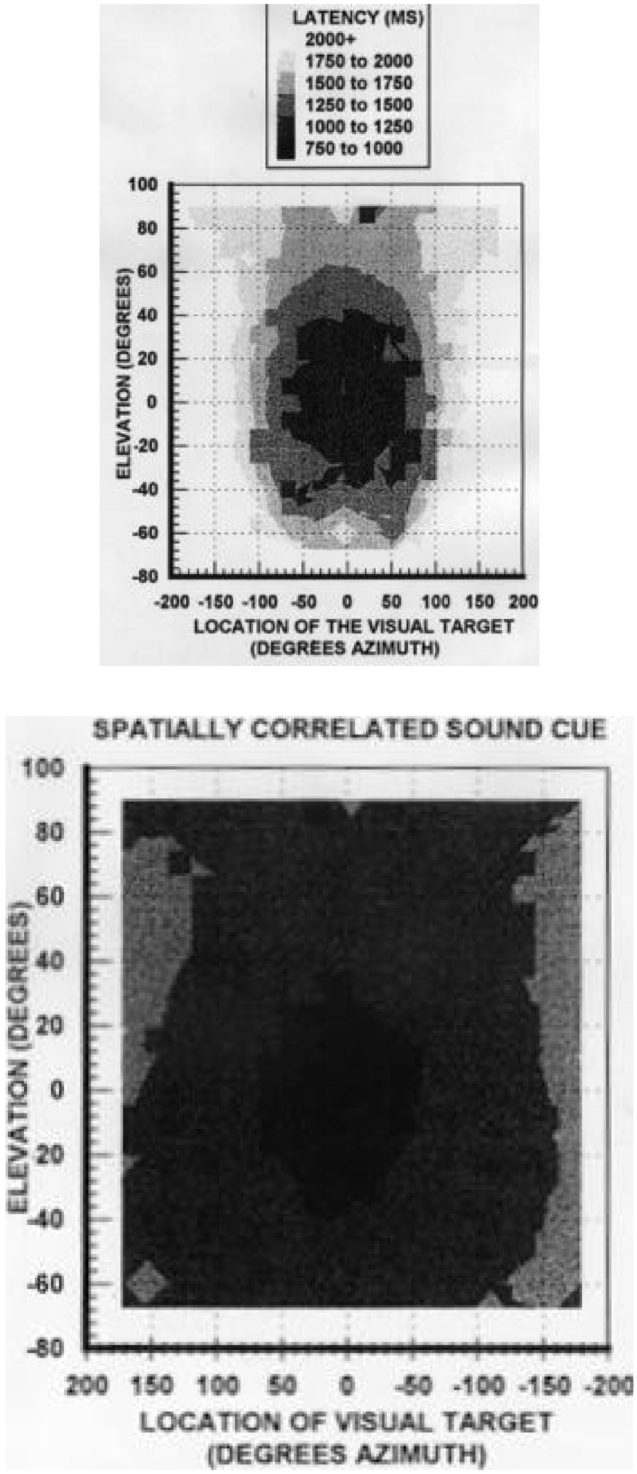

Figure 20. *Target search latency with and without 3-D audio cueing.*[95]

[95] D. R. Perrott, J. Cisneros, R. L. McKinley, and W. R. D'Angelo, "Aurally Aided Detection and Identification of Visual Targets," *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 105–106.

Adelbert W. Bronkhorst, J. A. Veltman, and Leo van Breda reported[96] that search times for a military target were not significantly different for a visual display, a 3-D audio display, or both together. All three conditions had significantly shorter search times than in the no-display condition. There was no significant difference among the four conditions in workload rating, however. The subjects were six Dutch helicopter pilots and two trained observers. The data were collected in a fixed-base flight simulator.

Adelbert W. Bronkhorst and J. A. Veltman[97] compared the search time and workload associated with a simulated target localization between 2-D and 3-D audio. The subjects were eight Royal Netherlands helicopter pilots. 3-D audio search times were less than times with 2-D audio. But the shortest search times occurred when both 2- and 3-D audio were presented (see Figure 21). The workload was not significantly different among the four conditions.

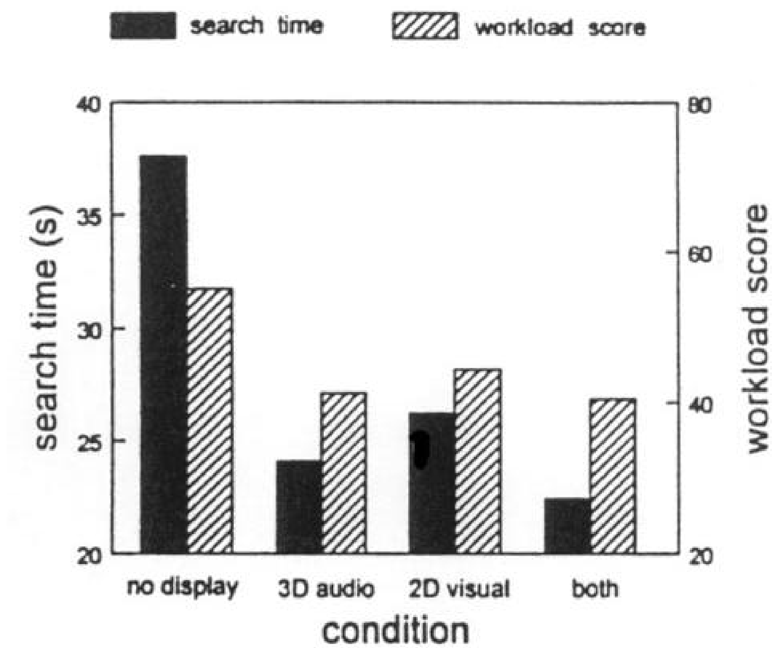

Figure 21. *Average search times and workload scores.*[98]

Twelve pilots participated in a second experiment that examined 3-D visual displays. Search time, tracking error, and workload were lowest when 3-D visual displays were used (see Figures 22 and 23).

[96] Adelbert W. Bronkhorst, J. A. (Hans) Veltman, and Leo van Breda, "Application of a Three-Dimensional Auditory Display in a Flight Task," *Human Factors*, Vol. 38, No. 1 (1996), pp. 23–33.
[97] Adelbert W. Bronkhorst and J. A. (Hans) Veltman, "Evaluation of a Three-Dimensional Auditory Display in Simulated Flight," *AGARD Proceedings Audio Effectiveness in Aviation*, AGARD-CP-596 (1996), pp. 5-1 to 5-6.
[98] Ibid., p. 5-4.

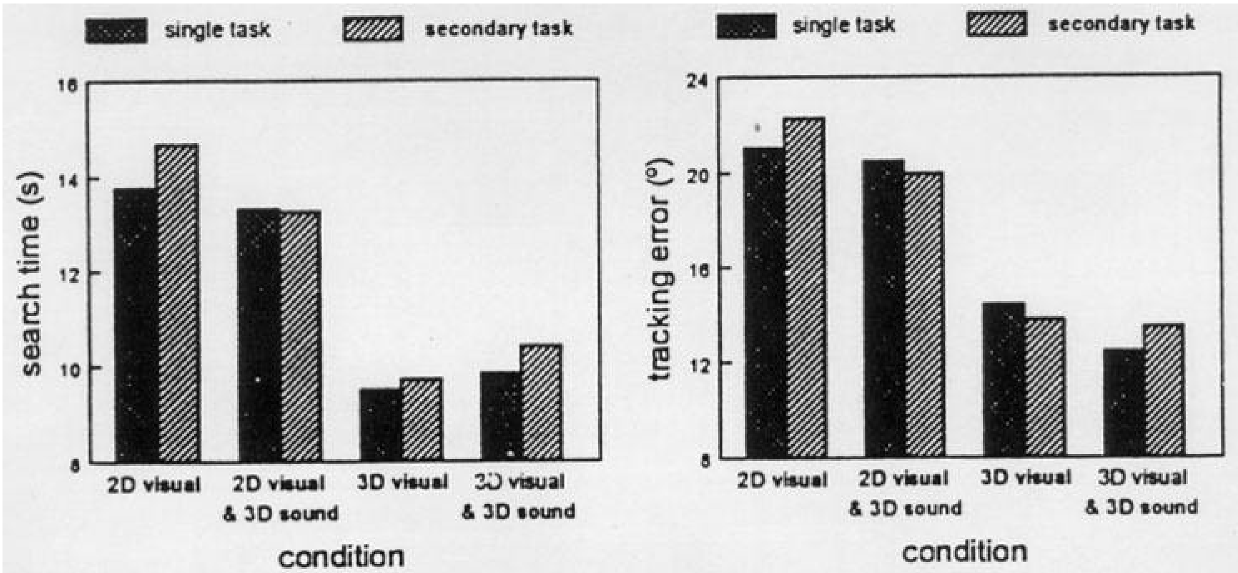

Figure 22. *Average search times and tracking errors.*[99]

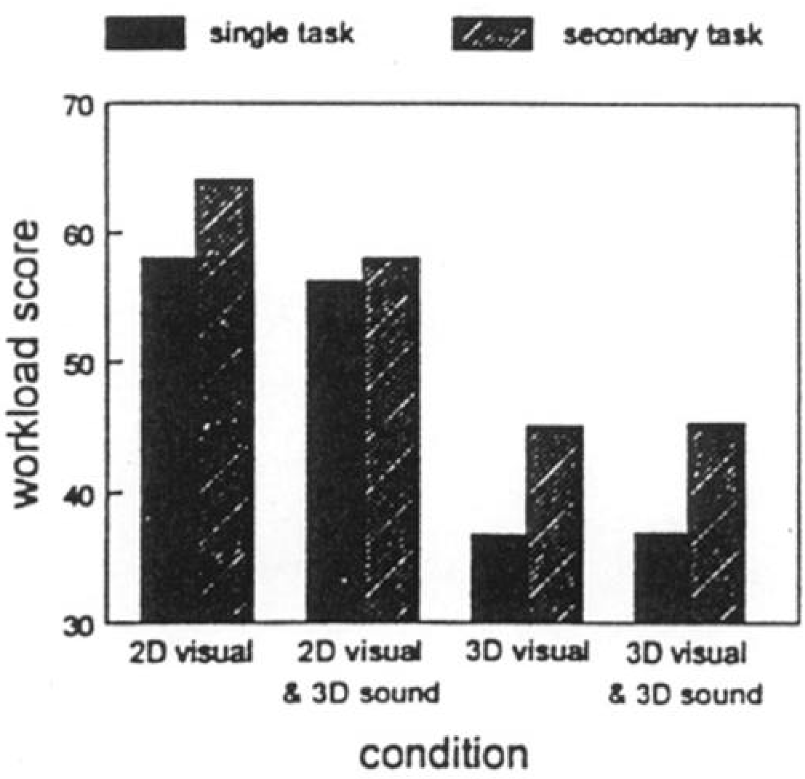

Figure 23. *Average workload scores.*[100]

[99] Adelbert W. Bronkhorst and J. A. (Hans) Veltman, "Evaluation of a Three-Dimensional Auditory Display in Simulated Flight," *AGARD Proceedings Audio Effectiveness in Aviation*, AGARD-CP-596 (1996), p. 5-5.
[100] Ibid.

A 500-ms improvement in target acquisition time was demonstrated when 3-D audio was added to the standard TCAS.[101] The subjects were 10 commercial airline crews. The test facility was the NASA Ames Crew-Vehicle Systems Research Facility Advanced Concepts Flight Simulator.

NASA[102, 103] compared the acquisition time of targets using the standard head-down TCAS and a 3-D audio presentation of the same information. The subjects were 10 two-person crews composed of airline pilots rated in Boeing 757, 767, 737-300/400, or 747-400 aircraft. Data were collected in the NASA-Ames Crew-Vehicle Systems Research Facility Advanced Concepts Flight Simulator. The results indicate a 500-ms improvement in acquiring targets using a 3-D audio version of the TCAS (2.13 s) rather than the standard TCAS (2.63 s).

The 500-ms improvement has also been reported in simple laboratory search tasks.[104] This improvement occurred 24 degrees from the fixation point of five subjects in an audiometric chamber for a 70-dB audio cue. The improvement was slightly less (300 ms) for a 40-dB audio cue.

3.3.1.5 Indicate Location of a Wingman

A lead pilot can receive an indication of the location of wingman using outputs from the aircraft's GPS receivers to establish their relative location.

3.3.1.6 Give Location Cues to Air Traffic Controllers

3-D audio has been used to provide location cues of incoming and departing aircraft by projecting the pilot's voice from the relative position in the air or on the ground.

3.3.1.7 Help the Blind Navigate

A blind person wears a computer instrumented with a GPS receiver and a detailed map database. The person's position is compared to the desired position and a 3-D signal is provided to keep the person on course.[105]

3.3.1.8 Provide Hands-Free Communication

A prototype 3-D audio system was evaluated during an operational exercise at the North American Aerospace Defense complex. The system consisted of a 3-D headset, a boom microphone, and a push-to-talk foot switch. Operator comments were positive, and six systems were procured.[106]

---

[101] Durand R. Begault and M. T. Pitman, "Three-dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays," *International Journal of Aviation Psychology*, Vol. 6, No. 1 (1996), pp. 79–93.
[102] Durand R. Begault and M. T. Pitman, 1995.
[103] Durand R. Begault and M. T. Pitman, "Three-dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays," *International Journal of Aviation Psychology*, Vol. 6, No. 1 (1996), pp. 79–93.
[104] T. Z. Strybel, J. M. Boucher, G. E. Fujawa, and C. S. Volp, "Auditory Spatial Cueing in Visual Search Tasks Effects of Amplitude, Contract, and Duration." *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 109–113.
[105] "Three-D Sounds Points Pilots Toward the Enemy," *Machine Design* (22 November 1999), pp. 40–41.
[106] D. A. North and W. R. D'Angelo, *3-Dimensional Audio Ergonomic Improvement Project for the NORAD CMOC* (AL/CF-TR-1997-0170). Wright-Patterson Air Force Base, OH: Armstrong Laboratory, 1997.

3.3.2 Problems

Problems of implementation include

1. Dual-channel equalization—for the human to detect direction, it is critical that the sound in each ear be equalized prior to delivery of the 3-D signal; this requires cross-talk cancellation in the earphones.
2. Vibration, which reduces hearing perception, especially at high vibrations (100,000 Hz).
3. Individual differences—head-related transfer functions have been developed to tailor 3-D to variations in ear canals. Some researchers have found increased rather than decreased localization error while using these functions.[107] These transfer functions have been enhanced using artificial neural networks.[108]
4. Noise—for signal-to-noise ratios less than 15 dB, noise can make localization more difficult; this is especially true of pure tones.
5. Communication—the same earphones used for the 3-D signal are used for communication, and there have been some problems of acceptance by transport pilots.
6. Postural adaptation—after head rotation, the perception of center is displaced in the direction of the original rotation.
7. Cones of confusion—3-D audio requires temporal disparity between signals to the left and right ear. Small or no disparities indicate that the sound is emanating from the vertical plane between the two ears, anywhere in this plane. The greatest confusion is up/down and front/back. Front/back reversals are common, back/front less so. For example, Durand R. Begault and E. M. Wenzel reported[109] 11 percent back/front reversals compared to 47 percent front/back. This was for an auditory target localization task in a sound isolation chamber.
8. Capability of synthesizers—there are differences in users' ability to determine direction of sound sources as a function of capability of the auditory localization cue synthesizers. Based on data from six male subjects, G. Valencia, M. A. Ericson, and J. R. Agnew reported[110] no significant differences between a system presenting two sound sources varying in azimuth coupled with head position (DIRAD) and a system presenting four sound sources varying in azimuth coupled with head position, evaluation, and distance (AL-204). Measures were mean magnitude error, mean response time, and mean percentage of reversals. However, there was a significant interaction between type of synthesizer and target location. Mean magnitude error was significantly greater for the AL-204 when the target was 0 to 59 degrees (zero was straight ahead, 59 degrees to the subjects' right). Furthermore, front/back reversals occurred with the AL-204 when the target was at 0 to 59, 240 to 299, or 300 to 359 degrees (see Figure 24). In a comparison

[107] D. S. Savick, *A Comparison of Various Types of Head-Related Transfer Functions for 3-D Sound in the Virtual Environment* (ARL-TR-1605). Aberdeen Proving Ground, MD: Army Research Laboratory, 1998.
[108] D. Reinhardt, *Neural Network Modeling of the Head-Related Transfer Function* (AFIT/GAM/ENC/98M-01). Dayton, OH: Air Force Institute of Technology, 1998.
[109] Durand R. Begault and E. M. Wenzel, "Headphone Localization of Speech," *Human Factors*, Vol. 35, No. 2 (1993), pp. 361–376.
[110] G. Valencia, M.A. Ericson, and J. R. Agnew, "A Comparison of Localization Performance With Two Auditory Cue Synthesizers," *Proceedings of the 1990 National Aerospace and Electronics Conference*, Vol. 2, pp. 749-754.

of the three experienced subjects versus the three inexperienced subjects, the experienced subjects had less magnitude error and fewer reversals.

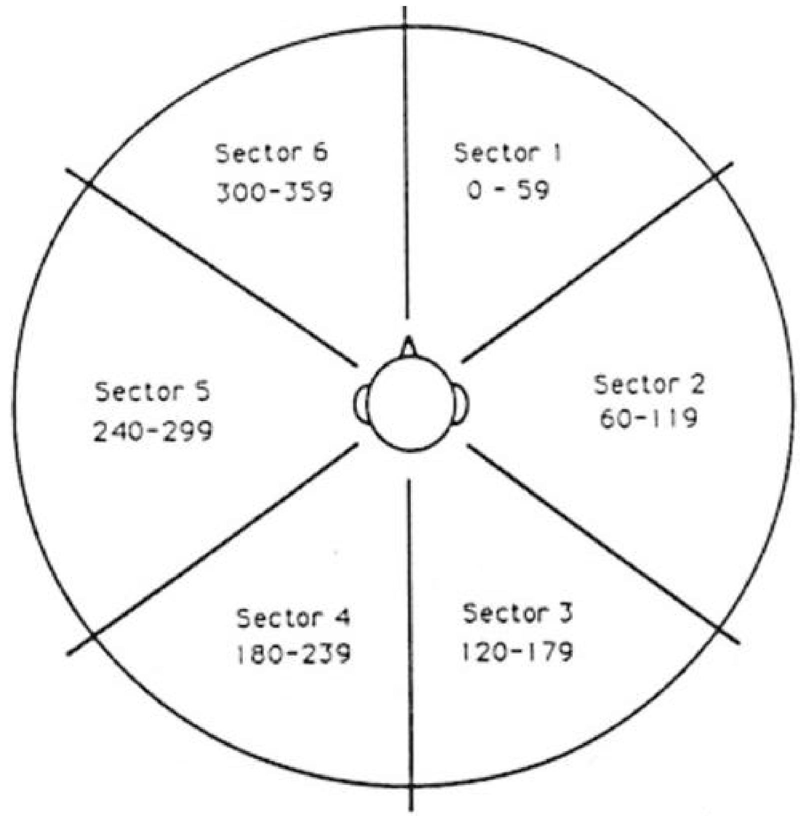

Figure 24. *Grouping of target locations into sectors.*[111]

9. Limited bandwidth—applying 3-D audio to military aircraft has been difficult due to the limited signal bandwidth over which to present the sound. Based on the data of three subjects, Robert B. King and Simon R. Oldfield concluded[112] that the ability to localize targets in elevation was lost when the signal was limited between 0 and 9 kHz or front/back between 0 and 7 kHz or between 10 and 16 kHz. They recommended a 0- to 16-kHz bandwidth.

10. Spectral proximity—the greater the spectral proximity, the lower the probability of correctly distinguishing sounds by either spatial separation or signal frequency (see Figure 25).

[111] G. Valencia, M.A. Ericson, and J. R. Agnew, "A Comparison of Localization Performance With Two Auditory Cue Synthesizers," *Proceedings of the 1990 National Aerospace and Electronics Conference*, Vol. 2, p. 751.

[112] Robert B. King and Simon R. Oldfield, "The Impact of Signal Bandwidth on Auditory Localization: Implications for the Design of Three-Dimensional Audio Displays," *Human Factors*, Vol. 39, No. 2 (1997), pp. 287–295.

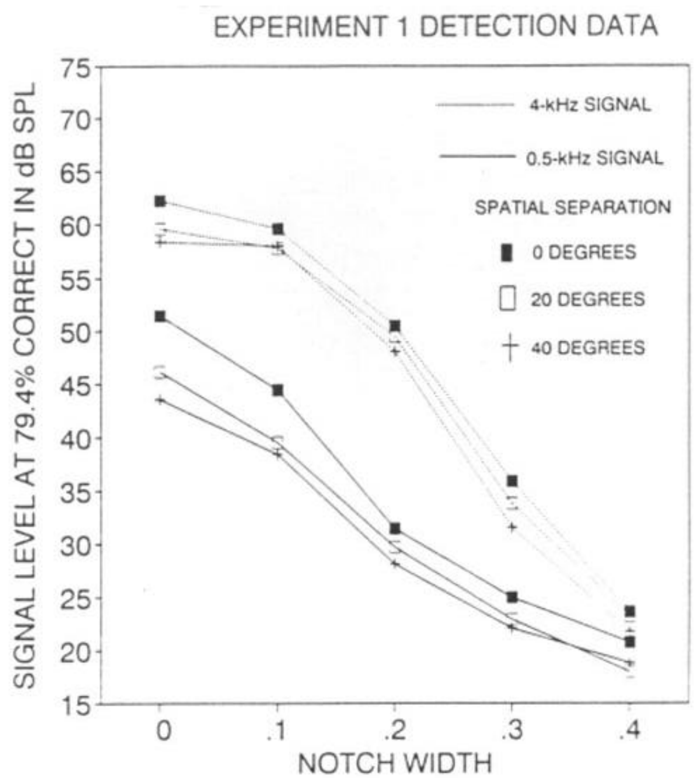

Figure 25. *Average detection level as a function of spectral proximity (notch width), spatial separation, and signal frequency.*[113]

[113] T. J. Doll and T. E. Hanna, "Spatial and Spectral Release From Masking in Three-Dimensional Auditory Displays," *Human Factors*, Vol. 37, No. 2 (1995), p. 345.

A comparison of audio technology is presented in Table 3.

Table 3. *Audio Technology Comparison*[114]

Type of processing	Dimensionality	Interactive controls	Perceptual performance	Headphone-compatible	Stereo speaker–compatible	Direct 3-D sound–compatible
Mono	0D	None (on/off)	Single-point source from speaker location	Yes	Yes	No
Stereo	1D (left/right)	Panning (left/right)	Sounds placed between speakers	Yes	Yes	No
Simple stereo extender	1D (spaciousness only)	None (on/off)	Sounds fill area around speakers	No	Yes	No
Advanced stereo extender (Qsound)	1D (left/right)	Panning (left/right)	Sounds placed on arc extending through speakers	No	Yes	No
Multi-speaker array (surround sound formats)	2-D (left/right, front/back)	Usually none (sound-tracks are pre-encoded)	Sounds placed on circle formed by speakers	Yes	Yes	No
Interactive 3-D audio (Aural 3-D)	3-D (left/right, front/back, up/down)	Full 3-D placement using XYZ coordinates	Sounds placed at any distance and position from listener	Yes	Yes	Yes

Beth Wenzel[115] described spatial auditory displays. Virtual acoustic environments require nonspatial source synthesis, sound field synthesis, and listener reception/directional synthesis. The performance advantages of 3-D sound are enhanced situational awareness (direct representation of spatial information, omnidirectional monitoring, reinforcement of information in other modalities, enhanced sense of presence) and enhanced multiple-channel presentation. Errors in the natural environment get worse in virtual environments. Visually dominant people seem to have more front/back reversals, probably since if nothing is seen, they conclude that the object must be behind them. Latencies of up to 500 ms are noticeable but do not greatly disrupt localization.

---

[114] Dave Bursky, "3-D Audio Technologies Provide Realistic Sound," *Electronic Design*, Vol. 44 (4 November 1996), p. 80.

[115] Information-gathering meeting, 15 April 1999. Beth Wenzel (bwenzel@mail.arc.nasa.gov) maintains a Spatial Auditory Display homepage at http://vision.arc.nasa.gov/~bwenzel/index.html.

3.4 Tailoring

One of the first steps in developing a user tailoring capability is to develop a language-independent knowledge base that "contains knowledge about user interface components and functions of the software applications."[116] Examples of tailored views are given in Figures 26 through 30.

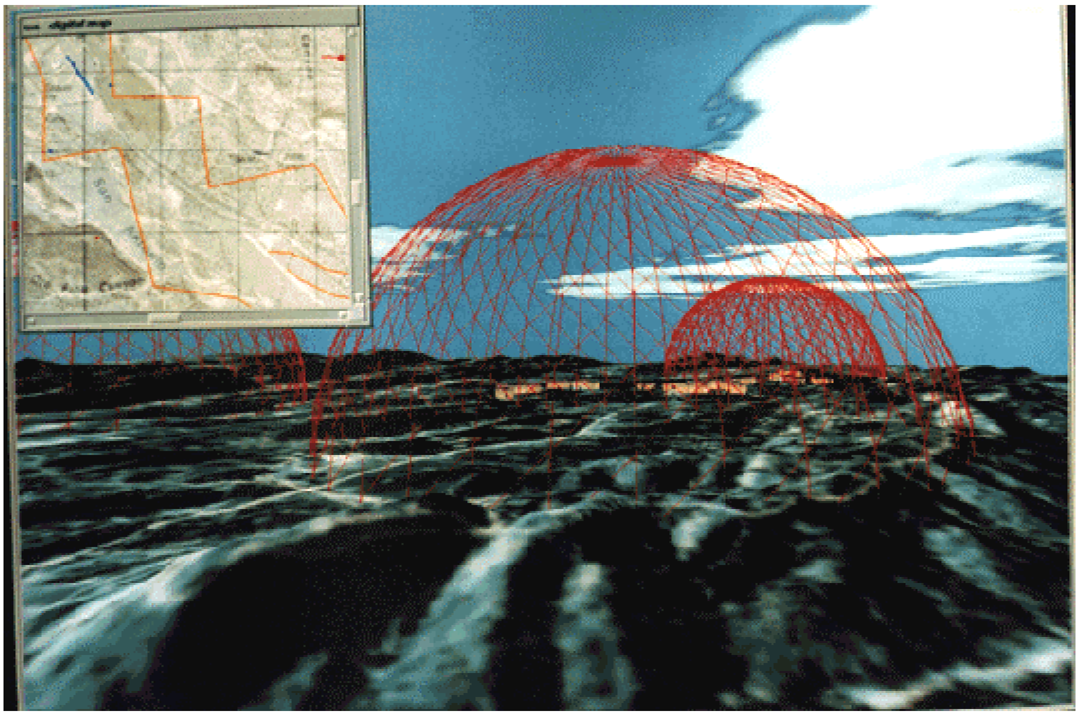

Figure 26. *Army commander tailored view.*

[116] E. A. Karkaletsis, C. D. Spyropoulos, and G. A. Vouros, "A Knowledge-Based Methodology for Supporting Multilingual and User-Tailored Interfaces," *Interacting With Computers*, Vol. 9 (1998), p. 312.

Figure 27. *Air Force commander tailored view.*

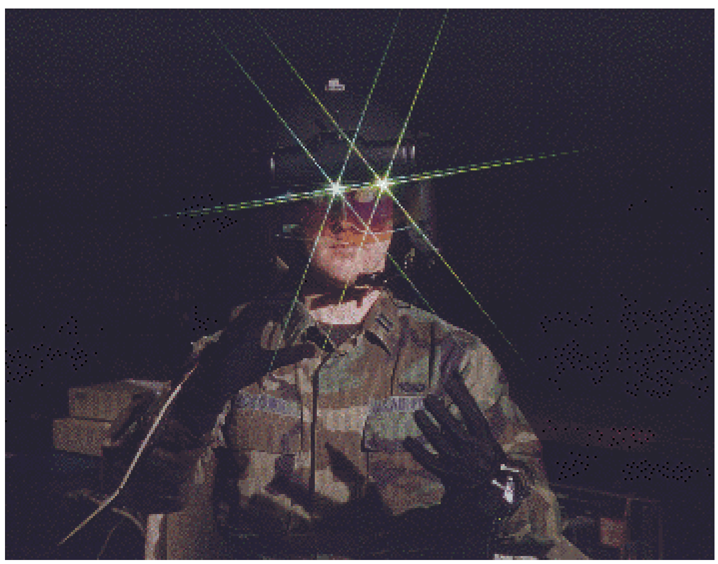

Figure 28. *Army aviator view during mission execution.*

Figure 29. *Air Force aviator view during mission rehearsal.*

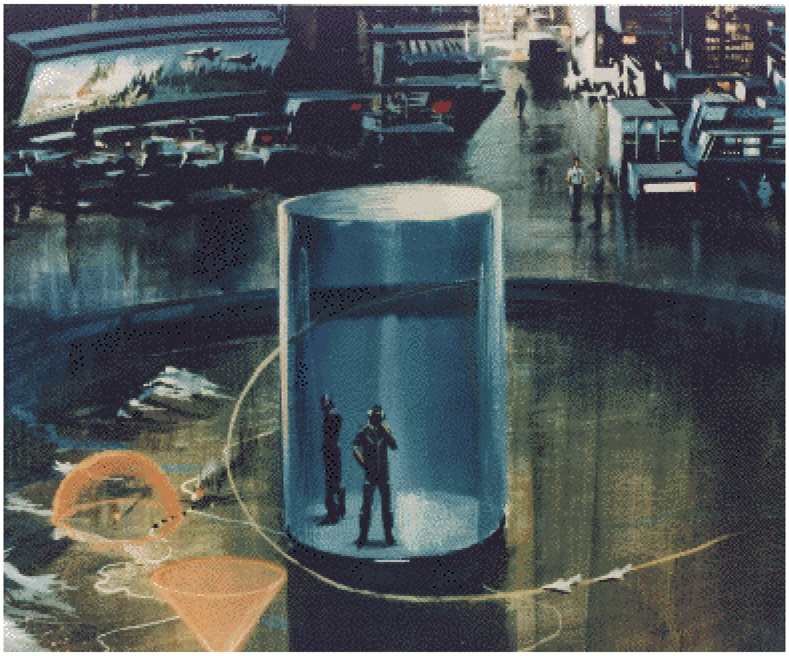

Figure 30. *Joint Force Commander view.*

Some COTS products already exist—Omnidesk, for example. This applet creates a user-configurable desktop for a web browser.[117] The companion, OmniFlow, allows the user to create a dependency graph of data.

[117] H. Lavana and F. Brglez, *OmniDesk and OmniFlows: Platform Independent Executable and User-Reconfigurable Desktops*. Research Triangle Park, NC: U.S. Army Research Office, 1997.

One DARPA program, Genoa,[118] is developing collective reasoning tools. The premise is: the earlier that crisis situations (stew pots) are identified and understood, the easier it is to arrive at pre-emptive or mitigative strategies—"Better decisions today and tomorrow through informed, structured collected reasoning." The Genoa ExtraNet includes the CIA, Defense Intelligence Agency, National Security Agency, commanders in chief, Joint Chiefs of Staff, Office of the Secretary of Defense, National Security Council and State Department. The ExtraNet is not being developed under Genoa but under information assurance work in other areas of DARPA. Analyzing and decision making are the focus of Genoa, which is extensible from the commander in chief to the lowest appropriate echelon. The Genoa process collaborates and shares information from the analysts and policy makers. The Thematic Argument Group is a virtual place in which arguments on a theme are worked by persons in distributed locations to develop a structured argument. Thematic Argument Groups are meant to be built and be destroyed very easily. They also can support existing organizational structures.

The Critical Information Package is a collection of information woven together into a structure. Critical Information Packages are intended to be persistent in type and modified in subsequent versions. A Critical Information Package contains structured arguments (i.e., critical intent model and structured evaluation and analysis system). The critical intent model is a program evaluation review technique chart for scheduling development of a capability such as weaponization. The structured evaluation and analysis system is a template for rolling up judgments in a stoplight manner—deciding, for instance, whether a cult has declared its intention to use terrorist acts. A Critical Information Package also includes virtual situation book libraries (to prepare multimedia presentations), supporting evidence (the raw data), and metadata (name, source, classification, authorization, access, confidence, description, keywords). The Genoa contractors are ISX, Global InfoTek, Pacific-Sierra Research, Syntek, CMU, and SAIC.

Genoa has four technologies: knowledge discovery, structured augmentation, corporate memory, and virtual collaborative environment. Knowledge discovery is being leveraged from other programs (such as Infomedia). It is an automated process to discover data trends, patterns, and anomalies and to filter out spurious data. Logically structured argumentation records complex analytic reasoning that must be readily assimilated, critiqued, and compared. This will provide tools via which analysts will argue. The critical intent model and the structured evaluation and analysis system are structured-argumentation tools, which focus analysis by leading users to drill down. These tools enable comparison of arguments to identify differences and reasons for their differences. Corporate memory is augmented support for comparing current situations to known past crises. It retains what you know, where you learned it, from whom you learned it, and what you did about it. Collaborative environment includes a Thematic Argument Group manager that provides business rules, task, and event management and user authorization to a multi-user, shareable application with enclave support.

Measuring Genoa's success includes asking, "How rich are the data arguments? Are better decisions being made? What is the diversity of the human decision-makers?" This will be shadowed this summer using actual analysts.

[118] Information-gathering meeting at DARPA, 19 May 1999.

The purpose of the Intelligent Collaboration and Visualization (IC&V) program[119] is to develop technology to support planning and execution. There are three key points. The first is collaboration among distributed systems connected with diverse bandwidths and accessed through a range of devices from handheld to room size. The second is collaboration among persons with sporadic connectivity and among changing personnel. The third point is determination of the technologies to select, sort, and search a multimedia environment. Pacific Command is the most challenging region for military crisis response due to vast distances, wide variation in communication methods, and multiple simultaneous crises.

IC&V was demonstrated at the Space and Naval Warfare Systems Command (SPAWAR) in May 1998 and at Pacific Command in October 1998. The following tools were involved in the demonstration:

- MASH, a multimedia architecture that scales across heterogeneous environments,[120] enables multimedia conferencing among hundreds of thousands of users. It transcodes multimedia streams, images, and protocols. It permits shared control of time-varying visualization. It is in trial use at Pacific Command.
- Visage Link[121] provides a collaborative visual medium and is being hardened for military use.
- Orbit Gold is a collaborative environment for people juggling multiple collaboration.
- The (Integrated Synchronous And Asynchronous Collaboration (ISAAAC) system is based on Habanero used to collaborate across heterogeneous computer systems.
- CSpace is an asynchronous collaboration across heterogeneous office applications. It extracts events from within commercial office applications, then constructs a graph structure representing all the changes in applications and enables users to maintain their own view and awareness of the state of the shared graph structure. This model could drive implicit collaboration.

Total Information Awareness program[122] is aimed at asymmetric warfare with a transnational threat. There are near-field (perimeter security, people tracking, face recognition, and news bulletin), transition zone, and far-field (databases, data mining, and heterogeneous search) levels of the problem. Near-field levels have less reaction time and fewer response options. Key components are data gathering, information discovery (model-driven search agents may be developed by industry), models and behavior (intent models, evidence models, model-driven search agents, and inference agents), and collective reasoning (argument templates from Genoa) moving from machine to human decision making. The project is seeking input; e-mail ideas to tia@darpa.mil.

A current DARPA program is Active Templates, which focuses on parameterizing problem-solving methods and uses a spreadsheet interface to build simple planning systems that are reactive to handle real-world activation.[123] The goal of the Active Templates program is to automate military operations, maintaining a causal situation model and providing incremental payoff as new automated functions are added and the causal model is improved. Although

[119] Ibid.
[120] See http://www-mash.cs.berkeley.edu/mash/.
[121] http://www.maya.com/visage/link/.
[122] Information-gathering meeting at DARPA, 19 May 1999.
[123] Ibid.

developed for the military, Active Templates are expected to have significant commercial payoff as the plans are adapted, merged, and updated with other plans. A goal is to make them user tailorable. Active Templates are being used at the Air Force Special Operations Command to determine features that must be added. BBN, ISI, and CMU have jumpstart efforts for this program.

An existing system, Quality-based Tactical Image Exploitation, is being integrated into the United States Navy's primary afloat command and control system. It tailors imagery information to user needs based on user preferences.[124]

Broadsword is a modular, object-oriented framework that provides "data brokering," auditing, and connectivity services to heterogeneous data sources.[125]

The Human-Centered Intelligent Systems Supporting Communication and Collaboration program is managed by Mike Shafto,[126] the Human-Centered Computing Group Manager in the NASA Ames Research Center Computational Sciences Division. Human-centered computing (HCC) is a software engineering methodology that improves both human and computer performance. "Human-centered" means that design is performed from a systems perspective, taking into account a scientific understanding of the nature of human and computer capabilities. As an engineering research area, HCC provides the methodology for integrating computer hardware and software with teams of human operators, to build systems that make best use of all human and computer resources. HCC embodies a "systems view" in which the interplay between human thought and action on the one hand, and hardware/software functionality on the other, is considered right from the start, rather than as an afterthought. Within this framework, NASA researchers are inventing and deploying innovative computational aids designed to complement human cognitive and perceptual capabilities. These aids rely both on computer-intensive data analysis and on human-centered visualization techniques.

The future vision is of work systems in which intelligent agents will enable teams and organizations to work together more effectively to achieve complex mission goals. Work system design requires articulating, simulating, and testing our understanding of dynamic interactions among people, technologies, and the physical and organizational environment. To enhance human performance in complex systems, NASA must advance theory, models, simulations, and enhancements of perception, cognition, learning, and communication. Examples of perception research topics include models of multimodal perception, speech production and speech perception, and communication and control. Examples of cognition research topics include human error, memory, cognitive capacity, attention, multitask performance, decision making, executive control, task interference, and fatigue.

The HCC approach to software engineering emphasizes participatory design and partnership between those who use and those who develop computer systems. Work practices and team learning are carefully analyzed by means of participant observation, ethnography, video

[124] P. Kaomea and W. Page, "A Flexible Information Manufacturing System for the Generation of Tailored Information Products," *Decision Support Systems*, Vol. 20 (1997), pp. 345–355.

[125] See Section 2.1 and http://www.if.afrl.af.mil/bsword/ for further information.

interaction analysis, prototyping, and evaluation in the context of real work settings. These analyses are used as the basis for the design of new automation and other kinds of computer systems. HCC promotes the use of formal modeling as a design tool for both software engineers and users, integrating multiple views: workflow, information processing tasks, reasoning, and action. Models are used to capture knowledge about current expertise and work practice, as well as to envision how new automation and innovative organizational concepts can improve team effectiveness for future missions.

Boeing, as part of the CPoF program, is working on the Human Computer Interface Manager component that intelligently tailors staff displays to remain in sync with the changing command post situation or context. The Human Computer Interface Manager context manager uses powerful inferencing models to interpret command post staff intent, based on staff interactions with the CPoF system. These inferences, along with other relevant context information, are then added to the Human Computer Interface Manager knowledge-based algorithms that select and configure appropriate presentation elements for display. The result of this program will be a fully functioning component, ready for integration and evaluation, with an aim toward eventual transition to a broad range of future command posts.

3.3.4 Task and User Modeling

To build intelligent systems that are truly helpful to people, people and their jobs must be understood. Process modeling examines the structure of the tasks and the environment. Cognitive modeling examines the user's problem-solving and decision-making behaviors as the tasks are performed. New computer tools for collaborative performance and human-machine interaction necessarily change how work is done, how people work together, and where work occurs. Modeling is therefore required to define the requirements for human-computer systems of the future.

NASA human factors scientists are concerned with mitigating human errors, ranging from frequent air traffic control and aviation incidents to the human/automation factors in the *Mir* collision. Recurring patterns of design-induced error attest to the inadequacy of current knowledge for the integration of expert human operators and advanced semi-automated systems.

Today, NASA engineers and contractors use many software tools to evaluate proposed designs for spacecraft, aircraft, and ground-control systems. Making these tools more intelligent and useful has the potential to allow the engineers to produce better designs in less time and at lower cost. HCC research will support intelligent design by combining several software technologies. The three fundamental technologies in an intelligent design tool are computer-aided design, data analysis, and optimization. These can be combined with other technologies—including automated reasoning, VR, data mining, data visualization, and neural networks—to create integrated, intelligent design tools.

Intelligent training systems can provide innovative and cost-effective solutions for the needs of the agency. First, amortized over the time a system is in use, computer-based training systems

[126] mshafto@mail.arc.nasa.gov.

are far less costly than human instructors. Furthermore, they provide a consistent, dependable "resident expertise" often difficult to maintain due to personnel attrition. There are likewise advantages to the trainee: the course curriculum (together with the pacing and presentation modes or media) can be tailored to the needs and preferences of the individual user. This saves the trainee time, alleviates boredom and unnecessary repetition, and ensures maximal learning effectiveness. The customization of training assumes greater importance as the agency acquires an increasingly diverse, heterogeneous workforce (for example, on the International Space Station).

Research in this domain focuses on applying advances in instructional science and technology, mission and vehicle simulation, and computer-based learning to meet agency-specific training requirements. Among the new technologies to be explored are more conversational, mixed-initiative interaction; web-based pedagogical browsers, and just-in-time training for remote distributed teams.

Immersive and virtual environments can provide an interactive capability for participants to intuitively and collaboratively explore complex, multidisciplinary simulations and data. This area has two components: The first includes a multimodal interface, which provides display of and control over complex 3-D data; these displays will use the visual, audio, and tactile senses. The other component is an extensible high-performance distributed software environment capable of integrating and co-registering time-varying data from a variety of sources, including computational simulation and experiment. This environment will enable the integrated and intuitive analysis of data by an integrated (though geographically distributed) team. Virtual environment technology extends the long-appreciated benefits for training, planning, analysis, and systems maintenance of aircraft simulation to a wide variety of new domains.

Human-user interaction with virtually any device imaginable may now be simulated in virtual environments for training, operational planning, or data visualization. However, the human is still a significant bottleneck limiting widespread, practical applications. Smooth, dexterous sensory-motor interaction that does not produce motion sickness and avoids untoward sensory-motor after-effects of extended use has still not been achieved. Virtual environment databases, which capture the level of detail present in the real environments that are simulated, are still awkward and time consuming to incorporate into virtual environment simulations. Overcoming these two major impediments will enable numerous NASA-specific virtual environment projects, such as *in situ* vehicle simulator training, telerobotics simulation and path planning, telemedicine, telesurgery, terrain visualization, atmospheric model visualization, guidance and training for mechanical assembly and maintenance, and visualization for virtual assembly.[127]

Task modeling and user modeling use intent inferencing and context understanding to tailor the information to the user, task, and available equipment. Technologies include information-needs models, dialog management, context understanding, and intent inferencing.

[127] See http://ic.arc.nasa.gov.

3.3.4.1 Information Needs Models

User models are being developed to improve the relevance of search results. These models can be used in conjunction with intelligent agents in the form of Enhanced User Needs (EUNs).[128] The combination of user models and intelligent filtering agents provides search results of large databases such as the Internet.

High-quality full-motion video images require high-capacity bandwidth. Such imagery has been used in video-mediated communication to emulate face-to-face communication. The visual cues provided by video-mediated communication include (1) gaze direction, (2) eye contact, (3) lip movements, (4) facial expressions and head movements, and (5) physical appearance.[129] Not all of these cues may be needed for effective performance. For example, H. Vons, R. Vertegaal, and G. van der Veer found no significant difference in problem solving among full-motion video with gaze direction, full-motion video without gaze direction, and still video with gaze direction. However, there was one significant difference in questionnaire responses: with only still video, it was harder to tell to whom a collaborative partner was talking. T. Kuro, N. Watanabe, S. Takano, and H. Kanayama developed a method to change mouth shape to match vocal speech.[130] They also identified the following as important to dialog: (1) "when speaking head movements are frequent, single blinks are typical, the interval between blinks is rather long" and (2) "when listening successive blinks are typical, one nods when approving, one tilts one's head when doubting, one shakes one's head when disagreeing."[131]

The University of North Carolina at Chapel Hill has been ranked top in the country for its computer graphics research and has been at the forefront for more than 30 years.[132] It is one of the five sites of the NSF's Graphics and Visualization Science and Technology Center. Research topics include flip-up head-mounted displays, see-through augmented-reality displays, volumetric displays, multi-user stereo displays,[133] image-generation hardware, modeling and simulation, low-latency viewer and object tracking, haptics, gesture-based interaction, six–degree-of-freedom controllers, anti-aliasing, automatic culling techniques, data fusion for augmented reality, telepresence, and post-rendering warping.[134]

The office of the future has a 12-person projection capability in which the images of persons at various locations can be projected onto four walls in each location. Sound is collocated with their projected image. The new item here is closed-loop calibration.[135]

[128] Sima C. Newell, "User Models and Filtering Agents for Improved Internet Information Retrieval," *User Modeling and User-Adapted Interaction*, Vol. 7, No. 4 (1997), pp. 223–237.
[129] H. Vons, R. Vertegaal, and G. van der Veer, "Mediating Human Interaction With Video: The Fallacy of the Talking Heads," *Displays*, Vol. 18 (1998), pp. 199–200.
[130] T. Kuro, N. Watanabe, S. Takano, and H. Kanayama, "A Proposal of Facial Picture Control Incorporating Several Essences of Conversation," *Systems and Computers in Japan*, Vol. 29, No. 7 (1998), pp. 57–64.
[131] Ibid., p. 59.
[132] Information-gathering meeting, 17 May 1999.
[133] Ibid.
[134] See http://www.cs.unc.edu/.
[135] See http://www.advanced.org.

Brent Seales identified two technical problems with telepresence: (1) calibration of the walls on which information is projected (a 10-year-old can do this in about 15 minutes) and calibration of the projectors (color is being matched with software) and (2) heavy requirements for bandwidth ($60,000 is spent on telephone lines and hardware alone in the consortium). Herman Tole identified the steps to overcome the calibration problem in the future: using a ceiling-mounted track system to provide set projector locations, using digital light projection to correct some of the optical distortions (such as the keystone effect), and using subliminal visual signals to keep the projectors calibrated. In addition, software algorithms are being developed to overcome the inconsistent colors in projectors and cameras.

Henry Fuchs identified two additional problem areas: (1) creating people-to-people telepresence and (2) creating people-to-information telepresence. The first requires matching the reality of person-to-person communication—that is, no time or shape distortion. Fuchs believes that the people-to-information interaction is much more difficult, since there is no underlying theory. Solving the distortion problem is done by computation. The computational cost depends on the geometric complexity of the background. The more complex the background, the greater the computational cost. There are geometrical distortions in the projector resulting from the assumption that each projector is a pinhole camera. Stretching and warping are needed. If the user is willing to tolerate a simple display surface, the computational cost plummets, since standard display algorithms can be used.

The minimum system is one projector and one camera. One projector will require a keystone correction (larger on top than on the bottom). Corrections are being developed both electronically and optically. Fuchs is developing software algorithms to correct this. The minimum useful configuration is two projectors and one camera. Two cameras are easy to triangulate, and two projectors blend the image properly.

Fuchs made several recommendations: First not using zoom, since the goal should be to make the system as natural as possible. The visual acuity of the camera-projector system matches the human eye—one arc minute. The technology for this is close to being developed. The resolution of the camera systems will still vary so that there will be one sweet spot. This highest resolution should be at head height. Second, not changing tilt, not even from session to session.

Projection rate is important for comfort level—more than 60 Hz is okay. Stereo displays have many difficulties (such as ghosting) for effective display. The instances when they are useful are few—in surgical procedures, for example. Holographic displays are good for simulating lenses for which there is not room. For example, Fuchs described an application of a light-emitting diode manufactured by Kopin that was projected onto and refracted through acrylic lenses to create a holographic image. This technique could be used to support augmented reality—for example, information overlays such as pagers and technical orders. Holographic displays can also create multifocus contact lenses in which every surface on the contacts has three different refractions.

Fuchs identified the problems with stereo: (1) it has to be calculated for every eye watching it and has to be right for everyone involved and (2) it has to be delivered to every single eye.

Polarization works for two persons but only if they are working at the same thing and are sitting close. A combination of private (calibrated to an individual's personal computer) versus public (calibrated to a group average) displays may solve the presentation problem. Private holographic displays in transparent glasses may also solve the presentation problem. Fuch's final recommendation was early deployment and use, to avoid unrealistic expectations.

To enable more realistic movement in a telepresence conference room, the University of North Carolina has developed software algorithms (single-constraint-at-a-time) and hardware (HiBall, a scalable tracking system for helmet-mounted displays).

The National Technology Alliance began in 1980 at the National Reconnaissance Office to close the gap between commercial and government information technology. The National Media Lab started in 1987, the National Information Display Lab started in 1990, and the National Center for Applied Technology was founded in 1997. Their mission is to empower government users to effectively and efficiently capitalize on technology emerging from commercial and consumer industry. The National Technology Alliance Technology Cycle begins with the user and then moves onto evaluations and technology assessment. This leads to research and development, creation of standards, and commercialization. The focus is on common problems that traverse many users and jobs. One example, the Imagery Display and Exploitation System, required greater resolution and higher reliability. ABP Metascan was interested in that resolution for radiology. This expanded the number of units in use from 1,000 in intelligence applications to 600,000 in use by radiologists. Orwin then developed a 5-million-pixel display. There are now four other manufacturers.

The Joint Operations Visualization Environment (JOVE) focuses on visualization for situation awareness. The JOVE motto is "The greatest thing is to get the true picture, whatever it is" (Winston Churchill). It provides both big picture and drill-down to get specific information. The system uses MIL-STD-2525A symbology. JOVE provides an intuitive presentation of a Common Operating Picture in four dimensions. It has two types of data—geospatial and relational. There are four configurations, from the boardroom system to an overall operational control to a portable control center and an airborne or truck-mounted system to remote access visualization on a laptop.

The Noise Robust Voice Control System was developed for use in a command environment. It enabled differentiating untethered multiple speakers.

John Riganati demonstrated iris recognition for security access control. He described a biometric network security using iris-based identify verification. This technology is being applied for automatic teller machines and for e-commerce. The literature indicates that the iris is invariant from age 6 months to death. A natural extension is to use goggles to create a "SCIFless" (sensitive compartmented information facility) in which a cleared person gets the information he or she needs to know. This identifies the right person. It could be married with SIREN, which would identify the right information for that person. The User Tailored Information Service is another project that is just beginning. It tailors simple systems, such as repetitive actions of reviewing logistics status, into user-unique, simple actions.

Barbara Connolly demonstrated ultra-resolution displays. One of these, the System Technology for Advanced Resolution, will be used in CPoF. John Fields stated that large ultrasound displays sound should (1) not show tiling, (2) have separate displays, and (3) be scalable from 2 to 30 feet at any aspect angle.

Advisable Planning determines the commander's intent and develops alternative courses of action in terminology that commanders can understand.[136] Visage is used for the interface. One tool, the Bed Down Critic, identifies inconsistencies and suggests changes. The complete Advisable Planning system guides the planner with high-level advice and understands the characteristics of alternatives. This demonstration was impressive but was seen as too immature to be used by the Air Force.

3.3.5 Dialog Management

Dialog management is a major issue in a work environment that includes the use of a wide range of databases, information domains, forms of analysis, planning, and command and control methods. As indicated earlier in this report, Broadsword provides a network-based infrastructure to support work in a domain with these features. The DARPA-sponsored HPKB also addresses technologies that involve dialog management in a large-scale (millions of bits of information with 100,00 axioms), diverse knowledge-base environment. The goal of the HPKB project is to produce technologies for developing very large, flexible, reusable knowledge bases. As information is extracted from different sources, knowledge-base technology is needed to *semantically integrate* meaning as this information focuses on a current situation and set of problems to be solved. It has been shown that pairwise integration does not scale; at best the aggregated systems evolve to suboptimal stovepiped systems. Teknowledge has approached the semantic integration problem by defining formal semantics for input and output across applications and knowledge bases used in the HPKB project, including inputs for a user working a problem in some domain. Dialog management begins with a template-based interface into which user-specified parameters can be inserted. Teknowledge's formal semantics are used in conjunction with two natural-language components, START and TextWise, to transform the input query into a legal Cyc query as a means of creating new knowledge for application to the user's problem. A related but different strategy for semantic integration was used in the SAIC integrated knowledge environment developed as part of the HPKB project.

3.3.6 Context Understanding

Context understanding and maintenance are important functions for the CPoF, since many of the human-machine interactions that occur there will be either incomplete or ambiguous when interpreted in isolation. In some cases, components such as language understanding will use contextual information to disambiguate the commanders' statements. In other cases, the Dialog Manager itself may play a role in interpreting the users' intent in complex interchanges. The Dialog Manager is the maintainer of the command post's processing context. It tracks people in the command center (their location), is aware of their roles on the staff team, and monitors their activities (working, resting, in a meeting). The dialog manager will also be knowledgeable about the capabilities of components such as the Battlespace Reasoning Manager. It will be capable of appropriately delegating tasks to other components when the users need access to planning, analysis, or simulation data.[137]

3.3.7 Intent Inferencing

Microsoft[138] is researching reasoning and intelligence, using Bayesian inference to exploit knowledge bases.

Different approaches have been developed to infer actor intent relative to the state of a work problem or to an application program, such as a work processor. Classical production system, neural net, and statistical-based mechanisms have been exploited in the computational architecture for a system to infer actor intent in order to provide context-sensitive support. Development of intent-inferencing technology initiated under the Air Force–DARPA Pilot's Associate program used a plan-goal-graph data structure and script-based reasoning to infer pilot intent. The system assessed multiple (potential) prestored plans based on event data to determine the active plan and to base decisions of pilot intentionality on state data relative to the active plan.[139, 140, 141, 142, 143] Inferred intent is used to select the presentation timing and representation of information to the pilot, as well as for making suggestions about types and forms of automated support to improve mission execution. More recently, E. Horvitz and colleagues have developed techniques to infer intent based on user-produced, free-text queries.[144] Their system uses probabilistic knowledge bases for interpreting user intent. Related work addresses intent inferencing for display management to better support time-critical decision-making.[145] Multi-attribute utility theory and Bayesian models of user beliefs are used to infer intent and use this knowledge as a basis for selecting information for presentation. Current research at AFRL/HE is investigating a third approach that infers intent based on a model of the situation awareness of the actor. This approach combines the use of a task network model, situation awareness mode, mental workload model, and human information process model with fuzzy logic, knowledge-based reasoning, and statistically based Bayesian belief reasoning to infer user intent. This knowledge is then used to adaptively modify the form, content, and modality of information—

---

[136] Information-gathering meeting at DARPA, 19 May 1999.
[137] http://www-code44.spawar.navy.mil/cpof/private/techpages/dialog.html.
[138] Information-gathering meeting at Microsoft, 16 April 1999.
[139] N. D. Geddes, "Intent Inferencing Using Scripts and Plans," *Proceedings of the First Annual Aerospace Applications of Artificial Intelligence Conference* (1985), pp. 160–172.
[140] N. D. Geddes and J. M. Hammer, "Automatic Display Management Using Dynamic Plans and Events," *Proceedings of the Sixth Symposium on Aviation Psychology* (30 April–4 May 1991).
[141] C.W. Howard, J. M. Hammer, and N. D. Geddes, "Information Management in a Pilot's Associate," *Proceedings of the 1988 Aerospace Applications of Artificial Intelligence Conference*, Vol. 1 (1988), pp. 339–349.
[142] Sewell et al., 1987.
[143] V. L. Shalin and N. D. Geddes, "Task Dependent Information Management in a Dynamic Environment: Concept and Measurement Issues," *Proceedings of the 1994 IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 3 (1994), pp. 2102–2107.
[144] D. Heckerman and E. Horvitz, "Inferring Informational Goals From Free-Text Queries: A Bayesian Approach," *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* (1998), pp. 230–237.
[145] E. Horvitz and M. Barry, "Display of Information for Time-Critical Decision Making," *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (1995), pp. 296–305.

visual, audio, or haptic—delivery to the user.[146] Simulation-based performance tests of this technology have been planned but remain to be executed.

All current approaches to intent inferencing incorporate mechanisms that are used to understand the problem context. The inferencing system reasons about current events, systems (for example, a weapon system), and actor data streams that are relative to some form of a domain model. N. D. Geddes uses a Plan-Goal_graph, E. Horvitz uses an attribute model, and S. S. Mulgund and G. L. Zacharias use a Bayesian belief net. Activities relative to the domain model are used to infer the user's intent.

[146] S. S. Mulgund and G. L. Zacharias, "A Situation-Driven Adaptive Pilot/Vehicle Interface," *Proceedings of the 3rd Annual Symposium on Human Interaction with Complex System* (1997).

Chapter 4: Collaboration

Collaboration technologies enhance the interaction between the decision makers and the JBI and also better interaction among decision makers themselves. Technologies include sharing, advanced white boarding, domain-specific workflow management, mixed-initiative systems, facilitation, and group interaction devices.

4.1 Sharing

Nick Flor proposed[147] four structures to collaboration: task, system structure, modifications, and system behavior. Collaborators either push or pull information among themselves to develop common representations of these four elements. Flor's theory is based on observation of two persons collaborating on a maintenance task.

Mark Young[148] provided a list of requirements for collaborative visualization: provide (1) the same data, same time, same view, and same aspect in geospatially referenced, dynamically updated object visualization; (2) object interaction interfaced to back-end services for processing support; (3) data layering and layer visibility control; (4) adjustable fidelity with continuous level-of-detail management; and (5) 2-D or 3-D whiteboard annotation support. He selected PI 3-D Virtual Whiteboard. Its attributes are a 100 percent pure Java2, web-based, client-server architecture; multiple clients per collaborative session; multiple sessions; centralized, federated visualization data servers; and clients to connect to servers and join session-supporting operations.

The Cspace project is developing techniques and tools to support a wide range of long-duration information-intensive collaborations, with emphasis on helping teams organize and manage their shared information and on helping collaborators manage their attention. Awareness of the Cspace infrastructure is intended to provide a common set of advanced collaborative services for tools that may not have been designed for collaborative use, including familiar single-user productivity tools such as Microsoft PowerPoint. Tools supported in the present prototype include the Windows NT file system, an outliner and whiteboard, and Microsoft PowerPoint. The infrastructure potentially can support any tool that has a suitable application program interface. Services include easily adjustable information awareness, fine-grained versioning and history-keeping, and a modeling capability that enables relationships to be defined and evolved among the parts of heterogeneous information objects. Additionally, there is a rich scheme for annotation, messaging, and linking, as well as a facility for structural differencing that provides visual awareness of changes or differences in shared objects such as file areas, presentation documents, and outlines.

Besides infrastructure development, the project is exploring applications that include business decision making (as part of the IC&V program), military command posts (as part of the CPoF program), and multilingual information management (as part of the Threat/Intelligence Data

[147] Nick V. Flor, "Side-by-Side Collaboration: A Case Study," *International Journal of Human-Computer Studies*, Vol. 49, No. 3 (1998), pp. 201–222.

[148] Information-gathering meeting at GTE, 15 April 1999.

Extraction System [TIDES] program). The project is also exploring support for software engineering teams.

The Cspace infrastructure is based on two key ideas: First, a common representational "fabric" model is used to manage fine-grained, shadow representations modified in subsequent versions for application-specific objects, as well as the structural models that relate to them. Second, an event-based scheme is used to maintain consistency among diverse representations and to provide awareness and messaging support for users. Events in this scheme are "situated" with respect to parts of the shared assets and models.

Information awareness is a principal concern. Participants in collaboration have different requirements for their awareness of changes to shared information, messages from other collaborators, and other notifications. These requirements for awareness may change rapidly with role, time, and task. For example, team managers may have a high need for awareness of member activities, which, for example, may increase prior to deadlines or meetings. Techniques being developed enable collaborators to effectively manage awareness levels when there is intense competition for their attention. In addition to developing technical concepts and research software, the project is undertaking behavioral evaluation using a combination of direct instrumentation, outcomes analysis, and observation and interviews. A major analytic field study of collaborative teams is being undertaken (involving more than 200 Master of Business Administration students), as well as several special-purpose behavioral experiments. Published behavioral results have given insight on phenomena such as information overload and shared mental models. The evaluation will also enable analysis of attention-management strategies, information retention, consensus formation, and roles of individuals in groups.

R&Tserve is a collaborative workspace for authors. It includes graphics support, a transaction archive, comments utility, help, automated table generation, and e-mail.[149]

Doug Olkein[150] described GTE's Info WorkSpace (IWS), a virtual online meeting place with data sharing. Communication is provided with desktop conferencing (asynchronous and real-time), distance learning, and mass briefing. IWS is a knowledge management search tool; it has registered user expertise and access to external intranet or Internet resources. It includes Microsoft, Placeware, Netscape, Databeam, and GTE products. The IWS toolbar has the following features: a whiteboard, a file cabinet, external conferencing, video, shared text, discussion groups, and a bulletin board. Supporting features include security, navigation aides, online help, user Rolodex, calendar, mail, and search. IWS also offers a one-to-one chat feature. The time to download applets is long. Therefore there is a low-bandwidth IWS. Initial release is in JEFX99. In the future, network administrators will be able to register applications; IWS would form the software backbone for integrated operations.

NASA accomplishes its work in distributed, multidisciplinary teams in a variety of public and private organizations. Development of efficient and effective design practices, data analysis,

[149] S. Abrams. "Web-Based Collaborative Publications System: R&Tserve," *Sixth Alumni Conference of the International Space University* (1997), p. 130.
[150] Information-gathering meeting, 15 April 1999.

mission monitoring, and control is possible through networked and portable computing and communication tools. Support systems for scientists and engineers are now being designed using model-based techniques for representing data, theories, devices, and operations. To move to the next generation of tools, to those that truly enhance collaborative performance, cognitive task analysis must be extended to integrate new forms of human-machine interaction and human cooperation across organizations. This especially requires understanding of how people formulate and share representations across disciplinary boundaries. Once researchers better understand the basic nature of interaction among human experts and intelligent software agents, a new generation of collaborative tools for science, system design, and mission operations can be built. Areas of particular applicability for these collaborative tools are the International Space Station, mission-critical software development, and ground-space (or, in aeronautics, surface-air) operations.

NASA Ames[151] is developing a set of intelligent collaboration and assistant systems. Postdoc is a multi-user, web-based application primarily for the storage and retrieval of documents. The Aviation Performance Measuring System[152] is a prototype for acquiring, analyzing, and interpreting data from flight data recorders on commercial aircraft. ScienceDesk[153] is a collaboratory system to assist scientists in performing distributed scientific work within geographically dispersed teams. It includes intelligent tools to control scientific hardware; to plan, conduct, and monitor working experiments; to store and index data sets; to develop and share scientific software models; and to support the overall scientific process. Intelligent Mobil Technologies[154] is producing portable computer systems that employ RF-based remote networking and intelligent software agents to users in remote locations. Another program is Distributed Intelligent Agents for Information Management and Sharing.[155] It supports dynamic and flexible organization of personal information repositories, distributed over the World Wide Web and sharable by multiple users. The repositories can be shared among persons with similar interests. Software agents do automatically discover new relevant information. Brahms is a multi-agent framework for modeling work practice. It identifies how information is shared, how social knowledge affects participation, which problem-solving methods are employed, and work quality.

Microsoft[156] stated that e-mail is evolving. Exchange Platinum will have partitioning, load-balanced clustering, native message standards, an active directory, enhanced workflow, Windows 2000 platform integration, Office 2000 integration for collaboration, and unified real-time, wireless messaging. The vast majority of collaboration work at Microsoft is on e-mail. Office 2000 goals are to have all products web-enabled, to embrace and extend industry standards, to design for the global user, to improve information sharing, and to reduce maintenance costs. The knowledge management product vision is to connect “the right people

151 Information-gathering meeting, 15 April 1999.
152 See http://ic.arc.nasa.gov/ic/projects/apms.
153 See http://ack.arc.nasa.gov:80/ic/projects/scidesk/index.html.
154 See http://ic.arc.nasa.gov/ic/projects/WNE/.
155 See http://ic.arc.nasa.gov/ic/projects/aim/current.html.
156 Information-gathering meeting at Microsoft, 16 April 1999.

and the right information through extensions for familiar business tools." Internet Explorer 5 provides collaboration and applicable sharing.

CollaborativE Video Analysis is a software tool that supports simultaneous video protocol analysis by multiple users. It is being developed at the University of Canterbury in Christchurch, New Zealand. The system enables both synchronous and asynchronous collaboration, synchronous multithreaded event logging, an animated direct manipulation interface, symbolic notation and visualization at different levels, quantitative analysis such as event counts and duration, event search, and reordering of video segments.[157]

The Naval Surface Warfare Center has developed a methodology to define user requirements for collaborative tactical computer interfaces. The methodology is called the Tactical Information GUI Engineering and Requirements Specification;[158] it has been applied to the redesign of the Naval Space Operations Center.

Susie Iacono[159] stated that there is not much research in group decision making and group decision-support systems becuse the highly structured, room-based, brainstorming systems previously developed did not work well. Now research has turned to team collaboration on the Web. What works best are the simplest technologies that are available to everyone, such as e-mail systems and web-based conferencing systems. However, users need to know what has happened since the last time they logged on. Hsinchun Chen of the University of Arizona has been developing 2-D and 3-D visualization to portray the current state of the knowledge. Learch and Crote of CMU are working on similar efforts. But people do not use these systems in the way expected or do not use them at all. Groups like to communicate naturally rather than in a highly structured way. It is also important for social structure to naturally emerge. A key is cooperation.

Susie Weisben of the University of Arizona showed that there are different ways to act in a team, and she is working on development of ways to portray the critical information. John Candy of the University of California at Berkeley is developing robots to provide physical presence to support people in dispersed locations. Issues being addressed are what kind of social interaction should these robots have and how to maintain visual memory of the remote space. Patrick Perona of Cal Tech is developing virtual characters to support tasks. Georgia Tech has the Classroom 2000 program, in which every bit of information is captured during the class—lecture notes, interactions, and video. Issues are storage and retrieval problems.

VR is a new start at DARPA to create a high-resolution environment so that distributed persons perceive that they are in the same environment.[160] Things considered are Dick Urban's

---

[157] A. Cockburn and T. Dale, "CEVA: A Tool for Collaborative Video Analysis," *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work* (1997), pp. 48–49.

[158] J. A. Bohan and D. F. Wallace, "Team Ergonomics and Human Engineering Methods for the Design of Collaborative Work Environments: A Case Study," *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (1997), pp. 1066–1070.

[159] Head of six program areas: information and data management, human-computer interaction, knowledge and cognitive systems, computation and social systems, robotics and human augmentation, and special projects.

[160] Information-gathering meeting at DARPA, 19 May 1999.

holographic glasses, CMU's video image combination, and USC's capturing of facial expressions and key movements to create avatar-like heads.

The Evolutionary Design of Complex Software is a joint AFRL/IF and DARPA program. The aim of the program is to develop technologies needed to support continuous evolutionary development of software systems for military weapon systems. A major goal is to create the ability to make the time and cost of making incremental changes to a large-scale software system proportional to the size of the change, as opposed to the size of the system. Individual technology development efforts are clustered into five areas: Architecture and Generation; Rationale Capture and Software Understanding; Information Management; High Assurance and Real-Time; and Dynamic Languages. Seventy-three projects are included across these five areas.

In the Information Management area, the Atlantis project is addressing workflow in a distributed collaborative environment. The goal is to devise new paradigms for representing processes to determine means by which the distributed software environment may assist teams of users in carrying out processes, and to discover mechanisms that permit in-progress processes to evolve compatibly. It is generally agreed that transaction models are inadequate for long-duration, interactive and cooperative activities. To address this issue, the Atlantis project is developing a transaction management component. It provides primitives for defining project-specific concurrency control policies. Another aspect of the workflow problem derives from the fact that large-scale software development often takes place across several independent organizations. As a result, independent entities wish to guard their own proprietary processes and tools while sharing data and process output (within security constraints). The Atlantis project is working on this problem by developing a model for "cooperating software processes."

Orbit is another project in the Evolutionary Design of Complex Software program that has developed a computer-supported collaborative work environment. It is attempting to leverage recent sociological theory on the nature of work to produce the next-generation computer-supported collaborative work system. One of its unique features is the use of Elvin, a scalable, distributed publish-subscribe event bus that supports content-based subscription. In contrast, the Common Object Request Broker Architecture uses a channel-based approach that results in all subscribers' receiving every event posted to the channel. Another distinguishing feature of Orbit is the Development of Courtyards. A Locale Service manages user sessions for groups. Courtyard provides a method to allow connection between Locales. Thus, all objects placed in a Courtyard are equally visible and accessible to the members of the included Locales. Furthermore the Orbit user interface services make it possible for a user to participate in multiple ongoing collaborations simultaneously, with the freedom to vary the level of interactivity as appropriate.

One notable accomplishment is that Orbit was used to build a collaborative environment for "distributed intelligence-gathering teams." The system was produced in less than a month. The developers believe that an equivalent stovepiped solution would have taken several years to complete.

AFRL has initiated a project to produce a Collaborative Enterprise Environment (CEE) that is aimed at reducing the time and cost involved in developing, testing, and fielding new military weapon systems. The CEE is a distributed virtual environment that supports the collaborative use of analysis, engineering design, and cost models along with systems, engagement, and campaign simulations to design and test system concepts virtually within a comprehensive operational context. The design emphasizes developing the product and process interactively. The CEE virtual environment includes connectivity and exploitation of the World Wide Web.

The CEE consists of decision support systems, resource browsing and assembly tools, and a "plug and play" communication infrastructure. Some important features include a Web-based user interface and interactive infrastructure; explicit process models for analysis, engineering design, and work domain business rules; and enterprise common object models. TANGO Interactive™, developed under DARPA sponsorship by researchers from Syracuse University, provides a candidate Java-based Web collaboratory system for the CEE. It provides utilities for setting up electronic communities provided with multimedia interaction tools. Video on Demand is a related project of the Northeast Parallel Architectures Center at Syracuse University. The goal of this effort is to produce a searchable video-on-demand system that supports user queries for video clips and an efficient video retrieval capability. The design employs the metadata concept and strictly partitioned continuous video data from metadata. Metadata provides descriptive information about the video that is stored in a database. The system supports both category-based and content-based queries. In a category-based query, an attribute of the video clip, contained in the metadata, is entered as a search term. Content-based searching involves a query formed on the basis of either a content-based data field or content descriptors that are matched to clip titles and annotations. All queries are entered through a Web browser. One distinguishing feature of this system is that video playback continues independently on the Web browser after the video client links with the server.

The Enterprise Common Object Model concept in the CEE is an attempt to establish well-formed, cross-cutting relations among a heterogeneous set of data generators and data users. An enterprise object is formed that can meet the needs of multiple users in different work domains, ranging from analysis to design to operations. For example, a satellite sensor can be used to produce a digital terrain image that is needed by a ground station, which, in turn, produces potential targets, based on "anomalies" in the terrain data. Anomalies meeting certain criteria may be used by the Theater Battle Management Core Systems in the development of an ATO. A bomb damage assessment report from an assigned aircraft may then provide probability-of-kill data to be included in the Enterprise Common Object. Different users can call on these common objects to support their unique work requirements.

Collaborative Virtual Workspace (CVW), developed by MITRE, provides a software-based medium to support temporally and geographically dispersed work teams who must synchronize their work in a variety of ways. It incorporates audio and videoconferencing capabilities along with document sharing and chat room features. In connection with collaborative tools like Microsoft NetMeeting, CVW provides a *persistent virtual workspace* for the use of applications, documents, and interpersonal interactions. CVW is structured as a virtual building consisting of floors and rooms, with each room providing a context for communication and application or document sharing. Because rooms, once established, persist, there is no requirement to set up network-based sessions or to know the location of users. CVW builds on work from Stephen White of the University of Waterloo and Pavel Curtis of Xerox PARC. IWS, the commercial version of this technology, is available from GTE.

CVW has been in use in MITRE as a prototype system for the past few years. An evaluation of CVW use over a 6-month period was conducted by Jane Mosier et al.[161] This study reviewed 6 weeks of data logs as a means of learning use patterns. In addition, five case studies were completed as a means of relating use to different types of work teams. In general, work teams exploited the features of the CVW that were most easily integrated into existing work processes. Some features, such a audio and videoconferencing, tended not to be used because alternatives already existed, and the required network infrastructure to support this functionality within CVW typically was not available for all team members. For this and other reasons, therefore, this evaluation was somewhat limited.

Of 196 users surveyed, 66 issued fewer than 11 communication commands during the sampled period. Thus, they were either passive listeners or inactive users of the system. The majority of more active users have maintained accounts for the system (sampled 9 months after the use survey), which provides a crude measure of perceived value. The most consistent finding from the case studies was that CVW appeared to be most useful for (1) providing a discussion area for quick and short-lived topics, including communication on topics and items that a person generally would not take the time to convey via e-mail or other means, and (2) quickly becoming current on what is happening in the project or office after being out of contact. The rooms provided a basis for rapid, synchronized discussions. Both group and private conversations are supported. A scrollback feature for a room provided the ability for a person who had been offline to quickly regain context and knowledge of the current state of work. Another important function: it allowed the team member to pick up little events that others might forget to mention." An abundance of casual conversation interspersed with more focused material, however, tended to interfere with the ability to effectively use scrollback. In general, some found CVW to be more convenient than separate e-mail, chat rooms, and document sharing tools; others did not.

4.2 Advanced White Boarding

Siewiorek[162] developed the C-130 Help Desk to provide technical support to Air National Guard and Reserve aircraft maintenance specialists. Maintainers working in a hangar or on the flight line request information from a single help desk. The sergeant at the help desk views what is on the requesters' displays, then manipulates the displays remotely to demonstrate the correct procedures for accessing the appropriate data. A similar system was developed for F-15s.

---

[161] Jane N. Mosier, T. L. Fanderclai, and K. K. Kennedy, *An Evaluation of CVW Use at MITRE.* MITRE Technical Report, 1998.

[162] Demonstration at Carnegie Mellon University, 17 March 1999.

A Mobile Communication and Computing Architecture system was developed to provide just-in-time information for mobile users. The system is a wearable computer that enables service engineers in the field to collaborate synchronously and asynchronously. The system mobile engineers share and build corporate memory by accessing information from multiple sites and while commuting. The system includes voice bulletin boards, video clips, and maintenance databases.[163]

Itsy, a prototype Compaq computer the size of a cigarette pack, will make collaboration easier. It is being used to process data at the user side and thus reduces the amount of information that must be transmitted. Itsy enables collaboration of disparate users.

The Center for Strategic Technology Research has developed an immersive environment, the Insight Lab, that uses barcodes to link paper and whiteboard printouts to multimedia stored in a computer.[164] The lab includes linked sticky notes, data reports, and electronic whiteboard images. Input is from voice commands, a wireless mouse, a wireless keyboard, and a barcode scanner. Information is conveyed via displays, tackable walls, an electronic whiteboard, and layered whiteboards. The CPoF will also use whiteboards.[165]

4.3 Domain-Specific Workflow Management

Workflow is "the sequence of actions or steps, in sequential or parallel arrangement that compromise a business process. An automated workflow is the workflow that is integrated with enabling information technology."[166]

One form of domain-specific workflow management is intelligent HCI. "An intelligent interface is one that provides tools to help minimize the cognitive distance between the mental model that the user has of the task and the way in which the task is presented to the user by the computer when the task is performed."[167] An intelligent interface has five components: domain-specific, domain-adaptation, dialog, presentation, and interaction toolkit. This categorization is known as the ARCH model. C. Kolski and E. LeStrugeon stated that there are five types of intelligent interfaces (from lowest to highest intelligence): flexible interface, human error–tolerant interface, adaptive interface, assistant operator, and intelligent agent.[168]

SPAWAR uses of Gensym's G2 Intelligent Systems for Operations Management: DARPA sponsored the developed of a Team-Based Access Control system for application in patient care. The system included a "hybrid access control model that incorporated the advantages of broad, role-based permissions across object types, yet required fine-grained, identity-based control on

[163] A. Smailagic, D. Siewiorek, A. Dahbura, and L. Bass, *MoCCA: A Mobile Communication and Computing Architecture.* Pittsburgh, PA: Carnegie Mellon University, 1999.

[164] B.M. Lange, M.A. Jones, and J. L. Meyers, "Insight Lab: An Immersive Team Environment Linking Paper, Displays, and Data," *Proceedings of the Conference on Human Factors in Computing Systems* (1998), pp. 18–23.

[165] See http://www.darpa.mil/iso/cpof/ for additional information.

[166] T.A. Nassif, *Supporting the Fleet: Taking Workflow to the Waterfront.* Monterey, CA: Naval Post Graduate School, 1995, p. 6.

[167] C. Kolski and E. LeStrugeon, "A Review of Intelligent Human-Machine Interfaces in the Light of the ARCH Model," *International Journal of Human-Computer Interaction*, Vol. 10, No. 3 (1998), p. 193.

[168] Ibid., p. 206.

individual users in certain roles and to individual object instances."[169] The focus was on team collaboration and control of workflows.

The Europeans have developed the Workflow on Intelligent and Distributed database Environment system—a conceptual model that includes "an organizational model as a basis for task assigned proposed for the project, advanced functionalities for exception handling, the concepts of multitask and supertasks for workflow modularization, and integrated transactional semantics."[170] They are also applying workflow management to the telecommunications business and have developed an architecture for this application. The architecture is composed of presentation blocks, function blocks, and data blocks.[171] On the basis of their ongoing efforts, M.C.A. Van de Graaf and G. J. Houben developed design guidelines.[172]

GTE[173] manages workflow by monitoring which software systems are being used. If a system is not used, it is ripped out; systems that are being used are modeled to identify efficiency enhancement.

The goal of the Planning and Decision Aids program[174] is to determine how to get courses of action to the commander in minutes rather than days. People are slow and make errors. Computers lack insight. The Planning and Decision Aids program has a family of tools for generative planning (Multiagent Planing and Visualization (MAPVIS) and System for Interactive Planning and Execution (SIPE4I)) and case-based planning (Joint Assistant for Deployment and Execution (JADE)), scheduling and resource allocation (airlift planning), workflow and process management, and mixed initiatives (Special Operations Flight Planning System (SOFPlan), TRIPS). The metrics to evaluate these technologies are planning speed, quality (rewards and risk, time and resource required/used, simplicity, flexibility), and understandability.

The JADE support system for TPFDD planning is the result of merging two technology integration experiments accomplished under a joint AFRL/DARPA program. The support system consists mainly of three different technologies: case-based reasoning, parallel structured search and retrieval, and generative reasoning and learning. When combined, these technologies provide the infrastructure to derive and support a mixed-initiative interface for an interactive planning system. JADE and its predecessors have been demonstrated at several military exercises.

As a mixed-initiative planning system, the GUI for JADE supports different ways for the user to form a query on the database and provides a way for the intelligent system to make appropriate

---

[169] R. K. Thomas, "Team-Based Access Control (TMAC): A Primitive for Applying Role-Based Controls in Collaborative Environments," *Proceedings of the Second ACM Workshop on Role-Based Access Control* (1997), p. 13.

[170] F. Casati, P. Grefen, B. Pernici, G. Pozzi, and G. Sanchez, *WIDE Workflow Model and Architecture.* Enchede, Netherlands: Technische University, 1997.

[171] W. Nijenhuis, W. Jonker, and P. Grefen, *Supporting Telecom Business Processes by Means of Workflow Management and Federated Databases.* Enschede, Netherlands: International Institute of Technology and Management, 1997.

[172] M. C. A. Van de Graaf and G. J. Houben, *Designing Effective Workflow Management Processes.* Eindhoven, Netherlands: Eindhoven University of Technology, 1996.

[173] Information-gathering meeting, 15 April 1999.

[174] Information-gathering meeting DARPA, 19 May 1999.

suggestions for actions to modify a plan to meet the current situation. This is made possible because the generative reasoning and learning technology uses derivational analogy as a method to capture lines of reasoning used in prior plan development that can provide a rationale for why certain plan modification may be needed in the current case.[175, 176] The planner can propose a case and suggest modifications based on the automatic input of a request derived directly from the commander's guidance, or the user can make tailored queries to initiate interactive work with the support tool. Queries can be formed at different levels of specificity. The basic query development process is template based. Given the user's case selection, the support agent provides plan modification issues and suggestions in a dialog window. This may involve bringing in information from other cases. This form of mixed-initiative interaction continues throughout the construction of individual force modules until a complete TPFDD is produced.

In addition to demonstrating several underlying artificial intelligence technologies useful for planning support, the JADE project has nicely illustrated the type of technology blending that is needed to support context-relevant and intelligent mixed-initiative work between a human user and the support technology. To date, the majority of the research effort has focused, however, on developing and integrating the individual reasoning and information-retrieval pieces of the system. More work is needed to include user modeling and enhanced task modeling to support a more robust mixed-initiative interface capability.

4.4 Mixed-Initiative Systems

"A mixed-initiative system is one in which both humans and machines can make contributions to a problem solution, often without being asked explicitly."[177] Mixed-initiative planning systems are being designed to exploit the strengths of humans and computers. "Humans are still better at formulating the planning tasks, collecting and circumscribing the relevant information, supplying estimates for uncertain factors, and various forms of visual or spatial reasoning that can be critical for many planning tasks. Machines are better at systematic searches of the spaces of possible plans for well-defined tasks, and in solving problems governed by large numbers of interacting constraints. Machines are also better at managing and communicating large amounts of data."[178] M. H. Burstein and D. V. McDermott identified key issues in mixed-initiative

[175] Jaime G. Carbonell. "Derivational Analogy: A Theory of Reconstructive Problemsolving and Expertise Acquisition," in R. S. Michalski, Jaime G. Carbonell, and T. M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach*, Vol. II. Morgan Kaufmann, 1986, pp. 371–392.

[176] M. Veloso, A. M. Mulvehill, and M. T. Cox, "Rationale-Supported Mixed-Initiative Case-Based Planning," *Proceedings of the Fourteenth National Conference on Artificial Intelligence and Ninth Innovative Applications of Artificial Intelligence Conference* (1997), pp. 171–179.

[177] Jaime Carbonell, cited in M. H. Burstein and D. V. McDermott, "Issues in the Development of Human-Computer Mixed-Initiative Planning," in Barbara Gorayska and Jacob L. May (eds.), *Cognitive Technology: In Search of a Humane Interface.* New York: Elsevier Science, 1996, p. 285.

[178] M. H. Burstein and D. V. McDermott, "Issues in the Development of Human-Computer Mixed-Initiative Planning," in Barbara Gorayska and Jacob L. May (eds.), *Cognitive Technology: In Search of a Humane Interface.* New York: Elsevier Science, 1996, p. 286.

planning systems. For search control management, the issues they listed[179] are control dialogs to establish collaborative patterns, variable speed and resolution response, decoupling and recombining plans, context registration, intent recognition, and plan analysis. Key issues in the representation of plans and plan-related information sharing are shared representations, abstractions, visualizations, uncertainty, versioning, author tracking, and change authority. Issues for plan revision management include maintaining continuity between plan versions, run-time replanning, and coordinating multi-agent planning tasks. Planning under uncertainty is a major issue in itself. Learning issues are user preference, prior plans and their effects, and general and domain-specific planning knowledge or heuristics. Interagent communications and coordination issues are distributed information management and maintenance of and timely access to shared plans. These authors stated that the important research areas are dialog-based task management, context registration, flexible and interactive visualizations, and information acquisition and management.

V. S. Subrahmanian of the University of Maryland described relevant programs.[180]

- Uncertainty management—There are three types of uncertainty (data, temporal, and spatial). The last two have large problem spaces; the first does not. ProbView is a query language that accommodates data uncertainty only. It was expanded to handle the other two types of uncertainty in temporal-probabilistic databases. The next step was the development of probabilistic object bases to handle storing object rather than relational databases. Probabilistic object bases are under development with funding from DARPA.
- Heterogeneous data or software access—This is an extension of the DARPA Integrated Intelligent Interfaces program. It includes a mediator (a program that integrates multiple databases). WebHermes is a platform for creating mediators for different applications. WebHermes (Heterogeneous Reasoning and Mediator System) includes two parts: (1) software integration enabling access to the software's external foreign functions and (2) semantic integration to logically merge data from multiple sources. Hermes provides a simple language to do both.
- The Interactive Maryland Platform for Agents Collaborating Together—An agent that should be able to build on any other piece of software and provide a valuable service. This platform enables agent collaboration.
- Multimedia databases and presentations—Multimedia content indexing and retrieval was developed to retrieve media objects from multiple sources by similarity. In addition, the Collaborative Heterogeneous Interactive Multimedia Platform was developed to present multimedia data. The platform is a framework for creating a living, dynamically updateable media presentation using queries.

Applications include logistics for the Army Logistics Integration Agency, Boeing's study of controlled flight into terrain, and U.S. Army Technical and Engineering Center missile siting.

Dr. Phil Emmerman of the Army Research Laboratory described efforts to reduce the footprint of the Tactical Operations Command from 50 persons to 8.[181] The new Command should have (1) extensibility across battle function areas, API applications, and layered battle function area–

[179] M. H. Burstein and D. V. McDermott, "Issues in the Development of Human-Computer Mixed-Initiative Planning," in Barbara Gorayska and Jacob L. May (eds.), *Cognitive Technology: In Search of a Humane Interface.* New York: Elsevier Science, 1996, pp. 289–295.
[180] Information-gathering meeting at the University of Maryland, 20 May 1999.
[181] Information-gathering meeting at the Army Research Laboratory, 20 May 1999.

specific applications; (2) scalability from corps to platforms, responsiveness, and fidelity; and (3) adaptability to handle information dynamics associated with new dynamics and information sources, database schemas, and situation-specific procedures. The reduction in footprint as well as the advances in information technology will result in changes in battle function areas. Thrust areas are

- dynamic environment with level of detail, tactical entities and features, multiresolution terrain, weather, and nuclear, biological, and chemical warfare
- multimodal human-computer interface
- loosely coupled 2-D and 3-D: both are needed to see the environment (3-D) and yet not get lost (2-D)
- multiresolution analysis
- software agents for monitoring, altering, retrieving, dissemination, and fusion

Intelligent systems applications are global or local adaptive view; responsive Tactical Operations Command–platform coupling; integrated distributed sensing, targeting, and engagement; multiresolution analysis with physics-based models for sensing, planning, and execution; Army Battle Command System or legacy system mediation; and real-time intelligence broadcast feeds. Future Battle Command Brigade and Below is a separate system.

There is a multimodal soldier-centered computer interface. The modes are touch, speech, gaze, gesture, natural language, and battlefield visualization. The next steps are to develop broad bandwidth. There are also nuclear, biological, and chemical warfare and weather battlefield modeling that provide high-resolution weather, terrain, and nuclear, biological, and chemical warfare visualization. A goal is to create intuitive visualization to support rapid and accurate situational awareness by providing aggregation/deaggregation and temporal compression/ decompression. The concepts include filters, lethality, visibility, and prediction. There is a need to visualize agents that have been developed and what they do.

The Combat Information Processor incldues a 2-D map that was tethered to a virtual geographic information system 3-D view. Weather is overlaid on the 3-D. Annotations are presented overlaid on imagery. All the data from different applications can be integrated and shown in a single system with two screens. Legacy systems are bogged down in providing the dynamic feeds. The 2-D and 3-D can be untethered. There is also a multimodal interface. A speech-recognition system provided free by Microsoft is being used and works better than other commercial speech engines. It is being used as a front end to the natural language parser. Blobology is spatial integration with time compression to show troop movements. The 2-D world is used to generate a 3-D view to create blobology that shows mass of forces. Other things that could be used to define blobs are vulnerability and fire power. This would be useful for planning and after-action-reviews. Some of these tools will be fielded soon. Others are still under development.

G. M. Ferguson[182] defined *mixed initiative* as "several participants can each make contributions to the plan under development through some form of communication." Ferguson noted that such "communication can be explicit, as in a natural language or graphical front-end, or implicit from an agent's observation of other agent's actions."[183] His lessons learned[184] in developing a prototype mixed-initiative planner were: (1) "mixed-initiative planning is fundamentally a process of communication"; (2) "it is fundamentally based on defeasible reasoning, that is, conclusions are subject to revision given new information or time to reason"; and (3) "there are more common sources of defeasibility, such as incomplete knowledge of the world, uncertain effects of actions, and the like."

A key problem in mixed-initiative systems is the development of an unambiguous yet natural vocabulary. This is especially difficult, according to H. Chen, since "people tend to use different terms to describe a similar concept, depending on their backgrounds, training, and experiences."[185] This is exacerbated by collaboration across geographic areas or time. In these cases, there can be as little as 20 percent overlap in the use of given words.

The Navy has designed a mixed-initiative system to support situational assessment in warfare. Plan recognition is a software program designed to deduce enemy goals based on overt enemy actions. A force group display, similar to a diagram of a football play, is used to graphically depict enemy intentions.[186]

Computer-supported collaborative writing has been extensively analyzed. Not surprisingly, the interactive behavior of collaborators is dependent on the system design and the experience of the users. However, in general, users employ collaborative writing systems for exploration, organization, and composition. The system is rarely used for collaboration.[187]

One form of mixed initiative is adaptive automation. Levels of automation are listed in Table 4.

---

[182] G. M. Ferguson, *Knowledge Representation and Reasoning for Mixed-Initiative Planning.* Rochester, NY: University of Rochester, 1995, p. iv.
[183] Ibid., p. 62.
[184] Ibid., p. 67.
[185] H. Chen, "Collaborative Systems: Solving the Vocabulary Problem," *Computer* (May 1994), pp. 58–66.
[186] S. Kushnier, C. H. Heithecker, J. A. Ballas, and D. C. McFarlane, "Situational Assessment Through Collaborative Human-Computer Interaction," *Naval Engineers Journal* (July 1996), pp. 41–51.
[187] Chaomei Chen, "Writing With Collaborative Hypertext: Analysis and Modeling," *Journal of the American Society for Information Science*, Vol. 48, No. 11 (1997), pp. 1049–1066.

Table 4. *Levels of Automation*[188]

Mode	**Operator's Role**	**System's Role**
Silent/manual	Decide and act	Passive
Informative	Decide and act Influence system behavior	Support
Cooperative	Decide and act Influence system behavior Override system	Decide and act Support Override operator
Automatic	 Request information Influence system behavior	Decide and act Provide information Respond to operator influence
Independent	Passive	Decide and act

4.5 Facilitation

One form of facilitation is groupware, a "computer software technology enhancing the ability of people to work together as a group."[189] A groupware system, Group Support Systems, has been designed for NASA. It is being made more portable.

One method of facilitating collaboration is the development of a graphical representation of collaborative search. Ariadne is one example. It records queries and results, "subsequently producing a visualization of the search process that can be reflected on, shared and discussed by interested parties."[190]

Linda Candy identified "allocation between user and system of automated and mediated tasks" as an area ripe for research.[191]

4.6 Group Interaction Devices

One of the interesting opportunities is to provide some innovative technologies that enable groups to collaborate across data with some novel visualization techniques. Two examples of that are 3-D displays and data walls that support multiple people to interact simultaneously.

4.6.1. 3-D Sand Table

One interaction device that is very interesting but still not proven in terms of value in the command center is the 3-D sand table. It enables a small group of people to interact directly with

[188] B. A. Chalmers, "Design Issues for a Decision Support System for a Modern Frigate," in K. Garner (ed.), *Proceedings of the Second Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment*. Patuxent River, MD: Naval Air Warfare Center Aircraft Division, 1997.
[189] G. P. Hamel and R. Wijesinghe, *Group Support Systems (GSS)* (NASA-CR-201381). Houston, TX: NASA Johnson Space Center, May 1996, p. 1.
[190] Ibid., p. 182.
[191] Linda Candy, "Computers and Creativity Support: Knowledge, Visualisation, and Collaboration," *Knowledge-Based Systems*, Vol. 10, No. 1 (1997), p. 11.

a 3-D view of terrain and units. An example of this is the Dragon system at the Naval Research Laboratory—one of the first examples of a VR responsive workbench. A number of large screen display systems now support group interaction with 3-D views. They all require special glasses for interaction with them. The screens can be vertical, horizontal, or tilted.

4.6.2 Data Wall

A very interesting aspect of the large-screen displays occurs when a small group tries to interact with the data wall. With today's systems, there is one mouse or pointer that the group shares. But in the future, there will be several groups working on multiple pointer systems. This will require not only some interesting hardware techniques for identifying users' pointers, but also enhancements to operating systems to allow more than a single mouse or pointer.

(This Page Intentionally Blank)

Appendix A: References

Abrams, S. "Web-Based Collaborative Publications System: R&Tserve," *Sixth Alumni Conference of the International Space University* (1997), pp. 130–132.

Bauck, Jerry, and Duane H. Cooper. "Generalization Transaural Stereo and Applications," *Journal of the Audio Engineering Society*, Vol. 44, No. 9 (1996), pp. 683–704.

Baumann, R., and R. Clavel. "Haptic Interface for Virtual Reality Based Minimally Invasive Surgery Simulation," *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Vol. 1 (1998), pp. 381–386.

Bederson, Benjamin, and Jonathan Meyer. "Implementing a Zooming User Interface: Experience Building Pad++," *Software—Practice and Experience*, Vol. 28, No. 10 (1998), pp. 1101–1135.

Begault, Durand R. "Challenges to the Successful Implementation of 3-D Sound," *Journal of the Audio Engineering Society*, Vol. 39, No. 11 (1991), pp. 864–870.

Begault, Durand R. "Perceptual Effects of Synthetic Reverberation on Three-Dimensional Audio Systems," *Journal of the Audio Engineering Society*, Vol. 40, No. 11 (1992), pp. 895–904.

Begault, Durand R. "Call Sign Intelligibility Improvement Using a Spatial Auditory Display," *Proceedings of the Space Operations, Application, and Research Conference* (1993), pp. 363–370.

Begault, Durand R. (1994). *3-D Sound for Virtual Reality and Multimedia.* Boston, MA: Academic Press Professional, 1994.

Begault, Durand R. "Head-Up Auditory Display Research at NASA Ames Research Center," *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, Vol. 1 (1995), pp. 114–118.

Begault, Durand R. "Virtual Acoustics: Evaluation of Psychoacoustic Parameters," *Third International Workshop on Human Interface Technology IWHIT '97*, Aizu, Japan, 1997.

Begault, Durand R., S. R. Ellis, and E. M. Wenzel. "Headphone and Head-Mounted Visual Displays for Virtual Environments," invited paper, 15th International Conference of the Audio Engineering Society.

Begault, Durand R., and M. T. Pitman. "3-D Audio Versus Head-Down TCAS Displays," Moffett Field, CA: NASA Ames Research Center, 1994.

Begault, Durand R., and M. T. Pitman. "Three-dimensional Audio Versus Head-Down Traffic Alert and Collision Avoidance System Displays," *International Journal of Aviation Psychology*, Vol. 6, No. 1 (1996), pp. 79–93.

Begault, Durand R., and E. M. Wenzel. "Headphone Localization of Speech," *Human Factors*, Vol. 35, No. 2 (1993), pp. 361–376.

Begault, Durand R., E. M. Wenzel, L. L. Tran, and M. R. Anderson. "Octave-Band Thresholds for Modeled Reverberant Fields" (preprint 4662), *Audio Engineering Society 104th Convention*, Amsterdam, 1998.

Blauert, Jens. *Spatial Hearing: The Psychophysics of Human Sound Localization* (revised edition). Cambridge, MA: MIT Press, 1997.

Boff, Kenneth R., and Janet E. Lincoln, eds. *Engineering Data Compendium: Human Perception and Performance.* Wright-Patterson Air Force Base, OH: Armstrong Aerospace Medical Research Laboratory, 1988.

Bohan, J. A., and D. F. Wallace. "Team Ergonomics and Human Engineering Methods for the Design of Collaborative Work Environments: A Case Study," *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (1997), pp. 1066–1070.

Brems, Douglas J., Michael D. Rabin, and Jill L. Wagget. "Using Natural Language Conventions in the User Interface Design of Automatic Speech Recognition Systems," *Human Factors*, Vol. 37, No. 2 (1995), pp. 265–282.

Bronkhorst, Adelbert W., and J. A. (Hans) Veltman. "Evaluation of a Three-Dimensional Auditory Display in Simulated Flight," *AGARD Proceedings Audio Effectiveness in Aviation*, AGARD-CP-596 (1996), pp. 5-1 to 5-6.

Bronkhorst, Adelbert W., J. A. (Hans) Veltman, and Leo van Breda. "Application of a Three-Dimensional Auditory Display in a Flight Task," *Human Factors*, Vol. 38, No. 1 (1996), pp. 23–33.

Bursky, Dave. "3-D Audio Technologies Provide Realistic Sound," *Electronic Design*, Vol. 44 (4 November 1996), pp. 79–96.

Burstein, M. H., and D. V. McDermott. "Issues in the Development of Human-Computer Mixed-Initiative Planning," in Barbara Gorayska and Jacob L. May (eds.), *Cognitive Technology: In Search of a Humane Interface.* New York: Elsevier Science, 1996.

Candy, Linda. "Computers and Creativity Support: Knowledge, Visualisation, and Collaboration," *Knowledge-Based Systems*, Vol. 10, No. 1 (1997), pp. 3–13.

Carbonell, Jaime G. "Derivational Analogy: A Theory of Reconstructive Problemsolving and Expertise Acquisition," in R. S. Michalski, Jaime G. Carbonell, and T. M. Mitchell (eds.), *Machine Learning: An Artificial Intelligence Approach,* Vol. II. Morgan Kaufmann, 1986, pp. 371–392.

Casati, F., P. Grefen, B. Pernici, G. Pozzi, and G. Sanchez. *WIDE Workflow Model and Architecture.* Enchede, Netherlands: Technische University, 1997.

Chalmers, B. A. "Design Issues for a Decision Support System for a Modern Frigate," in K. Garner (ed.), *Proceedings of the Second Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment.* Patuxent River, MD: Naval Air Warfare Center Aircraft Division, 1997.

Chang, C., J. Wang, and R. K. Chang. "Scientific Data Mining: A Case Study," *International Journal of Software Engineering and Knowledge Engineering*, Vol. 8, No. 1 (1998), pp. 77–96.

Chen, Chaomei. "Writing With Collaborative Hypertext: Analysis and Modeling," *Journal of the American Society for Information Science,* Vol. 48, No. 11 (1997), pp. 1049–1066.

Chen, H. "Collaborative Systems: Solving the Vocabulary Problem," *Computer* (May 1994), pp. 58–66.

Cockburn, A., and T. Dale. "CEVA: A Tool for Collaborative Video Analysis," *Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work* (1997), pp. 47–55.

Conway, L., and C. J. Cohen. “Video Mirroring and Iconic Gestures: Enhancing Basic Videophones to Provide Visual Coaching and Visual Control,” *IEEE Transactions on Consumer Electronics*, Vol. 44, No. 2 (1998), pp. 388–397.

Cooper, G. E., and R. P. Harper, Jr. *The Use of Pilot Ratings in the Evaluation of Aircraft Handling Qualities.* NASA Technical Report TN-D-5153, 1969.

Dhar, V. “Data Mining in Finance: Using Counterfactuals to Generate Knowledge From Organizational Information Systems,” *Information Systems*, Vol. 23, No. 7 (1998), pp. 423–437.

Doll, T. J., and T. E. Hanna. “Spatial and Spectral Release From Masking in Three-Dimensional Auditory Displays,” *Human Factors*, Vol. 37, No. 2 (1995), pp. 341–355.

Ellis, “Sensor Spatial Distortion, Visual Latency, and Update Rate Effects on 3-D Tracking in Virtual Environments.”

Ellis, Stephen R. “Presence of Mind: A Reaction to Thomas Sheridan’s ‘Further Musings on the Psychophysics of Presence,’” *Presence*, Vol. 5, No. 2 (1996), pp. 247–259.

Ellis, S. R., N. S. Dorighi, B. M. Menges, B. D. Adelstein, and R. H. Jacoby. “In Search of Equivalence Classes in Subjective Scales of Reality,” *Proceedings of HCI International ’97* (1997), pp. 873–876.

Endsley, M. R., S. A. Rosiles, H. Zhang, and J. Macedo. *Three-Dimensional Auditory Localization as a Cue for Spatial Disorientation in Aircraft.* Bolling Air Force Base, DC: Air Force Office of Scientific Research, 1996.

Ericson, M. A., R. L. McKinley, M. P. Kibbe, and D. J. Francis. “Laboratory and In-Flight Experiments to Evaluate 3-D Audio Display Technology,” *Proceedings of the Space Operations, Application, and Research Conference* (1993), pp. 371–377.

Ferguson, G. M. *Knowledge Representation and Reasoning for Mixed-Initiative Planning.* Rochester, NY: University of Rochester, 1995.

Flor, Nick V. “Side-by-Side Collaboration: A Case Study,” *International Journal of Human-Computer Studies*, Vol. 49, No. 3 (1998), pp. 201–222.

Forbes, T. W. “Auditory Signals for Instrument Flying,” *Journal of Acoustical Science*, Vol. 13 (1946), pp. 255–258.

Geddes, N. D. “Intent Inferencing Using Scripts and Plans,” *Proceedings of the First Annual Aerospace Applications of Artificial Intelligence Conference* (1985), pp. 160–172.

Geddes, N. D., and J. M. Hammer. “Automatic Display Management Using Dynamic Plans and Events,” *Proceedings of the Sixth Symposium on Aviation Psychology* (30 April–4 May 1991).

Gershon, N. “Moving Happily Through the World Wide Web,” *IEEE Computer Graphics and Applications*, Vol. 16, No. 2 (1966), pp. 72–75.

Guignard, J. C. and Irving, A. “Effects of Low-Frequency Vibration on Man,” *Engineering Index* (1960), 51.

Haas, E. C., C. Gainer, D. Wrightman, M. Couch, and R. Shilling. “Enhancing System Safety With 3-D Audio Displays,” *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (1997), pp. 868–872.

Haas, M. W., S. L. Beyer, L. B. Dennis, B. J. Brickman, M. M. Roe, W. T. Nelson, D. B. Snyder, A. L. Dixon, and R. L. Shaw. *An Evaluation of Advanced Multisensory Display Contents for Use in Future Tactical Aircraft* (AL/CF-TR-1997-0049). Wright-Patterson Air Force Base, OH: Armstrong Laboratory, March 1997.

Hamel, G. P. and R. Wijesinghe. *Group Support Systems (GSS)* (NASA-CR-201381). Houston, TX: NASA Johnson Space Center, May 1996.

Hatfield, F., E. A. Jenkins, and M. W. Jennings. *Eye/Voice Mission Planning Interface* (AL/CF-TR-1995-0204). Wright-Patterson Air Force Base, OH: Armstrong Laboratory, December 1995.

Hatfield, F., E. A. Jenkins, M. W. Jennings, and G. Calhoun. "Principles and Guidelines for the Design of Eye/Voice Interaction Dialogs," IEEE Paper 0-8186-7493-8, 1996.

Heckerman, D. and E. Horvitz. "Inferring Informational Goals From Free-Text Queries: A Bayesian Approach," *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence* (1998), pp. 230–237.

Hendrix, C., and W. Barfield. "Presence in Virtual Environments as a Function of Visual and Auditory Cues," *Proceedings of the Virtual Reality Annual International Symposium* (1995), pp. 74–82.

Howard, C. W., J. M. Hammer, and N. D. Geddes. "Information Management in a Pilot's Associate," *Proceedings of the 1988 Aerospace Applications of Artificial Intelligence Conference*, Vol. 1 (1988), pp. 339–349.

Horvitz, E., and M. Barry. "Display of Information for Time-Critical Decision Making," *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (1995), pp. 296–305.

Hummels, C., and P. J. Stappers. "Meaningful Gestures for Human Computer Interaction: Beyond Hand Postures," *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition* (1998), pp. 591–596.

Jane's Information Group. "AV-8B to Test 3-D Audio Displays," *International Defence Review*, Vol. 24, No. 6 (1 December 1991), p. 176.

Johnson, C. W., J. C. McCarthy, and P. C. Wright. "Using a Formal Language to Support Natural Language in Accident Reports," *Ergonomics*, Vol. 38, No. 6 (1995), pp. 1264–1282.

Julig, L. F., and J. L. Kaiwi. *Task-Oriented Quantitative Testing for Synthesized 3-D Auditory Displays.* San Diego, CA: Naval Command, Control, and Ocean Surveillance Center, August 1995.

Kaomea, P., and W. Page. "A Flexible Information Manufacturing System for the Generation of Tailored Information Products," *Decision Support Systems*, Vol. 20 (1997), pp. 345–355.

Karkaletsis, E. A., C. D. Spyropoulos, and G. A. Vouros. "A Knowledge-Based Methodology for Supporting Multilingual and User-Tailored Interfaces," *Interacting With Computers*, Vol. 9 (1998), pp. 311–333.

Kibbe, M. P., and D. J. Francis. "TAV-8B Flight Test Results for the Demonstration of an Airborne 3-D Audio Cuer," *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting* (1994), p. 987.

King, Robert B., and Simon R. Oldfield. “The Impact of Signal Bandwidth on Auditory Localization: Implications for the Design of Three-Dimensional Audio Displays,” *Human Factors*, Vol. 39, No. 2 (1997), pp. 287–295.

Kolski, C., and E. LeStrugeon. “A Review of Intelligent Human-Machine Interfaces in the Light of the ARCH Model,” *International Journal of Human-Computer Interaction*, Vol. 10, No. 3 (1998), pp. 193–231.

Kuo, Sen M., and G. H. Canfield. “Dual-Channel Audio Localization and Cross-Talk Cancellation for 3-D Sound Reproduction,” *IEEE Transactions on Consumer Electronics*, Vol. 43, No. 4 (1997), pp. 1189–1196.

Kuro, T., N. Watanabe, S. Takano, and H. Kanayama. “A Proposal of Facial Picture Control Incorporating Several Essences of Conversation,” *Systems and Computers in Japan*, Vol. 29, No. 7 (1998), pp. 57–64.

Kushnier, S., C. H. Heithecker, J. A. Ballas, and D. C. McFarlane. “Situational Assessment Through Collaborative Human-Computer Interaction,” *Naval Engineers Journal* (July 1996), pp. 41–51.

Lange, B. M., M. A. Jones, and J. L. Meyers. “Insight Lab: An Immersive Team Environment Linking Paper, Displays, and Data,” *Proceedings of the Conference on Human Factors in Computing Systems* (1998), pp. 18–23.

Lavana, H., and F. Brglez. *OmniDesk and OmniFlows: Platform Independent Executable and User-Reconfigurable Desktops*. Research Triangle Park, NC: U.S. Army Research Office, 1997.

Lee, M. D., R. W. Patterson, D. J. Folds, and D. A. Dotson. “Application of Simulated Three-Dimensional Audio Displays to Fighter Cockpits: A User Survey,” *Proceedings of the IEEE 1993 National Aerospace and Electronics Conference*, Vol. 2 (1993), pp. 654–660.

Lyons, T. J., K. K. Gillingham, D. C. Teas, W. R. Ercoline, and C. Oakley. “The Effects of Acoustic Orientation Cues on Instrument Flight Performance in a Flight Simulator,” *Aviation, Space, and Environmental Medicine,* Vol. 61 (1990), pp. 699–706.

Maier, D., M. E. Meredith, and L. Shapiro. “Selected Research Issues in Decision Support Databases,” *Journal of Intelligent Information Systems*, Vol. 11 (1998), pp. 169–191.

Malkewitz, R. “Head Pointing and Speech Control as Hands-Free Interface to Desktop Computing,” *ACM Conference on Assistive Technologies,* Vol. 3 (1998), pp. 182–188.

Mascaro, S., and H. H. Asada. “Hand-in-Glove Human-Machine Interface and Interactive Control: Task Process Modeling Using Dual Petri Nets,” *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, Vol. 2 (1998), pp. 1289–1295.

McKinley, R. L., W. R. D’Angelo, and M. A. Ericson. “Flight Demonstration of an Integrated 3-D Auditory Display for Communication, Threat Warning, and Targeting.” *AGARD Conference Proceedings Audio Effectiveness in Aviation,* AGARD-CP-596 (1996), pp. 6-1 to 6-13.

McKinley, R. L., W. R. D’Angelo, M. W. Haas, D. R. Perrot, W. T. Nelson, L. J. Hettinger, and B. J. Barickman. “An Initial Study of the Effects of 3-Dimensional Auditory Cueing on Visual Target Detection,” *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 119–123.

Mosier, Jane N., T. L. Fanderclai, and K. K. Kennedy. *An Evaluation of CVW Use at MITRE.* MITRE Technical Report, 1998.

Mulgund, S. S., and G. L. Zacharias. “A Situation-Driven Adaptive Pilot/Vehicle Interface,” *Proceedings of the 3rd Annual Symposium on Human Interaction with Complex System* (1997).

Nassif, T. A. *Supporting the Fleet: Taking Workflow to the Waterfront*. Monterey, CA: Naval Post Graduate School, 1995.

Newell, Sima C. “User Models and Filtering Agents for Improved Internet Information Retrieval,” *User Modeling and User-Adapted Interaction*, Vol. 7, No. 4 (1997), pp. 223–237.

Nijenhuis, W., W. Jonker, and P. Grefen. *Supporting Telecom Business Processes by Means of Workflow Management and Federated Databases*. Enschede, Netherlands: International Institute of Technology and Management, 1997.

North, D. A., and W. R. D’Angelo. *3-Dimensional Audio Ergonomic Improvement Project for the NORAD CMOC* (AL/CF-TR-1997-0170). Wright-Patterson Air Force Base, OH: Armstrong Laboratory, 1997.

Oviatt, S. “Interface Techniques for Minimizing Disfluent to Spoken Language Systems,” in *Human Factors in Computing Systems: “Celebrating Independence,”* CHI ’94 Conference Proceedings. New York: ACM, 1994.

Oviatt, S. “Multimodal Interfaces for Dynamic Interactive Maps,” in *Human Factors in Computing Systems: Common Ground*, CHI 96 Conference Proceedings. New York: ACM, 1996.

Oviatt, S., M. MacEachern, and G. Levow. “Predicting Hyperarticulate Speech During Human-Computer Error Resolution,” *Speech Communication,* Vol. 24 (1998), pp. 87–110.

Perlis, D., K. Purang, and C. Andersen. “Conversational Adequacy: Mistakes Are Essential,” *International Journal of Human-Computer Studies*, Vol. 48, No. 5, 1998, pp. 553–575.

Perrott, D. R., J. Cisneros, R. L. McKinley, and W. R. D’Angelo. “Aurally Aided Detection and Identification of Visual Targets,” *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 104–108.

Perrott, D. R., T. Sadralodabai, K. Saberi, and T. Z. Strybel. “Aurally Aided Visual Search in the Central Visual Field: Effects of Visual Load and Visual Enhancement of the Target,” *Human Factors,* Vol. 33, No. 4 (1991), pp. 389–400.

Plenge, G. “On the Difference Between Localization and Lateralization,” *Journal of the Acoustical Society of America*, Vol. 56 (1974), pp. 944–951.

Pope, S. T., and L. E. Fahlen. “The Use of 3-D Audio in a Synthetic Environment: An Aural Renderer for a Distributed Virtual Reality System,” *IEEE Virtual Reality Annual International Symposium* (1993), pp. 176–182.

Raghavan, Vijay V., Jitender S. Deogun, and Hayri Sever. “Introduction,” *Journal of the American Society for Information Science*, Special Topic Issue: Knowledge Discovery and Data Mining, Vol. 49, No. 5 (1998), pp. 397–402.

Reinhardt, D. *Neural Network Modeling of the Head-Related Transfer Function* (AFIT/GAM/ENC/98M-01). Dayton, OH: Air Force Institute of Technology, 1998.

Savick, D. S. *A Comparison of Various Types of Head-Related Transfer Functions for 3-D Sound in the Virtual Environment* (ARL-TR-1605). Aberdeen Proving Ground, MD: Army Research Laboratory, 1998.

Scarborough, E. *Enhancement of Audio Localization Cue Synthesis by Adding Environmental and Visual Cues.* School of Engineering of the Air Force Institute of Technology Air University, 1992.

Sewell et al., 1987.

Sewell, D. R. and N. D Geddes. “A Plan Based Method for Computer-Human System Design,” *Human Computer Interaction: INTERACT’90* (1990), pp. 283–288.

Shalin, V. L., and N. D. Geddes. “Task Dependent Information Management in a Dynamic Environment: Concept and Measurement Issues,” *Proceedings of the 1994 IEEE International Conference on Systems, Man, and Cybernetics,* Vol. 3 (1994), pp. 2102–2107.

Sharma, Rajeev, Vladimir I. Pavlovic, Thomas S. Huang, “Toward Multimodal Human-Computer Interface,” *Proceedings of the IEEE*, Vol. 86, No. 5 (1998), pp. 853–869.

Smailagic, A., D. Siewiorek, A. Dahbura, and L. Bass, L. *MoCCA: A Mobile Communication and Computing Architecture.* Pittsburgh, PA: CMU, 1999.

Sollenberger, R. L., and P. Milgram. “Effects of Stereoscopic and Rotational Displays in a Three-Dimensional Path-Tracing Task,” *Human Factors*, Vol. 35, No. 3 (1993), pp. 483–499.

Solz, T. J., J. M. Reising, T. Barry, and D. C. Hartsock, “Voice and Aided Hand Trackers to Designate Targets in 3-D Space. *Proceedings of the Society of Photo-Optical Instrumentation Engineers,* 2734, 1996, 2–11.

Strybel, T. Z., J. M. Boucher, G. E. Fujawa, and C. S. Volp. “Auditory Spatial Cueing in Visual Search Tasks Effects of Amplitude, Contract, and Duration.” *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting* (1995), pp. 109–113.

Thomas, R. K. “Team-Based Access Control (TMAC): A Primitive for Applying Role-Based Controls in Collaborative Environments,” *Proceedings of the Second ACM Workshop on Role-Based Access Control* (1997), pp. 13–19.

“Three-D Sounds Points Pilots Toward the Enemy,” *Machine Design* (22 November 1999), pp. 40–41.

Tierney, J. “Jung in Motion, Virtually, and Other Computer Fuzz,” *The New York Times,* (September 16, 1993), B1–4.

Trybula, W. J. “Data Mining and Knowledge Discovery,” *Annual Review of Information Science and Technology*, 32 (1997), pp. 197–229.

Twidale, M. and Nichols, D. “Designing Interfaces to Support Collaboration in Information Retrieval,” *Interacting With Computers,* Vol. 10 (1998), pp. 177–193.

Valencia, G., M. A. Ericson, and J. R. Agnew. “A Comparison of Localization Performance With Two Auditory Cue Synthesizers,” *Proceedings of the 1990 National Aerospace and Electronics Conference*, Vol. 2, pp. 749–754.

Van de Graaf, M. C. A., and G. J. Houben. *Designing Effective Workflow Management Processes.* Eindhoven, Netherlands: Eindhoven University of Technology, 1996.

Veloso, M., A. M. Mulvehill, and M. T. Cox. "Rationale-Supported Mixed-Initiative Case-Based Planning," *Proceedings of the Fourteenth National Conference on Artificial Intelligence and Ninth Innovative Applications of Artificial Intelligence Conference* (1997), pp. 171–179.

Verlinden, J. C., J. D. Bolter, and C. van der Mast. *Virtual Annotation: Verbal Communication in Virtual Reality* (PB95215505). Delft, Netherlands: Delft University of Technology, 1993.

Vons, H., R. Vertegaal, and G. van der Veer. "Mediating Human Interaction With Video: The Fallacy of the Talking Heads," *Displays,* Vol. 18 (1998), pp. 199–206.

Welch, Robert B. *Perceptual Modification: Adapting to Altered Sensory Environments.* New York: Academic Press, 1978.

Wenzel, E. M. "The Relative Contribution of Interaural Time and Magnitude Cues to Dynamic Sound Localization," *Proceedings of the IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics* (1995), pp. 80–83.

Wenzel, E. M. "Effectiveness of Interaural Delays Alone as Cues During Dynamic Sound Localization" (abstract), *Acoustical Society of America,* Vol. 100 (1996), p. 2608.

Wenzel, E. M., S. Fisher, P. K. Stone, and S. H. Foster. "A System for Three-Dimensional Acoustic 'Visualization' in a Virtual Environment Workstation," *Proceedings of the First IEEE Conference Visualization, Visualization '90* (1990), pp. 329–337.

Wightman, F. L. *Evaluation of Three-Dimensional Auditory Displays.* Madison, WI: University of Wisconsin, April 1995.

Witmer, Bob G., and Michael J. Singer. "Measuring Presence in Virtual environments: A Presence Questionnaire," *Presence,* Vol. 7, No. 3 (1998), pp. 225–240.

Wright, Maury. "3-D Audio: Above and Behind You or Moving Low Left to Right?" *EDN* (6 June 1996), pp. 87–99.

Appendix B: Glossary of Terms

3-D audio 3-D audio displays.

Advanced whiteboarding Creation and sharing explanations and summary of information.

Alert and notification of events Many of the fuselets will be performing various kinds of alerts or detecting changes. There needs to be a *language* that users can use to describe what needs to be monitored. Rather than have a low level for setting up specific alerts, the user needs a language for describing the *policy* at a level meaningful to the user. Not only is the language important, but translating the user's requests into meaningful actions (including generating intelligent agents) is a major challenge in this area.

Annotation Attachment of explanations and caveats to expressions by users and others.

Automatic formatting and filtering Tailoring the information to the user, task, and equipment available.

Context understanding Real-time understanding of user(s)' situation and tasks at hand.

Conversational query and dialog User expressions of information needs and possibly desired sources.

Database "An organized collection of stored data."[192]

Data cleaning The "process of examining data and determining the existence of incorrect characters or mistransmitted information."[193]

Data mining The "process employed to analyze patterns in data and extract information."[194]

Data visualization 3-D visual displays, including animation.

Data warehouse A "repository of information that includes historical data and possible current information."[195]

Dialog management Embedded management of relationships among user(s)' expressions.

Dimensional database A database "that stores one or more kinds of base facts and connects them to dimensional information."[196]

Domain-specific gesturing Translations of gestural expressions.

Domain-specific workflow management Management of allocation of tasks, information, and decisions among participants.

Drill down Drill-down capabilities for explaining presentations.

Dynamically adaptable The system can *learn* from its experience. It can accept an explicit model of the user or the task, but over time it will be able to *infer* such a model.

Facilitation Support of group processes for discussion and decision making.

Gentle slope system Incremental capabilities require only incremental investment in training.

---

[192] W. J. Trybula, "Data Mining and Knowledge Discovery," *Annual Review of Information Science and Technology*, 32 (1997), p. 199.
[193] Ibid.
[194] Ibid.
[195] Ibid.
[196] D. Maier, M. E. Meredith, and L. Shapiro, "Selected Research Issues in Decision Support Databases," *Journal of Intelligent Information Systems*, Vol. 11 (1998), p. 173.

Information needs models Embedded understanding of information needs for situations and tasks.

Intent inferencing Real-time understanding of user(s)' goals, plans, and preferences.

Interactive analysis and query This includes the capability to drill down, do cluster analysis and data mining, and throughout the analysis, present the information in a way most meaningful to the user. This is also a *language* issue.

Knowledge discovery A "process of transforming data into previously unknown or unsuspected relationships that can be employed as indicators of future actions."[197]

Mixed initiative Human-machine partnership in problem solving.

Natural language Translations of natural language expressions.

Nontraditional senses Olfactory, tactile queuing.

Online analytical processing "The application of traditional query-and-reporting programs to describe and extract what is in a database."[198]

Online transaction processing "The method of automatically handling data as they are entered into a system."[199]

Pattern analysis "The application of a program to analyze data and look for relationships."[200]

Sharing Interaction via shared representations of information.

Speech Translations of vocalized expressions.

Tailored presentations This is also a *language* issue. One wants to provide the user with greater power to tailor the presentations of information to meet user needs. This will vary considerably depending on who the user is and what the user needs.

Tailoring Adaptation of presentations to particular users and current tasks.

Task-centered information discovery Using context understanding and intent inferencing to provide information relevant to the task the user is currently performing.

Undiscovered public knowledge "The creation of knowledge by acquiring similar but apparently unrelated information from textual databases with different domain information."[201]

[197] W. J. Trybula, "Data Mining and Knowledge Discovery," *Annual Review of Information Science and Technology*, 32 (1997), p. 199.
[198] Ibid.
[199] Ibid.
[200] Ibid.
[201] Ibid., p. 200.

Use-driven information dissemination Using context understanding to provide the user with the right information in the right format at the right time.

User tailorability Being able to use speech, natural language, and zooming.

Validation "The process of insuring the accuracy of data, beyond the process of data cleaning."[202]

[202] Ibid., p. 199.

(This Page Intentionally Blank)

Appendix C: Interact Technologies Survey

One goal of this year's study, *Information Management to Support the Warrior*, was for the Interact panel to identify products (application tools) that support the interact segment of the Joint Battlespace InfoSphere concept. Part of this effort included a survey of program managers, primarily within the Air Force, but also the Navy and Army. The spreadsheet on the following pages is a summary of the survey, and includes pointers to numerous sources where further information on the technologies may be found.

(Survey begins on next page)

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Perception					
3-D Visualization	3-D visual displays, including animation				
		AFRL/IFEC	COTS & GOTS	YES	A. Hall
		AFRL/IFSB	YES	YES	P. Jedrysik
		AFRL/IFTB	COTS	YES	M. Foresti
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	IC&V	J. Scholtz
		MITRE	Yes	CPoF	W. Page
		Navy—SPAWAR	YES	YES	J .Clarkson & M. Lasher
3-D Audio	3-D audio displays				
		AFRL/IFEC	COTS & GOTS	YES	D. Benincasa
		Army—Ft. Monmouth	YES	NO	John Soos
		MITRE		YES	N. Gershon
		MITRE		YES	B. Wright
		MITRE		YES	S. Eick
		MITRE		YES	R. Rao
		Navy—SPAWAR	YES	YES	G. Osga
Natural Language	Natural language presentations—visual or audio				
		AFRL/IFEC	COTS & GOTS	YES	C. Pine, W.Gadz, & D. Ventimiglia
		AFRL/IFTD	YES	NO	D. White
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	TIDES (new)	G. Strong
		Navy—SPAWAR	YES	YES	B. Sundheim

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Explanation	Drill-down capabilities for explaining presentations				
		AFRL/IFTB	GOTS	YES	P. Lucas
		AFRL/IFTD	YES	NO	D. White
		Army— Ft. Monmouth	YES	YES	John Soos
Tailoring	Adaptation of presentations to particular users & current tasks				
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	NO	D. White
		Army— Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	YES	YES	J. Clarkson
Understanding					
Modeling	Representation & manipulation of relationships among entities				
		AFRL/IFEC	COTS & GOTS	YES	J. Mucks
		AFRL/IFSB	YES	YES	A. Sisti & B. McQuay
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	R. Dziegiel
		MITRE	YES	YES	H. Carpenter
		Navy SPAWAR	YES	YES	D. Hardy
Simulation	Representation & manipulation of dynamic relationships				
		AFRL/IFEC	GOTS	YES	D. Ventimiglia
		AFRL/IFSB	YES	YES	A. Sisti & B. McQuay
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	MIC	R. Dziegiel
		Army— Ft. Monmouth	YES	YES	John Soos
		MITRE	YES	YES	H. Carpenter
		Navy SPAWAR	YES	YES	D. Hardy

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Sensitivity	Assessment of assumptions & their impact on what user is seeing				
		AFRL/IFTD	YES	WinWin	R. Dziegiel
What if?	Assessment of likely consequences of courses of action				
		AFRL/IFEC	GOTS	YES	J. Parker
		AFRL/IFSB	YES	YES	A. Sisti & B. McQuay
		AFRL/IFTD	YES	YES	R. Dziegiel
		Army—Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	NO/YES	YES	B. Schlichter & E. Allen
Decision					
Structuring	Representation of alternatives, attributes, & consequences				
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	J. Crowter & R. Dziegiel
		Army—Ft. Monmouth	YES	YES	John Soos
		MITRE	YES?	YES	P. Lehner
		Navy SPAWAR	?/YES	YES	R. Larsen & J. Morrison
Uncertainty Portrayal	Representation of missing, unreliable, indeterminate, & complex info.				
		AFRL/IFEC	GOTS	YES	D. Benincasa
		AFRL/IFTD	YES	YES	J. Crowter & R. Dziegiel
		AFRL/IFTE	YES	YES	L. Popyack
		Army—Ft. Monmouth	YES	YES	John Soos
		MITRE	NO	YES	N. Gershon
		Navy SPAWAR	?/YES	YES	B. Schlichter, R. Larsen, & J. Morrison
Tradeoff Management	Representation and assessment of benefits & costs				
		AFRL/IFTD	YES	YES	J. Crowter & R. Dziegiel
		Army—Ft. Monmouth	YES	YES	John Soos

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Advice	Representation of alternatives, attributes, & consequences				
		AFRL/IFSB	YES	YES	A. Sisti & B. McQuay
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	J. Crowter & R. Dziegiel
		AFRL/IFTE	YES	YES	L. Popyack
		Army—Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	?/YES	YES	L. Anderson
Communication					
Query Language	User expressions of information needs & possibly desired sources				
		AFRL/IFEC	COTS	YES	D. Ventimiglia
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	Communicator	G. Strong
		MITRE	YES	YES	A. Rosenthal
		Navy SPAWAR	YES	YES	
Natural Language	Translations of natural language expressions				
		AFRL/IFEC	COTS & GOTS	YES	C. Pine, W.Gadz, & D. Ventimiglia
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	TIDES (new)	G. Strong
		MITRE	YES	YES	G. Strong & L. Hirschman
		Navy SPAWAR	YES	YES	B. Sundheim
Speech	Translations of vocalized expressions				
		AFRL/IFEC	COTS	YES	D. Benincasa
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	TIDES (new)	G. Strong
		MITRE	YES	YES	G. Strong & L. Hirschman
		Navy SPAWAR	YES	YES	C. St. Clair

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Gesturing	Translations of gestural expressions				
		AFRL/IFEC	COTS	NOT ACTIVE	J. Gregory
		Army—Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	YES	YES	J. Clarkson
Annotation	Attachment of explanations & caveats to expressions by users & others				
		AFRL/IFEC	COTS	NOT ACTIVE	J. Parker
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	TIDES (new)	G. Strong
Collaboration					
Sharing	Interaction via shared representations of information				
		AFRL/IFEC	COTS	YES	C. Flynn
		AFRL/IFSB	YES	YES	B. McQuay
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	J. Milligan
		Army—Ft. Monmouth	YES	YES	John Soos
		DARPA	NO	IC&V	J. Scholtz
		MITRE	YES	YES	E. Rhode
		Navy SPAWAR	YES	YES	J. Weatherford & L. Duffy
Explanation	Creation and sharing explanations & summaries of information				
		AFRL/IFEC	COTS	YES	C. Flynn
		AFRL/IFSB	YES	YES	B. McQuay
		AFRL/IFTB	GOTS	YES	
		Army—Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	YES	YES	J. Weatherford & L. Duffy

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Facilitation	Support of group processes for discussion and decision making				
		AFRL/IFEC	COTS	YES	C. Flynn
		AFRL/IFSB	YES	YES	J. Smith & B. McQuay
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	J. Milligan
		Army— Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	YES	YES	J. Weatherford & L. Duffy
Workflow Management	Mgt. of allocation of tasks, information, & decisions among participants				**G. Osga**
		AFRL/IFEC	COTS	YES	C. Flynn
		AFRL/IFSB	YES	YES	B. McQuay
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	J. Milligan
		Army— Ft. Monmouth	YES	YES	John Soos
		Navy SPAWAR	NO/YES	YES	G. Osga
User Modeling					
Information Needs Models	Embedded understanding of information needs for situations & tasks				**G. Osga**
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	C. Burns
		Army— Ft. Monmouth	YES	YES	J. Peace
		Navy SPAWAR	NO/YES	YES	G. Osga
Dialog Management	Embedded management of relationships among user(s)’ expressions				
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	C. Burns
		MITRE	YES	YES	W. Page & L. Harper
		Navy SPAWAR	NO/YES	YES	Dan Lulue

Technologies for Presentation & Interaction	Definition/Explanation	Responding Organization	COTS/GOTS (Yes or No)	Current Program (Yes or No)	Program Manager (Name)
Context Understanding	Real-time understanding of user(s)' situation & tasks at hand				
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	C. Burns
		DARPA	NO	Communicator	G. Strong
		MITRE	YES	YES	W. Page & L. Harper
		Navy SPAWAR	NO/YES	YES	J. Morrison & G. Osga
Intent Inferencing	Real-time understanding of user(s)' goals, plans, & preferences				
		AFRL/IFEC	GOTS	YES	J. Parker
		AFRL/IFTB	GOTS	YES	
		AFRL/IFTD	YES	YES	C. Burns
		DARPA	NO	Communicator	G. Strong
		Navy SPAWAR	NO/YES	YES	J. Morrison, R. Larsen, & G. Osga

(This Page Intentionally Blank)

Appendix D: Abbreviations

ACP	Airspace Control Plan
Adtrees	All-Dimension trees
AFOSR	Air Force Office of Scientific Research
AFRL	Air Force Research Laboratory
ARPI	Advanced Research Projects Agency Rome Planning Initiative
API	Applicication Program Interface
ASR	automatic speech recognition
ATO	air tasking order
AVI	Automatic Vehicle Identification
CD-ROM	Compact Disc Read-Only Memory
CEE	Collaborative Enterprise Environment
CIA	Central Intelligence Agency
CMU	Carnegie Mellon University
COGs	Centers of Gravity
COTS	commercial off-the-shelf
CPE	Common Prototyping Environment
CPoF	Command Post of the Future
CVW	Collaborative Virtual Workspace
DARPA	Defense Advanced Research Projects Agency
dB	decibels
EDGE	Enhanced Geo Data Environment
EEG	electroencephalogram
EUN	Enhanced User Need
ExInit	Exercise Intialization
GOTS	government off-the-shelf
GPS	Global Positioning System
GUI	graphical user interface
HCC	Human-Centered Computing
HCI	human-computer interaction
HMRS	Hand Motion Gesture Recognition System
HPKB	High-Performance Knowledge Base
Hz	hertz
IC&V	Intelligent Collaboration and Visualization
IE	information extraction
IEEE	Institute of Electrical and Electronics Engineers
IF	information
IFD	Integrated Feasibility Demonstration

ISAAAC	Integrated Synchronous and Synchronous Collaboration
HCC	human-centered computing
IFD	Integrated Feasibility Demonstration
ISR	intelligence, surveillance, and reconnaissance
IWS	Info WorkSpace
JBI	Joint Battlespace InfoSphere
JADE	Joint Assistant for Deployment and Execution
JAOC	Joint Air Operations Center
JOVE	Joint Operations Visualization Environment
MAPVIS	Multiagent Planning and Visiualization
MFC	Microsoft© Foundation Classes
MIT	Massachusetts Institute of Technology
MVC	Model View Controller
NASA	National Aeronautics and Space Administration
NSF	National Science Foundation
OGL	Open Graphics Library
RF	radio frequency
SAB	Scientific Advisory Board
SAIC	Science Applications International Corporation
SCIF	sensitive compartmented information facility
SIPE4I	System for Interactive Planning & Execution
SOFPlan	Special Operations Flight Planning System
SPAWAR	Space and Naval Warfare
TCAS	Traffic Alert and Collision Avoidance System
TIDES	Threat/Intelligence Data Extraction System
TIE	Technology Integration Experiment
TPFDD	Time-Phased Force Deployment Data
TRIPS	The Rochester Interactive Planning System
U.S.	United States (of America)
USC	University of Southern California
VDI	Visible Decisions Inc.
VGA	video graphics array
VR	virtual reality
VRS	voice recognition system
WebHermes	Web Heterogeneous Reasoning and Mediator System
ZUI	zooming user interface

Report Distribution

Headquarters Air Force

SAF/OS	Secretary of the Air Force
AF/CC	Chief of Staff
AF/CV	Vice Chief of Staff
AF/CVA	Assistant Vice Chief of Staff
AF/HO	Historian
AF/ST	Chief Scientist
AF/SC	Communications and Information
AF/SG	Surgeon General
AF/SF	Security Forces
AF/TE	Test and Evaluation

Assistant Secretary for Acquisition

SAF/AQ	Assistant Secretary for Acquisition
SAF/AQ	Military Director, USAF Scientific Advisory Board
SAF/AQI	Information Dominance
SAF/AQL	Special Programs
SAF/AQP	Global Power
SAF/AQQ	Global Reach
SAF/AQR	Science, Technology and Engineering
SAF/AQS	Space and Nuclear Deterrence
SAF/AQX	Management Policy and Program Integration

Deputy Chief of Staff, Air and Space Operations

AF/XO	DCS, Air and Space Operations
AF/XOC	Command and Control
AF/XOI	Intelligence, Surveillance and Reconnaissance
AF/XOJ	Joint Matters
AF/XOO	Operations and Training
AF/XOR	Operational Requirements

Deputy Chief of Staff, Installations and Logistics

AF/IL	DCS, Installations and Logistics
AF/ILX	Plans and Integration

Deputy Chief of Staff, Plans and Programs

AF/XP	DCS, Plans and Programs
AF/XPI	Information and Systems
AF/XPM	Manpower, Organization and Quality
AF/XPP	Programs
AF/XPX	Strategic Planning
AF/XPY	Analysis

Initial Distribution (continued)

Deputy Chief of Staff, Personnel

AF/DP	DCS, Personnel

Office of the Secretary of Defense

USD (A&T)	Under Secretary for Acquisition and Technology
USD (A&T)/DSB	Defense Science Board
DARPA	Defense Advanced Research Projects Agency
DISA	Defense Information Systems Agency
DIA	Defense Intelligence Agency
BMDO	Ballistic Missile Defense Office

Other Air Force Organizations

AFMC	Air Force Materiel Command
– CC	- Commander, Air Force Materiel Command
– EN	- Directorate of Engineering and Technical Management
– AFRL	- Air Force Research Laboratory
– SMC	- Space and Missile Systems Center
– ESC	- Electronic Systems Center
– ASC	- Aeronautics Systems Center
– HSC	- Human Systems Center
– AFOSR	- Air Force Office of Scientific Research
ACC	Air Combat Command
– CC	- Commander, Air Combat Command
– AC2ISRC	- Aerospace Command and Control Agency
AMC	Air Mobility Command
AFSPC	Air Force Space Command
PACAF	Pacific Air Forces
USAFE	U.S. Air Forces Europe
AETC	Air Education and Training Command
– AU	- Air University
AFOTEC	Air Force Test and Evaluation Center
AFSOC	Air Force Special Operations Command
AIA	Air Intelligence Agency
NAIC	National Air Intelligence Center
USAFA	U.S. Air Force Academy
NGB/CF	National Guard Bureau
AFSAA	Air Force Studies and Analysis Agency
USSPACECOM	U.S. Space Command

U.S. Army

ASB	Army Science Board

U.S. Navy

NRAC	Naval Research Advisory Committee
SPAWAR-SSC	Space and Naval Warfare Systems Center, San Diego
Naval Studies Board	

Initial Distribution (continued)

U.S. Marine Corps

DC/S (A)	Deputy Chief of Staff for Aviation

Joint Staff

JCS	Office of the Vice Chairman
USJFCOM	U.S. Joint Forces Command
J2	Intelligence
J3/5	Operations
J4	Logistics
J6	Command, Control, Communications & Computer Systems
J7	Joint Training
J8	Strategy, Requirements & Integration
J9	Joint Experimentation

Other

Study Participants
Aerospace Corporation
ANSER
MITRE
RAND

(This page intentionally left blank)

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave Blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1999	**Final, January 1999 - December 1999**

4. TITLE AND SUBTITLE	5. FUNDING NUMBERS
Building the Joint Battlespace Infosphere, Volume 2: Interactive Information Technologies	
6. AUTHOR(S) J. McCarthy, D. Frost, R. Katz, R. Sproull, C. Morefield, V. Gawron	
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES) AF/SB Pentagon Washington, DC 20330-1180	8. PERFORMING ORGANIZATION REPORT NUMBER SAB-TR-99-02
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) SAF/OS AF/CC Pentagon Washington, DC 20330-1670	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Public Distribution authorized, February 2000	

13. ABSTRACT (Maximum 200 Words)

Building the Joint Battlespace Infosphere

The Joint Battlespace Infosphere (JBI) is a combat information management system that provides individual users with the specific information required for their functional responsibilities during crisis or conflict. The JBI integrates data from a wide variety of sources, aggregates this information, and distributes the information in the appropriate form and level of detail to users at all echelons. The JBI was originally described in the 1998 USAF Scientific Advisory Board (SAB) report *Information Management to Support the Warrior.*

In Chapter 4 of Volume 1 of the report, some interaction technologies were described in the context JBI functions: command, planning, execution, and combat support. In Volume 2, a much wider variety of interaction technologies is examined in greater detail. The goal of Volume 2 is to ensure that the masterpiece that is the JBI technical infrastructure is not partnered with clumsy, outdated user interfaces. Furthermore, the goal of the volume is to make JBI developers plan for future interaction technologies and not simply project current interaction techniques onto the JBI of the future.

The volume places interaction techniques into three categories: 1. C*apture*, which is the input of information to the JBI. 2. *Presentation* which is concerned with how the users perceive information. 3. C*ollaboration* which focuses on shared workspaces for multiple users.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Battlespace Infosphere, C4I, command and control, decision support tools, combat support, data fusion, fuselet, battlespace awareness, information management, information access, information grid, information manipulation, collaboration, force templates	116
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	None

This Page Intentionally Blank