NASA Technical Memorandum 109415

NASA STI Program
Coordinating Council
Eleventh Meeting

September 9, 1993

NASA STI Modernization Plan

This publication was prepared by the NASA Center for AeroSpace Information,
800 Elkridge Landing Road, Linthicum Heights, MD 21090-2934, (301) 621-0390.

TABLE OF CONTENTS

NASA STI PROGRAM
COORDINATING COUNCIL MEETING

NASA STI Modernization Plan

September 9, 1993
10:00 am - 4:00 pm
Crystal City Gateway 4
Conference Room

Attendees

NASA/Code JTT
Katie Bajis
Barbara Bauldock
Rick Dunbar
Denise Duncan
Jim Erwin
Dr. Linda Hill
Karen Holloway
Judy Hunter
Karen Kaye
Elliott Linder
Harry Needleman
Ann Normyle
LouAnn Scanlan
Ron Sepic
Roland Ridgeway
Patt Sullivan
Ardeth Taber-Dudas
Paula Trott
Dick Tuey
Kay Voglewede

NASA/CASI
Carl Eberline
Kevin Callahan
Wanda Colquitt
Bob Ferris
Walter Heiland
Gail Hodge
Rich Hughes
Mark Jeschke
Stephen Mullen
Jackie Streeks
Eric Vogle
Chuck Walsh

DTIC
Huddy Haller
Roberta Schoen

UUCOM
Rick Dunbar

MITRE
Howard Markham

AIAA
Tony Lenti
Dave Purdy

FASTC/SC
Tom Reinhardt

NASA/IIA
Bonnie Carroll

OSTI/DOE
Mark Fornwall

Welcome and Introductions

Patt Sullivan
NASA STI Program

NASA STI Modernization Plan

Karen Kaye
NASA STI Program

Engineering Review Board

Judy Hunter
NASA STI Program

RECON Replacement Project

Dr. Linda Hill
NASA STI Program

Electronic SCAN (Selected Current Aerospace Notices)

Rick Dunbar
NASA STI OProgram

Machine Translation Project

Katie Bajis
NASA STI Program

Multimedia

Karen Kaye
NASA STI Program

Electronic Document Interchange

Dick Tuey
NASA STI Program

Electronic Document Interchange

Jim Erwin
NASA STI Program

NAM (NASA Access Mechanism)

Judy Hunter
NASA STI Program

Network Upgrades

Roland Ridgeway
NASA STI Program

Standards in the Architectural Effort

Howard Markham
MITRE

NASA STI Program

Coordinating Council

The NASA Scientific and Technical Information (STI) Program Coordinating Council consists of participants from NASA Headquarters, NASA Centers, and NASA contractors. The Coordinating Council meets periodically to exchange information and pursue topics of vital interest to the NASA STI Program.

Coordinating Council Meetings

First Meeting	NASA RECON Database	May 23, 1990
Second Meeting	International Acquisition	July 23, 1990
Third Meeting	STI Strategic Plan	November 29, 1990
Fourth Meeting	NACA Documents Database Project	February 7, 1991
Fifth Meeting	Quality	July 1, 1991
Sixth Meeting	Who Are Our Key Users?	October 25, 1991
Seventh Meeting	Acquisitions	January 23, 1992
Eighth Meeting	Using the Internet	June 5, 1992
Ninth Meeting	Total Quality Management	October 28, 1992
Tenth Meeting	Information Retrieval: The Role of Controlled Vocabularies	April 22, 1993
Eleventh Meeting	NASA STI Modernization Plan	September 9, 1993

Document Preparation

The following transcription was prepared from the audio tape of the session by the staff of the NASA Center for AeroSpace Information and reviewed by the speakers. The transcription is intended to give the substance of the presentations and does not attempt to exhaustively report comments from the audience. Accompanying viewgraphs immediately follow each presentation.

N94- 36842

Welcome and Introductions

Patt Sullivan welcomed the assembled guests and introduced the speakers.

NASA STI Modernization Plan

Karen Kaye

Strategic Plan

Back in 1990, the STI program began work on a strategic plan (Viewgraphs 1 and 2). We had regular meetings, and identified where the program should go. That resulted in a document that specified goals for the STI Program. Those goals necessitated an upgrade of our current technology base. For example, one goal was to enhance the quality of our products and services, with a focus on the customer. Of course, in order to improve quality, we really needed to improve the underlying technology base. Another goal was to enhance and improve access to STI resources for the user community to make it easier for them to get our information. Next, we needed to increase the scope of and access to foreign materials.

Current Operations

We also needed to improve current operations. We had an analysis done of operations at the NASA Center for AeroSpace Information a year ago, and they identified some specific items that could be improved in the short-term to improve current operations. It was soon apparent that, in addition to those things, we really needed to upgrade the technology base to make major gains in what needed to be done.

User Studies

In 1990, we had the first of four user studies that indicated a need for improvements in specific areas. Getting input from our users is something that will last, that will continue to drive, what we're doing in terms of modernization. In 1991, an independent committee headed by Dr. Rosen recommended modernization of the STI infrastructure (Viewgraph 3). In 1992, a technology focus group was established by Gladys Cotter to identify technologies that could be leveraged to improve products and services.

Infrastructure Upgrade Plan

In 1993, we completed an infrastructure upgrade plan that included background on why we were doing modernization. It presented details of the results of user studies in a matrix form. The plan analyzed the current situation, our baseline, and looked at what we needed to do to upgrade the baseline and migrate to a modernized system. One of the focuses of the document was getting the funding we needed to go forward with the modernization. We were successful in getting the funding, and we have received the first increment, about two million dollars, which is being used to fund improvement projects.

Engineering Review Board

In 1993, an Engineering Review Board was established. The purpose of that board was to provide oversight to all of the projects that were encompassed by the modernization plan. Also in 1993, an architectural framework working group was established to look at the overall architectural issues that needed to be addressed in order to be sure that everything making up our modernization plan would work together and that we would be able to exchange information within NASA and with our exchange partners.

Modernization

What do we mean by modernization? First of all, we've used a number of terms for modernization during the life span of this plan (Viewgraph 4). As a matter of fact, we called it *modernization* at one time. The document itself was called *Infrastructure Upgrade Plan;* now we refer to it as *reinventing the STI Program.* The jargon is important: if you have the key words, it helps you to bring your issues and ideas to upper management. Additionally, what we're doing encompasses STI and user systems and all the support systems and services. It doesn't just deal with what the end user will see, but everything that makes what the end user receives possible. Our operating system spans the full document life cycle. We're using the new definition of *document* - essentially any type of recorded information that can be delivered to an individual. This includes video - everything - not just what we used to regard as a technical report. Our target architecture moves from highly centralized to distributed capabilities. Our time frame is five years, beginning in 1993.

Modernization Vision

What is our modernization vision? First of all, it's a virtual library (Viewgraph 5). That means access to information in a seamless fashion, so that we have at our fingertips all information that we need without physically having to go anywhere. Tied to that, we're talking about desk-top information access and delivery. Yet, we haven't stopped with the desktop; we have extended our information delivery vision to include your Newton, your portable personal assistant, even your car information system. It's *just in time* information delivery where and

when you need it. We're also including online translations. We're dealing here in terms of modernizing with commercial off-the-shelf and government off-the-shelf systems. Our focus is not to develop from scratch in-house, but to go and buy and integrate. In the case of government off-the-shelf, we don't even have to buy it.

Shared Network Environment

What is the modern STI information environment? What we're really looking at is a shared network environment (Viewgraph 6). We're talking about having access to full-text and images, having information on CD-ROM, having audio and video input and output, providing on-demand print, both local and remote, and optical archiving, on-demand translation, and a gateway. In that regard, we're looking at NAM, the NASA Access Mechanism, as our gateway. It's a project that got started in advance of modernization. There was a realization that this was a technology that was needed. The NAM virtual library includes people. It provides a peer-locator service. This is a very important service that came about as a result of our user studies. Scientists and engineers want to be able to contact others with similar expertise and be able to exchange information. Our user study showed that a peer locator service was really wanted.

Modernization Approach and Phases

What is our modernization approach? We have a very structured approach that includes incremental development of prototypes and testing (Viewgraph 7). Integral to our approach is our user and technical requirements. Planning is very, very integral. We're dealing with structured project plans and looking at how everything fits together, so that we can have an overall plan in which everything works. Since all of these things feed into each other, we're using a phased approach. We're talking about selecting and acquiring the technologies we need to support our modernization, integrating these technologies, building them into our infrastructure and doing rapid prototyping. What are our modernization phases? Of the five phases, the first phase begins in 1993 (Viewgraph 8). The mainframe replacement is the result of the recognition that we are dealing with antiquated IBM systems that carry high maintenance costs. We know now that there are new technologies out there that can do many of the same operations on a smaller scale system at a lower cost.

Network Upgrade and STIMS/RECON Replacement

The second item in Phase 1 is our network upgrade. We recognized a need at the NASA Center for AeroSpace Information (CASI) for some of the basic hardware, software and networking capabilities that have been utilized at some other locations. We are looking at modernizing the entire facility, making automation available to whoever needs it. Another item is our STIMS/RECON replacement. We are also considering bringing in a machine translation system. The graphical user interface (GUI) gateway front-end is the NASA Access

Mechanism (NAM). We also have a video multimedia support equipment item to support the handling of non-print information and multi-media.

Network Upgrade

In Phase 2, we will be upgrading the mainframe replacement. We will also be doing a network upgrade and bringing in additional equipment, including an optical imaging system. Additionally, more work will be done on NAM. In Phase 2, we will also see EDI, Electronic Document Interchange, getting past the trial, prototype stage and being implemented. In Phase 3, we have a second mainframe replacement, enhanced full-text and image retrieval, an enhanced optical imaging system, and EDI (Viewgraph 9). We are really looking at enhancements and bringing these systems up to speed, so that they will meet our operational needs. In Phase 4, we are again looking at upgrades of systems that have already been identified. We are also going to be looking at expert systems.

Network Capability

In Phase 5, we are again looking at upgrading our network capability. Phase 5 is way out in 1997, so it's hard to predict what will be available then and what we'll do. We're looking for gigabit transmission speeds on the network out that far. That will enable us to do things like deliver video at the desk-top in a far greater capacity. We're looking toward having the bandwidth needed to provide products and services that now seem pretty far-out to some of us. We are also looking at further enhancement of our optical imaging and electronic document interchange systems. Everything we've identified in Phase 1 is underway in one capacity or another, except for the mainframe replacement (Viewgraph 10). The reason that we're not doing that yet is that we're letting our software decisions drive our hardware decisions. So, the choice of the RECON replacement will determine what kind of hardware we'll use to replace our current mainframe.

Modernization Challenge

What is a modernization challenge? A big challenge is the overall structuring of these systems, seeing to it that they all work together (Viewgraph 11). We also need to identify the optimum windows of opportunity that will balance the issues and constraints. Some of these constraints may be technology constraints. One of the first things we did when we started this effort was to put together a technology map that looks at existing and emerging technologies that could be leveraged into the program. The optimum window of opportunity is very dependent upon the state of these technologies; this is particularity true as we move forward in our long-term modernization plan. We have endeavored to identify those technologies that will impact us, but there may be wonderful new technologies that emerge that we will be able to integrate easily into our operations to realize our modernization vision.

Evolutionary Modernization

Additionally, we are looking at achieving evolutionary modernization. This term has to be key during the conceptualization of the modernization program. We will start on a small scale; this will happen over the next five years. We must keep up with our user involvement with projects so that we can continue to meet our users' needs.

NASA STI PROGRAM
SCIENTIFIC & TECHNICAL INFORMATION

Modernization Overview

Karen Kaye
September 9, 1993

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Modernization Background

1990 **STI Program Management Team develops strategic plan**

1990 **Plan goals necessitate upgrade of technology base**

- **Enhance quality of our products and services through a focus on the customer**
- **Enhance and improve access to STI resources for the user community**
- **Increase the scope of access to foreign materials**
- **Improve current operations**

1990 onward **User studies indicate need for improvements**

KK-100. p2 1/12/94

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Modernization Background (cont'd)

1991	**Rosen committee recommends modernization of STI infrastructure**
1992	**Technology Focus Group established**
1993	**Infrastructure Upgrade Plan completed**
1993	**First funding increment received**
1993	**Engineering Review Board established**
1993	**Architectural Framework Working Group established**

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Modernization Overview

- **Encompasses STI end user systems and corresponding support systems/services**
- **Operating concept spans full "document" life-cycle**
- **Target architecture is a hybrid of centralized and distributed capabilities**
- **Timeframe is 5 years beginning 1993**

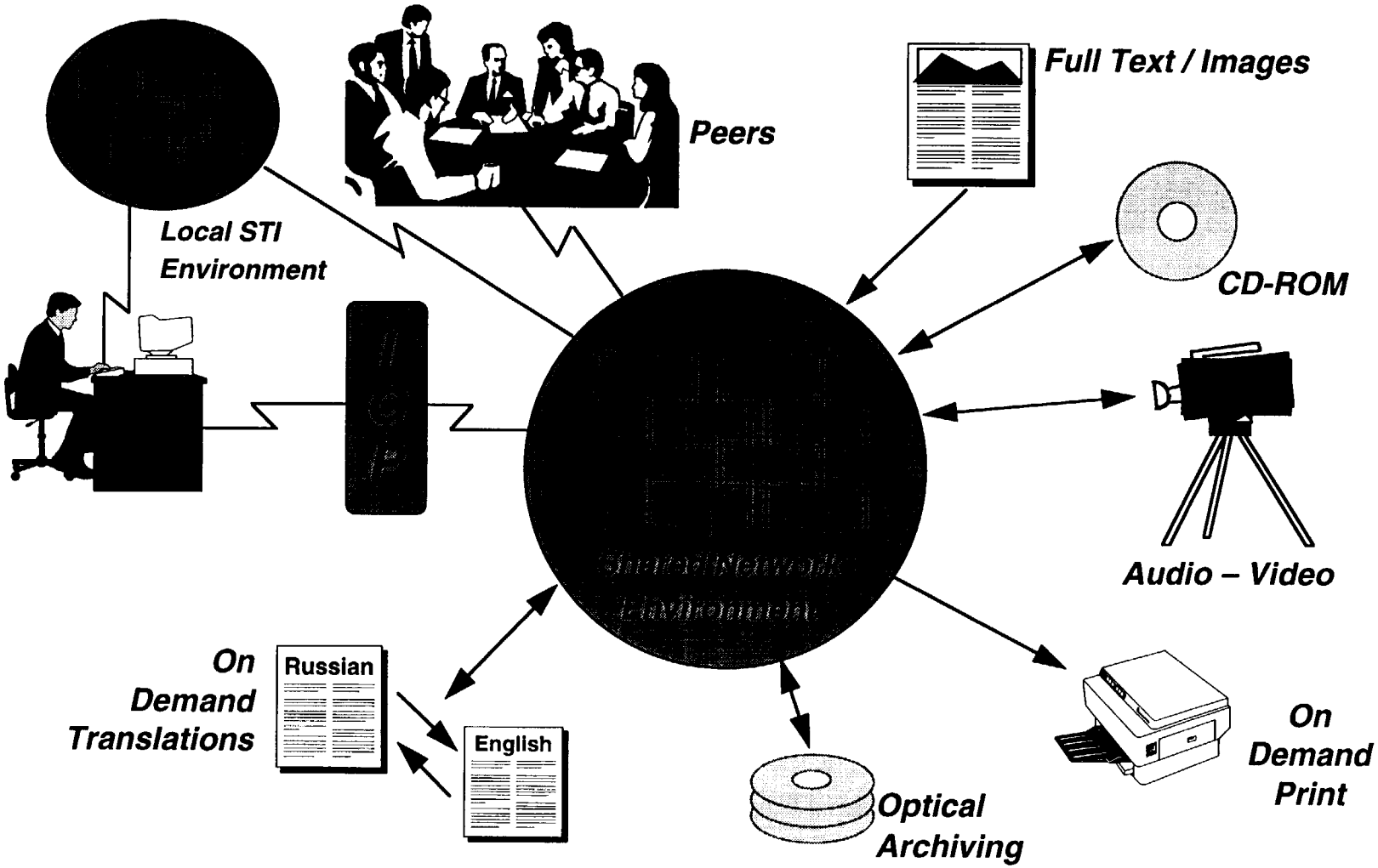
NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Modernization Vision

- **Virtual library (includes people)**
- **Desktop access and delivery**
- **Online translations**
- **COTS / GOTS**

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

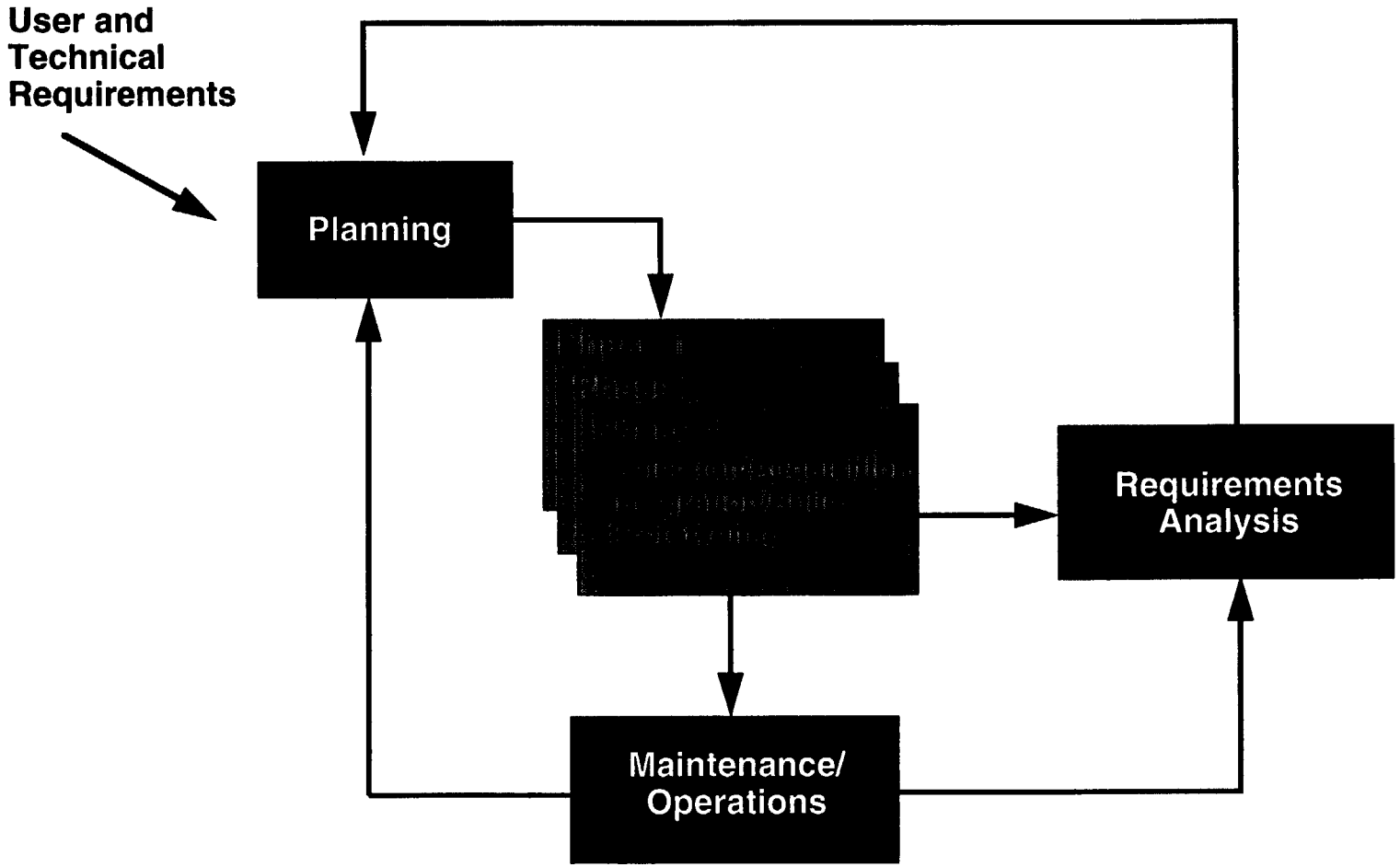
Modern STI Information Environment



KK-100. p6 1/12/94

NASA STI PROGRAM
SCIENTIFIC & TECHNICAL INFORMATION

Modernization Approach



NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Modernization Phases

1993 - 1997

Phase I - 1993

- Mainframe Replacement
- Network Upgrade I
- STIMS / RECON Replacement
    - Video / Multimedia Support Equipment
    - Electronic Document Interchange / Dissemination (3 month trial)
- Machine Translation System
- GUI / Gateway Front-end (NAM)

Phase 2 - 1994

- Mainframe Replacement Upgrade
- Network Upgrade II
- Optical Imaging System
- GUI / Gateway Front-end (NAM)
- Electronic Document Interchange/ Dissemination

13

KK-100. p8 1/12/94

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Modernization Phases (cont'd)

Phase 3 - 1995

- Second Mainframe Replacement
- Enhanced Full Text / Image Retrieval
- Enhanced Optical Imaging System
- Continue Electronic Document Interchange / Dissemination

Phase 4 - 1996

- Second Mainframe Replacement Upgrade
- Translation System Upgrade
- Second Enhancement Optical Imaging System
- Continue Electronic Document Interchange / Dissemination

Phase 5 - 1997

- Upgrade Network Capability
- Data Manipulation Tools
- Third Enhancement Optical Imaging System
- Continue Electronic Document Interchange / Dissemination

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Phase I Project Status

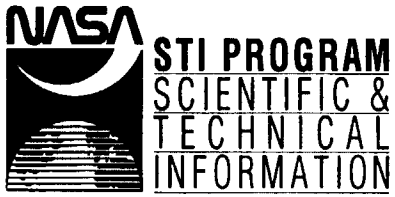
Phase 1 - 1993

- **Mainframe Replacement - Planned**
- **Network Upgrade I - Underway**
- **STMIS / RECON Replacement - Underway**
- **Machine Translation System - Underway**
- **GUI / Gateway Front-end (NAM) - Underway**

- **Video / Multimedia Support Equipment - Underway**
- **Electronic Document Interchange / Dissemination - Underway**

15

KK-100. p10 1/12/94



The Modernization Challenge

- **To identify the optimum windows of opportunity balancing all issues and constraints over the long and short term**
- **To achieve evolutionary modernization in annual increments over the next five years according to plan**
- **To sustain user involvement and meet user needs in measurable ways**

N94-36843

The Engineering Review Board

Judy Hunter

Background

The Engineering Review Board (ERB) was officially formed in 1993 (Viewgraph 1). It was put together in order to create our information infrastructure upgrade plan, our overall modernization effort. The Engineering Review Board is a permanent panel that meets regularly in order to coordinate all of the projects (Viewgraph 2). These people need to be in the position where they can look beyond this specific project to see how that one project fits into the whole system from a program- and system-wide perspective. The major project under review now is the RECON Replacement Project (Viewgraph 3). The next major focus will be on the NAM Lessons Learned document which we'll look at in October. The membership of the ERB consists of our program director, Gladys Cotter, and managers from each of our sections: Jim Erwin from Information Services, Judy Hunter from Special Projects, Barbara Bauldock from Budget, and Karen Kaye and Kristen Ostengaard, who help us with our strategic plan and with our long-range plan for the program (Viewgraph 4). Everyone on the board is a voting member. A quorum is three out of four of the managers, including the program director, and one or two of the staff people.

ERB Role

The role of the ERB is to actually approve concepts (Viewgraph 5). The ERB looks at a proposal from a system-wide perspective. The idea is to assure that all of the user requirements are met (Viewgraph 6). They look at the technical documentation. The board makes sure that you don't just get approval for the next step if your documentation is not in order; naturally they provide procurement and budget oversight. At the end of the project, they take a look at what actually happened in that project, and they make decisions about where or where not a project might go. In some cases, and hopefully in most of the cases, the project managers will come to the board with recommendations, and the board can choose form those recommendations (Viewgraph 7). The role of the ERB, just to summarize, is to look at the interfaces across the board, to make sure that all of the individual systems will, in the end, work together, that we're following whatever standards we decide to adhere to in order to reduce duplication. This will ensure that five or ten years down the line, these systems will still be working together and complementing each other, not working against each other.

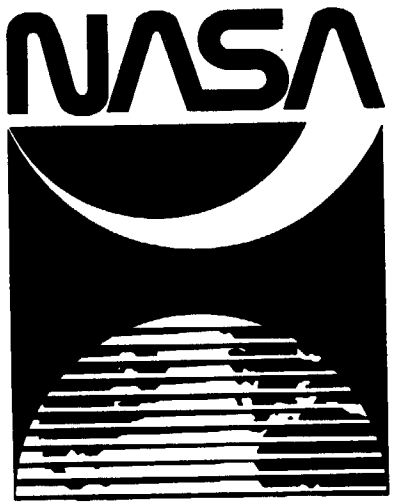
Future Directions

The future directions are to continue what we've been doing; that is, look at the projects; take all of the input from the Board, from all of the projects that have been done, and feed them into other projects in our architectural framework; make sure that all of the input we get from our sources, from all of our projects, are fed into anything new; and begin to make sure that all the interfaces work (Viewgraph 8). We need to develop; we know this. We're in the process of figuring out exactly the best way to do it in our environment, to develop new procedures for system product changes. If there's a major change, then things need to be reviewed.

NASA Scientific and Technical Information Program

The STI Program Engineering Review Board

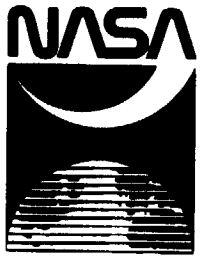
– An Overview –



STI PROGRAM
SCIENTIFIC &
TECHNICAL
INFORMATION

Presentation to the STI Coordinating Council

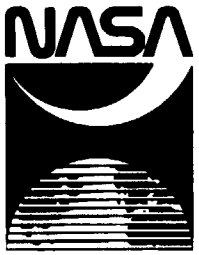
9 September 1993



STI PROGRAM
SCIENTIFIC &
TECHNICAL
INFORMATION

The Concept

- **The STI Engineering Review Board (ERB) is**
  - **A permanent panel**
  - **That convenes on a regular basis**
  - **To coordinate development projects and procurements**
  - **From a Program- and system-wide strategic perspective**



STI PROGRAM
SCIENTIFIC &
TECHNICAL
INFORMATION

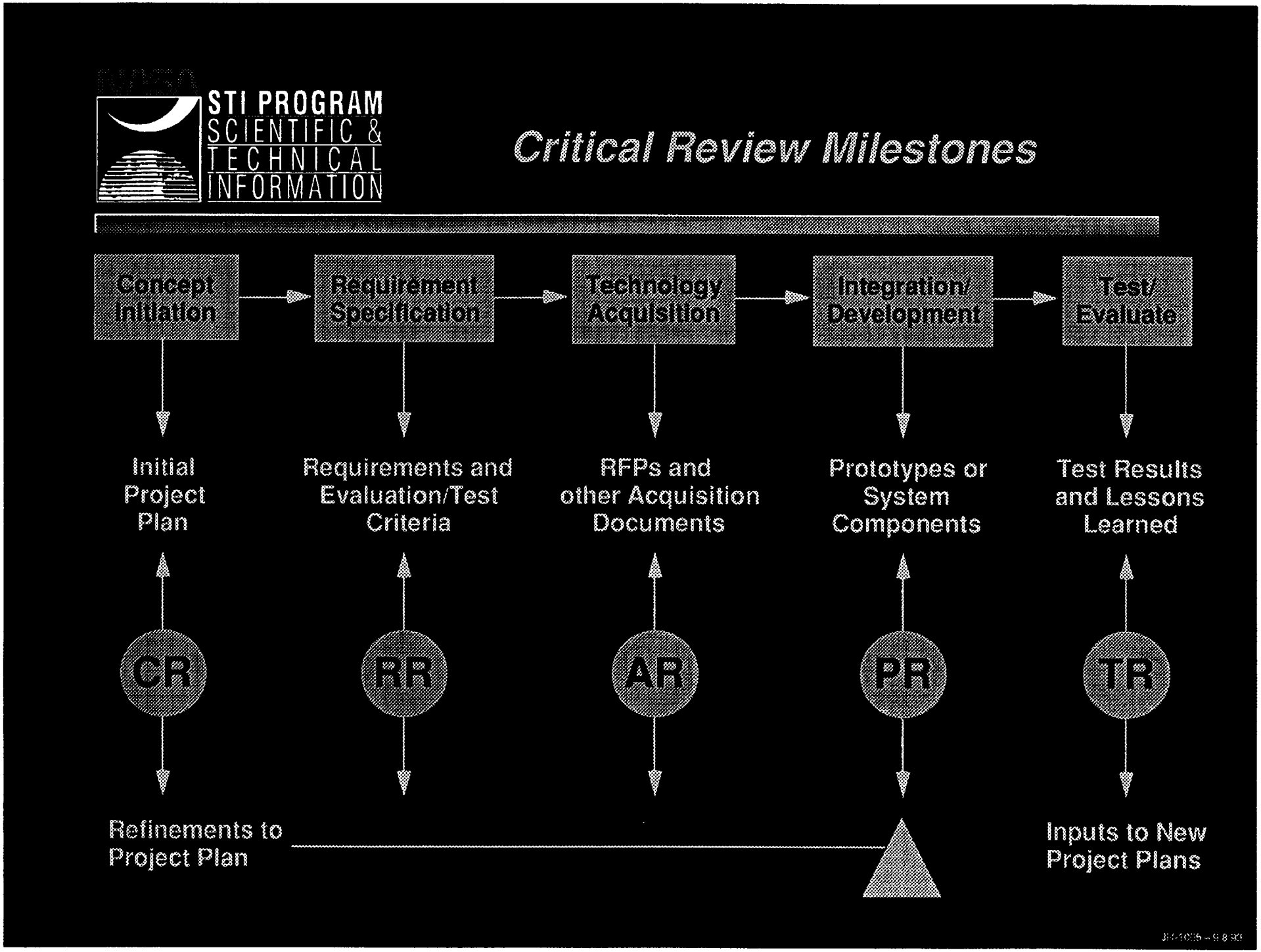
History

- **Formed in late 1992**
- **In anticipation of the "reinvention" of the STI Program**
- **First convened to support development of the STI Infrastructure Upgrade Plan (published 1 Feb 93)**
- **Currently reviewing the RECON Replacement project**
- **Next focus is NAM Lessons Learned**

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Internal ERB Organization

- **Membership consists of**
  - **Program Director**
  - **Managers representing Information Services, Special Projects, and Budget branches**
  - **Staff in the strategic planning areas**
- **Quorum requires a vote from 3 of 4 Managers and 1 of 2 Staff members**


STI PROGRAM
SCIENTIFIC & TECHNICAL INFORMATION
Critical Review Milestones
Concept Initiation
Requirement Specification
Technology Acquisition
Integration/ Development
Test/ Evaluate
Initial Project Plan
Requirements and Evaluation/Test Criteria
RFPs and other Acquisition Documents
Prototypes or System Components
Test Results and Lessons Learned
CR
RR
AR
PR
TR
Refinements to Project Plan
Inputs to New Project Plans

Role in Critical Reviews

- **Approval of project proposals and plans**
- **Assurance that user requirements are addressed**
- **Assessment of technical documentation**
- **Procurement and budget oversight**
- **Review of "lessons learned" at project conclusion**

Integration Role

- **Review of project and technology interfaces across the board**
  - **to assess workability and cost**
  - **reduce duplication and waste**
- **Proponent of standards utilization and compliance**
  - **in support of the longevity and interoperability of the COTs/GOTS technology selected by each project**

25

JH-1007 – 9/8/93

Future Directions

- **Input to STI Standards Project with ongoing interface for standards enforcement; Use of the STI Architectural Framework**
- **Interface to user-based focus/working groups e.g. a Quality Assurance Team**
- **Development of new procedures for system and product change processing**

Dr. Linda Hill
RECON Replacement Project

Background

The RECON Replacement Project is one of the first projects of the modernization plan. The intent is to replace one of the central elements of the service that the STI Program provides to our user community (Viewgraph 1). This drawing represents the environment in which the new information storage and retrieval system will reside. The replacement will be a commercial, off-the-shelf package providing the search and retrieval and database management functions that we need to support the program.

Project Components

The components of this project, in the center box of the drawing, are a search and retrieval engine, the database management package, and a system interface that will come with the system. The system is depicted as a kind of client/server architecture, which means different things in different environments. The environment we're bringing the new system into has to be understood. For example, we are not necessarily replacing the current Input Processing System (IPS) at CASI. It is not part of the procurement. Similarly, we have other systems - document ordering, registration, accounting, and photo-composition systems. These are not necessarily being replaced in this project.

Interface Requirements

Now, what that means is that we must be very sensitive to the interface requirements, and we have to know what interfaces are going to have to be adapted to the newly procured system. Users may very well have their own local clients, their own local interfaces, such as NASA's NAM gateway product or the NOTIS MDAS (Multiple Database Access System). The interface could be one of the many systems that are being developed at the Centers. We would like to have the system designed in such a way that the local interface could get to the new system through the system's interface or directly to the search engine, through the Z39.50 protocol, for example. Most of the time, the users are going to be using other databases; they can go directly to the other databases (for example, STN or DIALOG) or they can use the local interface to get to both our system and to the other databases.

Alternate Search Engines

We also don't see the new retrieval system as the only path to the data that will be available in the environment. We can provide alternate search engines, as illustrated on the left side of the drawing. We can establish proprietary indexes for alternate search engines that would access our databases. We are not, within this procurement, dealing with the image databases and the multimedia systems. What we are requiring in this procurement is that the system be able to interface with such systems.

Project Management Team

The Project Management Team is a small team that was designed to get the project going and moving on the fast track (Viewgraph 2). On the team are myself, Karen Holloway, Harry Needleman, Roland Ridgeway, and Gail Hodge. The procurement itself is an RMS procurement for CASI. The lines between the Project Management Team and the RMS box are intended to represent the very close relationships that we have and the fact that we're working hand-in-hand with the RMS staff to move this project along, to do the project planning and all that's entailed.

RECON Replacement ListServ

One of the methods that we've been using very successfully through this process to ensure communication and maximum understanding and knowledge of what we're doing, and to provide maximum opportunity to contribute input into decision making on various issues, is our ListServ, our RECON Replacement ListServ. There are 81 subscribers to the List now; we mail to those to whom we can't easily communicate electronically.

Project Elements

With this illustration, I'd like to give you some idea of some of the different elements of the project (Viewgraph 3). The first step was developing a requirements document - the functionalities, the architecture, and the capabilities that we wanted in the system. These requirements were turned into a Request for Proposal (RFP) because we discovered that the procurement method that we needed to use was the formal RFP process. The RFP has been sent out. The proposals are due on September 21. At that point, an evaluation process will begin. The evaluation process will result in a report and a recommendation. We are expecting the report to be done by the end of the year or in early 1994.

Record and File Restructuring

Since this replacement of RECON is a huge change to current operating methods, it gives us the opportunity to look at how the data are structured, both the records and the files. So, we have a very intense effort to restructure the record format and the file structure. This will eventually result in a record/file structure which is different from the one we have now. What we decided to do, in the spirit of the use of standards for the international exchange of information, was to use Z39.2, which is the bibliographic information interchange standard which is the basis for MARC, as our record framework for the re-design.

COSATI and MARC Formats

Some of you will realize what a radical step this is. The COSATI format, which is used by NASA, has been used by Federal Government agencies for their bibliographic record structure. Each agency has implemented its own version of COSATI, resulting in records that are similar to one another but not exactly the same. Libraries, on the other hand, use the MARC format for their monographic-type materials - their books, music scores, archives, manuscripts, maps, etc. The MARC format has not been applied, as we are applying it, to journal literature and technical reports. Because MARC is an international standard, we decided to adapt it to our types of records because many of the data elements in a MARC record are the same as we need for our records, and it provides the flexibility we need to create additional data elements.

Draft Record Design

We now have the draft record design; it is being reviewed internally by CASI for impact analysis. It will then be reviewed by the JTT staff, particularly the international group so that we're sure we have the data elements needed to track our agreements with other contributors to our database. It will also be discussed with a much wider group, with the CENDI Cataloging Working Group, with the MARC community and so forth. We completed an inventory of all existing data elements in RECON records. We are starting on plans for data clean-up during conversion. We have the chance to continue with the database upgrade project in which we identified quite a number of corrections that should be done, but we didn't have the resources. Primarily, it came down to the inflexibility of STIMS/RECON. It was just too difficult to do the corrections in this environment. So, we said, "We'd like to do these things someday, but we can't do them now." Now, we have the opportunity to take care of those things.

Plans

We will be doing prototype conversions and testing, and we will be doing an evaluation of file collections. Essentially, right now, RECON and STIMS are structured in files to represent certain characteristics of the records which we will be able to represent by data elements. We will not have to have a whole file to represent the distribution limitations, for example. We'll be combining as many files as possible. Down the line, we know we have to deal with installation of both hardware and software. Preliminary plans for this stage are being made. We're going through the process of software procurement first. We'll start the hardware procurement as soon as we know what we need for the system. We will have an acceptance testing period for the software. We will do phased file conversion, doing the main ones first. We will keep parallel systems up for a period of time because this is central to our service, and we have to ensure that the new system is stable before cutting over to it. We know that we will have to make modifications to system interfaces. There will have to be training, both internally and for the user community, and user documentation will have to be developed. We will also have to have a promotional effort. The goal of the project was to finish by the end of 1994. Because of the delays that we experienced in the RFP process, the project will extend into mid-1995. Our goal right now is to actually cut over to the new system by the end of 1995, saying goodbye to our RECON workhorse, which has done a very good job for us, but which is ready for retirement.

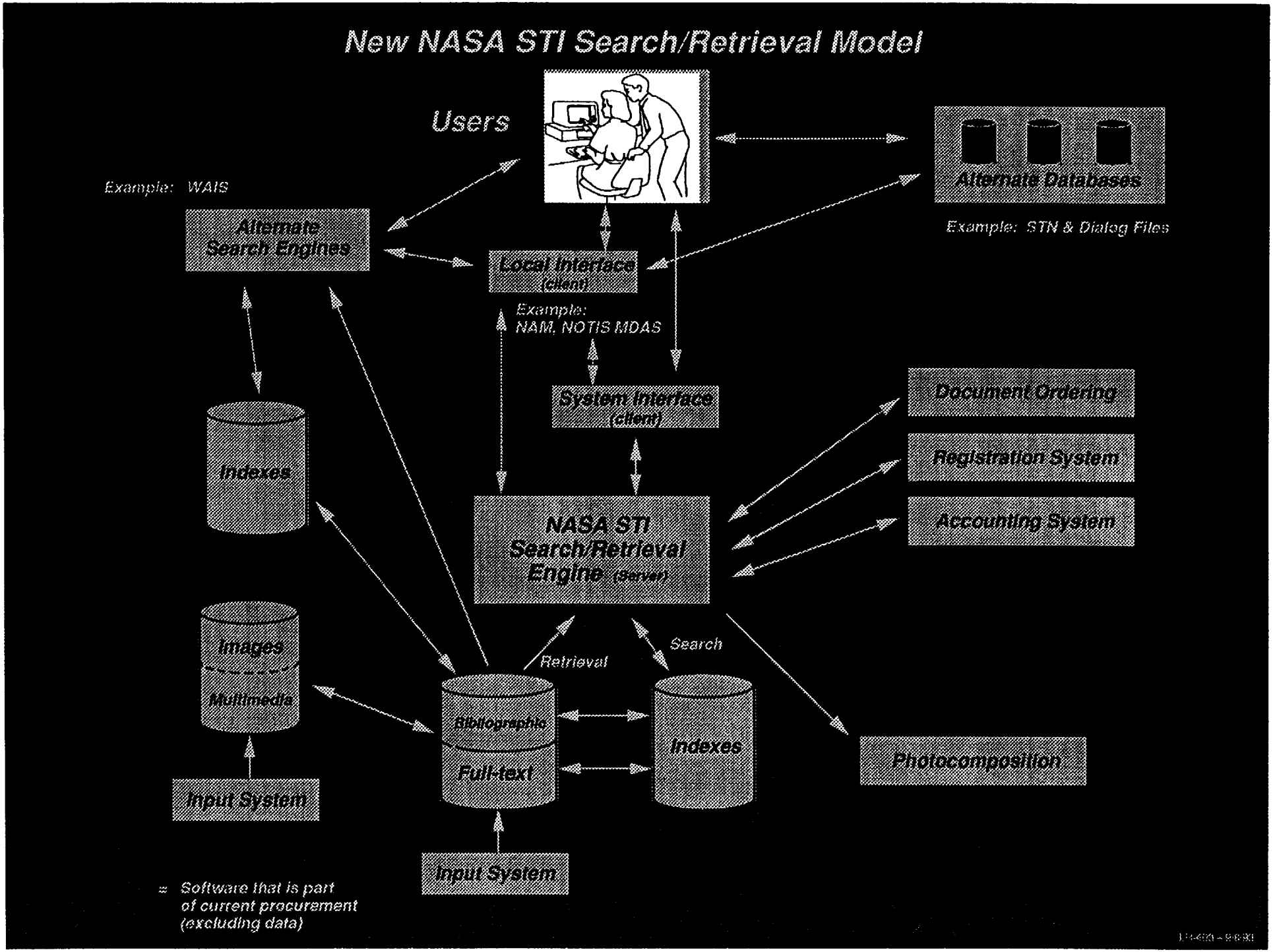
Are there any questions?

Question: What are we looking for in the replacement of the RECON database management system?

Answer: We are looking for a very sophisticated Boolean-based system. We are looking for a standards-based, open-architecture kind of system that we can use to interface all of our packages. We are looking for the capability to have flexible record structures so that we can integrate such things as pointers or links over to document images or other kinds of data.

Question: Will we be moving towards full-text capability?

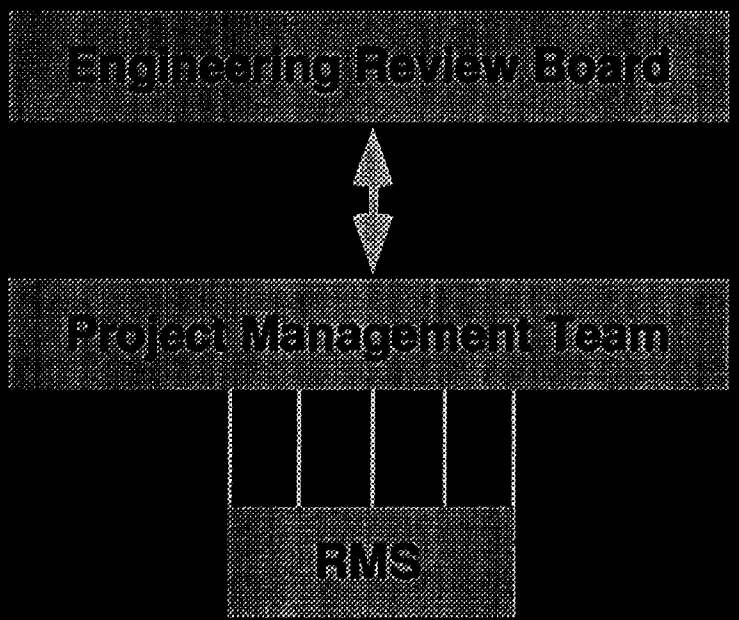
Answer: Yes, we are.


New NASA STI Search/Retrieval Model
Users
Example: WAIS
Alternate Search Engines
Local Interface (client)
Example: NAM, NOTIS MDAS
Alternate Databases
Example: STN & Dialog Files
System Interface (client)
Indexes
Document Ordering
Registration System
Accounting System
NASA STI Search/Retrieval Engine (Server)
Images
Multimedia
Retrieval
Search
Bibliographic
Full-text
Indexes
Photocomposition
Input System
Input System
= Software that is part of current procurement (excluding data)

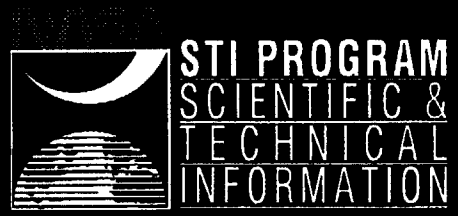
STI PROGRAM
SCIENTIFIC & TECHNICAL INFORMATION

STIMS/RECON Replacement Project

Methodology



- ListServ
- ViTS
- Presentations
- Mailings
- Status Reports
- Working Group



STIMS/RECON Replacement Project

Overview of Project Plan

Procurement

Requirements
Procurement Method
RFP
Evaluation Method
Award

Database Redesign

Logical Record Structure
Logical File Structure
Physical Record/File Structure

Pre-Implementation

Data Mapping
Data Cleanup
Prototype Conversions
Testing
Evaluation of File Collections

Implementation

Installation
Acceptance Testing
Phased File Conversion
Parallel Systems
Modifications to System Interfaces
Training: Internal
User Documentation

33

Electronic SCAN (Selected Current Aerospace Notices)

Rick Dunbar

Electronic and Paper SCAN

Essentially, the STI Program has had the paper SCAN for a long time, so STI management figured, why not have it online? (Viewgraphs 1 and 2). What happens is that a couple of times a month, CASI makes a SCAN on paper, and they make a big file electronically and automatically File Transfer Protocol (FTP) it over to a machine here, and we slice it and dice it into the SCAN topics and make it available by a Gopher, anonymous FTP, and LISTSERV so that people can get it whenever they like.

Electronic Ordering Methods

Essentially, it's available three ways (Viewgraph 3).

1) You can FTP to a ftp.sti.nasa.gov, login as the user *anonymous*, and get the topics that way. It's a little bit cumbersome because it goes by the SCAN topic numbers to keep the path name short so you don't have to type *aerodynamics of whatever* to see a couple of files.

2) Gopher is the easiest way to browse it because you can just sit and take your time and click and point around and see what's there.

3) LISTSERV is nice if you just like to read mail or you don't have real Internet access; like, if you only got NASA Mail, you can still get it electronically and look at it, so it's kind of handy. I've also thought about putting it into a world-wide web because those interfaces are nicer, actually, than Gopher.

File Transfer Protocol

Here's my final, my big statement about FTP (Viewgraph 4). If you can't type FTP on your machine and have it work, talk to whoever maintains your machine and have them make it work or show you what you are doing wrong because, if you're on the network and you've got access to the Internet using TCI/IP, you should have it.

Gopher

Gopher clients are a little bit different (Viewgraph 5). There's a zillion of them and I didn't try to list them here. There are Gopher clients for whatever you use. If you have DOS, there's a Gopher for DOS. If you have Microsoft Windows, there are Gopher clients for you. The DOS clients are a pain because, for every vendor's TCIP/IP package, you have to write a client for that package. The Windows implementation. What they're doing with this implementation is writing to a library that essentially knows how to talk to the package so that you can write a Gopher client that looks nice and can run on anyone's TCIP/IP package that runs on Windows, and the X client's okay as well.

LISTSERV

LISTSERV is a little bit more complicated because you've actually got to send e-mail and say, "I want a certain list" (Viewgraph 6). Essentially, what you do is send a note to listserv@sti.nasa.gov and subscribe to the list you want. It looks like this. You just put that in one of your mail messages. Users can also subscribe to multiple lists at the same time. If you want all of the SCAN's, there's a list called *scan-all-topics*, and it will give you a lot of mail twice a week. If you have more discriminating taste, you can pick what you want and you'll get mail in logical hunks. Here is the first page of the many pages of the list (Viewgraphs 7 - 18). If you subscribe to the SCAN-02, you get all the little sub-things underneath it. If you take a step out where it says *Aeronautics*, you'll see that in each of these main categories there's one that's indented that's called *General*; if you subscribe to that one, you get everything under the main category. By subscribing to scan-01, you'll receive everything under the main Aeronautics heading.

Electronic SCAN

Rick Dunbar
rdunbar@sti.nasa.gov

September 9, 1993

Current Process

- o SCAN is created bi-monthly at CASI
- o File is transferred using FTP to nova.sti.nasa.gov
- o File reformatted and broken into various SCAN topics and made available via gopher and anonymous FTP.
- o SCAN's E-mailed to people on the various listserv lists

How can I access SCAN

o Anonymous ftp on ftp.sti.nasa.gov in /scan directory

o Gopher using gopher server gopher.sti.nasa.gov

o Listserv by requests to listserv@sti.nasa.gov

o Perhaps via WWW in the future

FTP Clients

o These come with your TCP/IP Software. If you have problems talk with your systems/network administrator.

Gopher Clients

o There are a variety of gopher clients available for DOC, MacOS, Miscrosoft Windows and X Windows.

The DOS Clients are picky about which TCP/IP you are running.

The MS-Windows clients are (for the most part) using winsock so that you don't have to write a client for each TCP/IP implementation.

o To get a current list of gopher clients ftp to ftp.sti.nasa.gov and get the file /scan/Gopher-Clients.

Listserv Lists

o To find out what the topics are send E-mail to scan@sti.nasa.gov

o To receive the SCAN's via E-mail send a note to listserv@sti.nasa.gov and put 'subscribe' commnads in the body of the message. For example:

1) To subscribe to all SCAN's use:

subscribe SCAN-ALL-TOPICS Your Name

2) To subscribe to the three Aerodynamics topics use:

subscribe SCAN-02 Your Name

3) To subscribe to a single topic in a group use:

subscribe SCAN-02-01 Your NAME

02 AERODYNAMICS
02-01 AERODYNAMICS CHARACTERISTICS
02-02 AERODYNAMICS OF BODIES

COMPLETE LIST OF SCAN TOPICS:

AERONAUTICS
- 01 AERONAUTICS (GENERAL)
- NO TOPICS AVAILABLE
- 02 AERODYNAMICS
- 02-01 AERODYNAMICS CHARACTERISTICS
- 02-02 AERODYNAMICS OF BODIES
- 02-03 AIRFOIL AND WING AERODYNAMICS
- 03 AIR TRANSPORTATION AND SAFETY
- 03-01 COMMERCIAL AND GENERAL AVIATION
- 03-02 HELICOPTERS AND GROUND EFFECT MACHINES
- 03-03 STOL/VTOL AIRCRAFT
- 03-04 SUPERSONIC TRANSPORT
- 03-05 AIRCRAFT NOISE AND SONIC BOOM
- 03-06 AIRCRAFT SAFETY AND SAFETY DEVICES
- 03-07 CLEAR AIR TURBULENCE
- 04 AIRCRAFT COMMUNICATIONS AND NAVIGATIONS
- NO TOPICS AVAILABLE
- 05 AIRCRAFT DESIGN, TESTING AND PERFORMANCE
- 05-01 HYDRAULIC AND PNEUMATIC SYSTEMS
- 05-02 AUXILIARY ELECTRICAL SYSTEMS
- 06 AIRCRAFT INSTRUMENTATION
- NO TOPICS AVAILABLE
- 07 AIRCRAFT PROPULSION AND POWER
- 07-01 JET PROPULSION
- 08 AIRCRAFT STABILITY AND CONTROL
- NO TOPICS AVAILABLE
- 09 RESEARCH AND SUPPORT FACILITIES (AIR)
- 09-01 WIND TUNNELS

ASTRONAUTICS
- 12 ASTRONAUTICS (GENERAL)
- NO TOPICS AVAILABLE
- 13 ASTRODYNAMICS
- 13-01 CELESTIAL MECHANICS AND ORBITAL CALCULATIONS
- 14 GROUND SUPPORT SYSTEMS AND FACILITIES (SPACE)
- 14-01 SPACECRAFT GROUND SUPPORT
- 14-02 TEST FACILITIES
- 14-03 SIMULATORS AND SIMULATION
- 14-04 STERILIZATION
- 15 LAUNCH VEHICLES AND SPACE VEHICLES
- 15-01 LAUNCH VEHICLES
- 15-02 SOUNDING ROCKETS
- 15-03 SPACE PROBES
- 15-04 SCIENTIFIC SATELLITES
- 15-05 REENTRY VEHICLES
- 15-06 U.S.S.R SPACECRAFT
- 16 SPACE TRANSPORTATION
- 16-01 SPACE TRANSPORTATION AND MANNED SPACECRAFT
- 17 SPACE COMMUNICATIONS, SPACECRAFT COMMUNICATIONS, COMMAND AND TRACKING
- 17-01 SPACE COMMUNICATIONS
- 17-02 NAVIGATION SYSTEMS
- 17-03 GUIDANCE SYSTEMS
- 17-04 TRACKING
- 18 SPACECRAFT DESIGN, TESTING AND PERFORMANCE
- 18-01 SPACECRAFT ATTITUDE CONTROL AND STABILIZATION
- 18-02 RENDEZVOUS AND DOCKING
- 18-03 SPACE STATIONS
- 19 SPACECRAFT INSTRUMENTATION
- 19-01 SPACECRAFT AND AIRCRAFT INSTRUMENTATION
- 19-02 SENSORS AND TRANSDUCERS
- 20 SPACECRAFT PROPULSION AND POWER
- 20-01 ROCKET ENGINES, NOZZLES AND THRUST CHAMBERS
- 20-02 AUXILIARY PROPULSION
- 20-03 ELECTRIC PROPULSION

CHEMISTRY AND MATERIALS
- 23 CHEMISTRY AND MATERIALS (GENERAL)
- 23-01 CHEMICAL ANALYSIS
- 23-02 CHEMICAL PROCESSES AND ENGINEERING
- 23-03 LUMINESCENCE
- 23-04 PHOTOCHEMISTRY
- 24 COMPOSITE MATERIALS
- 24-01 REINFORCED MATERIALS AND FIBERS
- 24-02 COMPOSITE MATERIALS
- 25 INORGANIC AND PHYSICAL CHEMISTRY
- 25-01 CORROSION
- 25-02 METAL CRYSTALS
- 25-03 COATINGS
- 25-04 ELECTROCHEMISTRY
- 26 METALLIC MATERIALS
- 26-01 ALUMINUM
- 26-02 BERYLLIUM
- 26-03 LIQUID METALS
- 26-04 STEEL
- 26-05 TITANIUM
- 26-06 REFRACTORY METALS
- 26-07 METALLURGY
- 27 NONMETALLIC MATERIALS
- 27-01 PLASTICS
- 27-02 ADHESIVES
- 27-03 CERAMICS
- 27-04 ELASTOMERS
- 27-05 GRAPHITE
- 27-06 POLYMERS
- 28 PROPELLANTS AND FUELS
- 28-01 LIQUID PROPELLANTS
- 28-02 SOLID PROPELLANTS
- 29 MATERIALS PROCESSING
- NO TOPICS AVAILABLE

ENGINEERING
- 31 ENGINEERING (GENERAL)
- NO TOPICS AVAILABLE
- 32 COMMUNICATIONS AND RADAR
- 32-01 COMMUNICATION SATELLITES
- 32-02 COMMUNICATION EQUIPMENT
- 32-03 COMMUNICATION SYSTEMS
- 32-04 TELEMETRY
- 32-05 RADIO NOISE
- 32-06 COMMUNICATION THEORY
- 33 ELECTRONICS AND ELECTRICAL ENGINEERING
- 33-01 RADAR EQUIPMENT
- 33-02 SEMICONDUCTORS AND TRANSISTORS
- 33-03 ANTENNAS
- 33-04 ELECTRONIC COMPONENTS
- 33-05 CIRCUITRY
- 33-06 ELECTRICAL EQUIPMENT
- 33-07 AMPLIFIERS
- 33-08 FEEDBACK AND CONTROL THEORY
- 33-09 ELECTROMAGNETIC RADIATION
- 33-10 MICROELECTRONICS
- 33-11 MICROWAVE AND SUBMILLIMETER WAVE TECHNOLOGY
- 33-12 MAGNETISM
- 34 FLUID MECHANICS AND HEAT TRANSFER
- 34-01 BOUNDARY LAYER TECHNOLOGY
- 34-02 GAS DYNAMICS
- 34-03 FLUIDICS

34-04 FLUID FLOW
34-05 COMBUSTION PHYSICS
34-06 HEAT TRANSFER, BASIC
34-07 REENTRY HEAT TRANSFER
34-08 THERMAL PROTECTION
34-09 ABLATION
34-10 CRYOGENICS
35 INSTRUMENTATION AND PHOTOGRAPHY
35-01 PHOTOGRAPHY
35-02 INFRARED TECHNOLOGY
35-03 INSTRUMENT STANDARDS AND CALIBRATION TECHNIQUES
35-04 TEMPERATURE MEASUREMENT
35-05 PRESSURE MEASUREMENT
35-06 DISPLAY SYSTEMS
35-07 DATA RECORDING
35-08 GAS FLOW MEASUREMENT
36 LASERS AND MASERS
36-01 LASERS AND MASERS
36-02 LASER APPLICATIONS
37 MECHANICAL ENGINEERING
37-01 BEARINGS AND GEARS
37-02 LUBRICATION AND LUBRICANTS
37-03 MACHINING
37-04 FRICTION AND WEAR
37-05 SEALS
37-06 WELDING
37-07 METAL FORMING
37-08 PUMPS
37-09 VACUUM TECHNOLOGY
37-10 NONDESTRUCTIVE TESTING
37-11 TURBOMACHINERY
38 QUALITY ASSURANCE AND RELIABILITY
38-01 QUALITY CONTROL AND RELIABILITY
39 STRUCTURAL MECHANICS
39-01 SHELLS
39-02 STRESSES AND LOADS
39-03 STRUCTURE VIBRATION AND DAMPING
39-04 IMPACT PHENOMENA
39-05 STRUCTURAL FATIGUE
39-06 SANDWICH CONSTRUCTION
39-07 STRESS ANALYSIS
39-08 STRUCTURAL TESTS AND RELIABILITY

GEOSCIENCES
42 GEOSCIENCES (GENERAL)
NO TOPICS AVAILABLE
43 EARTH RESOURCES AND REMOTE SENSING
43-01 EARTH RESOURCES
43-02 GEODESY AND CARTOGRAPHY
44 ENERGY PRODUCTION AND CONVERSION
44-01 ENERGY RESOURCES
44-02 FUEL CELLS AND CHEMICAL BATTERIES
44-03 SOLAR SPACE POWER
44-04 NUCLEAR AUXILIARY POWER
45 ENVIRONMENT POLLUTION
45-01 ENVIRONMENT POLLUTION CONTROL
46 GEOPHYSICS
46-01 UPPER EARTH ATMOSPHERE
46-02 GEOLOGY AND SEISMOLOGY
46-03 GEOMAGNETISM
47 METEOROLOGY AND CLIMATOLOGY
47-01 METEOROLOGICAL SATELLITES
47-02 WEATHER FORECASTING
47-03 MICROMETEOROLOGY
47-04 CLOUD RESEARCH

47-05 METEOROLOGICAL INSTRUMENTS
48 OCEANOGRAPHY
48-01 WATER RESOURCES AND OCEANOGRAPHY

LIFE SCIENCES
51 LIFE SCIENCES (GENERAL)
51-01 BIOLOGY (GENERAL)
51-02 BIOCHEMISTRY
52 AEROSPACE MEDICINE
52-01 AEROSPACE MEDICINE
52-02 CLINICAL MEDICINE
52-03 PHYSIOLOGICAL FACTORS
52-04 BIOLOGICAL RADIATION EFFECTS
53 BEHAVIORAL SCIENCES
53-01 PSYCHOLOGICAL FACTORS
54 MAN/SYSTEMS TECHNOLOGY AND LIFE SUPPORT
54-01 LIFE SUPPORT SYSTEMS
54-02 CREW SAFETY AND PROTECTIVE CLOTHING
54-03 HUMAN ENGINEERING
54-04 MAN-MACHINE SYSTEMS
54-05 BIOINSTRUMENTATION
54-06 ROBOTICS
55 SPACE BIOLOGY
55-01 EXTRATERRESTRIAL LIFE

MATHEMATICAL AND COMPUTER SCIENCES
59 MATHEMATICAL AND COMPUTER SCIENCES (GENERAL)
59-01 APPLIED MATHEMATICS
59-02 DATA PROCESSING
60 COMPUTER OPERATIONS AND HARDWARE
60-01 DIGITAL AND ANALOG COMPUTERS
60-02 AIRBORNE OR SPACEBORNE COMPUTERS
61 COMPUTER PROGRAMMING AND SOFTWARE
61-01 COMPUTER SOFTWARE
61-02 CAD/CAM
62 COMPUTER SYSTEMS
NO TOPICS AVAILABLE
63 CYBERNETICS
63-01 CYBERNETICS AND BIONICS
63-02 ARTIFICIAL INTELLIGENCE
64 NUMERICAL ANALYSIS
64-01 NUMERICAL ANALYSIS
65 STATISTICS AND PROBABILITY
65-01 PROBABILITY AND STATISTICS
66 SYSTEMS ANALYSIS
NO TOPICS AVAILABLE
67 THEORETICAL MATHEMATICS
NO TOPICS AVAILABLE

PHYSICS
70 PHYSICS (GENERAL)
NO TOPICS AVAILABLE
71 ACOUSTICS
71-01 ACOUSTICS
71-02 ULTRASONICS
72 ATOMIC AND MOLECULAR PHYSICS
72-01 ATOMIC PHYSICS
72-02 MOLECULAR PHYSICS
73 NUCLEAR AND HIGH-ENERGY PHYSICS
73-01 NUCLEAR PHYSICS
73-02 RADIOACTIVITY
74 OPTICS
74-01 OPTICS
74-02 LIGHT
75 PLASMA PHYSICS

75-01 PLASMA APPLICATIONS
75-02 PLASMA DYNAMICS
75-03 MAGNETOHYDRODYNAMICS
76 SOLID-STATE PHYSICS
76-01 SOLID STATE DEVICES
76-02 SUPERCONDUCTIVITY
76-03 DIELECTRICS
76-04 EPITAXIAL DEPOSITION
77 THERMODYNAMICS AND STATISTICAL PHYSICS
NO TOPICS AVAILABLE

SOCIAL SCIENCES
80 SOCIAL SCIENCES (GENERAL)
NO TOPICS AVAILABLE
81 ADMINISTRATION AND MANAGEMENT
81-01 AEROSPACE MANAGEMENT
82 DOCUMENTATION AND INFORMATION SCIENCE
82-01 INFORMATION TECHNOLOGY
83 ECONOMICS AND COST ANALYSIS
NO TOPICS AVAILABLE
84 LAW, POLITICAL SCIENCE AND SPACE POLICY
84-01 WORLD SPACE PROGRAMS AND AEROSPACE LAW
84-02 SPACE COMMERCIALIZATION
85 URBAN TECHNOLOGY AND TRANSPORTATION
85-01 URBAN TECHNOLOGY AND TRANSPORTATION

SPACE SCIENCES
88 SPACE SCIENCES (GENERAL)
NO TOPICS AVAILABLE
89 ASTRONOMY
89-01 SOLAR ASTRONOMY
89-02 STELLAR ASTRONOMY AND COSMOLOGY
89-03 METEORS AND METEORITES
90 ASTROPHYSICS
90-01 GRAVITATION
90-02 ASTROPHYSICAL PLASMAS
91 LUNAR AND PLANETARY EXPLORATION
91-01 THE MOON
91-02 PLANETARY SCIENCES AND EXPLORATION
92 SOLAR PHYSICS
NO TOPICS AVAILABLE
93 SPACE RADIATION
93-01 COSMIC RADIATION
93-02 SOLAR RADIATION AND ACTIVITY
93-03 RADIATION BELTS

---------- LISTSERVER Statistics ----------

79 users subscribed to 192 lists.

User's E-mail Address	Number of Listserv Subscriptions
TISO@UDAVXB.OCA.UDAYTON.EDU	17
PEANDER@HOLOGRAM.LERC.NASA.GOV	11
THORNTON@VNET.IBM.COM	10
DSGMAD@CDSLR1.GSFC.NASA.GOV	9
AE773@FREENET.CARLETON.CA	8
JEFF344@VOODOO.LERC.NASA.GOV	8
GHOETKER@STI.NASA.GOV	7
TLYONS%HRTR1.SPAN@FEDEX.MSFC.NASA.GOV	6
FSASHPS@COBY.LERC.NASA.GOV	6
DJLESCO@ARIEL.LERC.NASA.GOV	6
FSJRS@OZ.LERC.NASA.GOV	5
MLN@BLEARG.LARC.NASA.GOV	5
@VTVM1.CC.VT.EDU:DEWOLF@VTVM1.CC.VT.EDU	5
BLEHRER@NHQVAX.HQ.NASA.GOV	5
TRIMET1@AIP.ORG	4
AMY@SCOTLAND.LERC.NASA.GOV	4
JERWIN@STI.NASA.GOV	4
DRESHFIELD#M#_ROBERT_L@LIMS-A1.LERC.NASA.GOV	3
STEVES@ECN.PURDUE.EDU	3
JGRANT@MAIL.CASI.NASA.GOV	3
JOHN@COSMIC.COSMIC.UGA.EDU	2
BMCCARTH@MICKEY.ENG.GULFAERO.COM	2
BAAKLINI#M#_GEORGE@LIMS-A1.LERC.NASA.GOV	2
TOMPOSKI@NPT.NUWC.NAVY.MIL	2
SCAN-26-04@ROCKET.COM	1
GGOTT@BLEARG.LARC.NASA.GOV	1
SCAN-76-02@ROCKET.COM	1
SCAN-44-02@ROCKET.COM	1
SCAN-15-05@ROCKET.COM	1
SCAN-34-07@ROCKET.COM	1
SCAN-34-08@ROCKET.COM	1
SCAN-18-01@ROCKET.COM	1
MSWENSON@ATC.BOEING.COM	1
SCAN-16-01@ROCKET.COM	1
0004229010@MCIMAIL.COM	1
SCAN-26-03@ROCKET.COM	1
SCAN-76-01@ROCKET.COM	1
SCAN-23-01@ROCKET.COM	1
SCAN-34-09@ROCKET.COM	1
SCAN-28-02@ROCKET.COM	1
TDOWLING@LIB.WASHINGTON.EDU	1
SCAN-20-03@ROCKET.COM	1
SCAN-18-02@ROCKET.COM	1
SCAN-15-03@ROCKET.COM	1
SCAN-28-01@ROCKET.COM	1
JPARKER@AURORA.MSFC.NASA.GOV	1
SCAN-26-06@ROCKET.COM	1
SCAN-19-02@ROCKET.COM	1
MSHAPIRO@MAIL.CASI.NASA.GOV	1
SCAN-26-01@ROCKET.COM	1
SCAN-92@ROCKET.COM	1
SCAN-32-01@ROCKET.COM	1
SCAN-25-02@ROCKET.COM	1
SCAN-34-01@ROCKET.COM	1
SCAN-27-02@ROCKET.COM	1
SCAN-26-07@ROCKET.COM	1
BLEHRER@HQ.NASA.GOV	1
SCAN-33-09@ROCKET.COM	1
SCAN-23-03@ROCKET.COM	1

SCAN-75-01@ROCKET.COM	1
SCAN-43-01@ROCKET.COM	1
SCAN-15-02@ROCKET.COM	1
SCAN-27-03@ROCKET.COM	1
RHUGHES@MAIL.CASI.NASA.GOV	1
SCAN-45-01@ROCKET.COM	1
SCAN-76-03@ROCKET.COM	1
SCAN-27-05@ROCKET.COM	1
SCAN-23-04@ROCKET.COM	1
SCAN-75-02@ROCKET.COM	1
SCAN-14-02@ROCKET.COM	1
SCAN-13-01@ROCKET.COM	1
SCAN-25-04@ROCKET.COM	1
SCAN-34-03@ROCKET.COM	1
SCAN-25-01@ROCKET.COM	1
SCAN-43-02@ROCKET.COM	1
MCCREIGHT@NASAMAIL.NASA.GOV	1
SCAN-15-01@ROCKET.COM	1
SCAN-35-08@ROCKET.COM	1
SCAN-20-01@ROCKET.COM	1

11

Users subscribed to 135 SCAN topics
Total of 192 subscriptions

Listserv list	Number of User Subscriptions
SCAN-34-01	4
SCAN-ALL-TOPICS	4
SCAN-20-01	4
SCAN-02	3
SCAN-02-01	3
SCAN-61-01	3
SCAN-61	3
SCAN-NOTIFY	3
SCAN-34-04	2
SCAN-25-04	2
SCAN-34-06	2
SCAN-24	2
SCAN-23-01	2
SCAN-75-03	2
SCAN-81	2
SCAN-91-02	2
SCAN-60-01	2
SCAN-74	2
SCAN-19-02	2
SCAN-03-01	2
SCAN-07	2
SCAN-01	2
SCAN-55-01	2
SCAN-60	2
SCAN-20-03	2
SCAN-23-03	2
SCAN-15	2
SCAN-18-02	2
SCAN-02-03	2
SCAN-59	2
SCAN-39	2
SCAN-63	2
SCAN-03-07	2
SCAN-33	2
SCAN-35-08	2
SCAN-17	2
SCAN-19	2
SCAN-76-01	2
SCAN-07-01	2
SCAN-23-04	2
SCAN-47	2
SCAN-64-01	2
SCAN-36	2
SCAN-16-01	2
SCAN-13-01	2
SCAN-26	1
SCAN-26-06	1
SCAN-32-01	1
SCAN-81-01	1
SCAN-27-03	1
SCAN-34-10	1
SCAN-18-01	1
SCAN-34-02	1
SCAN-02-02	1
SCAN-24-01	1
SCAN-76-03	1
SCAN-26-04	1
SCAN-37	1
SCAN-15-01	1
SCAN-47-01	1
SCAN-03-04	1

```
------------------------------------------------------------
                  Gopher Usage Report
            Beginning Sep 1, 1993 @ 00:54:27
            Ending    Sep 8, 1993 @ 20:42:33
------------------------------------------------------------

                                    Successful     Failed
                                    ----------     ------
otal Number of Connections:
otal Number of Files Retrieved:          222
otal Number of Directories Accessed:     274


ccess Report for Directories and Files

irectory or File Accessed
---------------------------------------------------------------------
can/Changes                              3 ***
Current-SCAN                            13 **********
Ordering-Information                     5 ****
Organization                             3 ***
PRICES                                   6 *****
README                                   8 ******
SCAN-ACCESS-METHODS                     10 ********
SCAN-TOPICS                              9 *******
archive                                 23 ******************
.01                                      2 **
.02                                      2 **
..02-01                                  3 ***
....1993-03-03                           1 *
....1993-09-02                           2 **
..02-03                                  1 *
..04                                     3 ***
..05                                     1 *
..07                                     4 ***
...07-01                                 4 ***
....1993-03-03                           1 *
....1993-03-17                           1 *
....1993-08-05                           1 *
....1993-09-02                           2 **
..12                                     1 *
..14                                     1 *
..16                                     1 *
...16-01                                 1 *
....1993-03-03                           1 *
..17                                     1 *
...17-01                                 1 *
....1993-03-09                           1 *
....1993-08-20                           1 *
....1993-09-02                           1 *
..20                                     1 *
...20-01                                 1 *
....1993-06-16                           1 *
...20-02                                 1 *
...20-03                                 2 **
....1993-03-03                           1 *
....1993-05-18                           1 *
....1993-06-02                           1 *
....1993-06-16                           1 *
....1993-07-02                           1 *
....1993-07-16                           1 *
....1993-08-05                           1 *
```



```
....1993-08-20                           1 *
....1993-09-02                           2 **
..23                                     4 ***
...23-01                                 1 *
....1993-03-03                           1 *
...23-02                                 1 *
....1993-08-05                           1 *
...23-04                                 2 **
....1993-03-03                           1 *
....1993-03-09                           1 *
....1993-07-02                           1 *
....1993-07-16                           1 *
....1993-08-05                           1 *
..25                                     1 *
...25-04                                 1 *
....1993-03-03                           1 *
....1993-06-16                           1 *
..26                                     1 *
...26-06                                 1 *
....1993-09-02                           1 *
..27/27-01                               1 *
....1993-03-03                           1 *
..28                                     1 *
...28-01                                 1 *
....1993-03-03                           1 *
....1993-03-09                           1 *
....1993-03-17                           1 *
....1993-04-07                           1 *
....1993-06-16                           1 *
....1993-07-02                           1 *
....1993-09-02                           1 *
..33                                     2 **
...33-08                                 2 **
....1993-03-03                           1 *
....1993-08-05                           1 *
....1993-08-18                           2 **
....1993-08-20                           1 *
...33-12                                 1 *
....1993-03-03                           1 *
..45                                     1 *
...45-01                                 1 *
....1993-08-20                           1 *
..47/47-01                               1 *
...47-02                                 1 *
....1993-03-03                           1 *
...47-05                                 1 *
..52                                     1 *
...52-01                                 1 *
....1993-03-03                           1 *
...52-02                                 1 *
....1993-03-03                           1 *
..54/54-06                               1 *
....1993-03-03                           1 *
....1993-03-09                           1 *
....1993-03-17                           1 *
....1993-06-16                           1 *
....1993-07-02                           1 *
....1993-07-16                           1 *
....1993-08-05                           1 *
....1993-08-18                           1 *
....1993-08-20                           1 *
....1993-09-02                           1 *
..59                                     1 *
...59-01                                 1 *
....1993-07-02                           1 *
```

```
.63                                1 *
..63-02                            3 ***
...1993-03-03                      1 *
...1993-03-09                      1 *
...1993-03-17                      1 *
...1993-04-07                      1 *
...1993-04-16                      1 *
...1993-05-06                      1 *
...1993-05-18                      1 *
...1993-06-02                      1 *
...1993-06-16                      1 *
....1993-07-02                     1 *
....1993-07-16                     1 *
....1993-08-18                     1 *
....1993-08-20                     2 **
....1993-09-02                     1 *
..64                               1 *
...64-01                           1 *
....1993-09-02                     1 *
..71                               1 *
...71-01                           1 *
....1993-08-05                     1 *
....1993-08-18                     2 **
..75/75-03                         1 *
....1993-03-03                     1 *
....1993-08-18                     1 *
....1993-08-20                     2 **
....1993-09-02                     3 ***
..82                               3 ***
...82-01                           1 *
....1993-08-05                     1 *
....1993-08-18                     2 **
....1993-08-20                     2 **
....1993-09-02                     3 ***
..88                               4 ***
..89                               2 **
...89-02                           1 *
....1993-03-03                     1 *
....1993-09-02                     2 **
..91                               1 *
...91-01                           1 *
....1993-09-02                     2 **
...91-02                           1 *
....1993-03-09                     1 *
....1993-08-05                    42 ******************************
.current                          15 **********
..01                               7 ******
...02                              7 ******
....02-01                          2 **
...05                              1 *
....05-02                          1 *
...07                              1 *
....07-01                          1 *
...09                              1 *
....09-01                          4 ***
..12                               2 **
...15                              1 *
....15-03                          1 *
...16                              1 *
...17                              3 ***
...20                              2 **
....20-01                          1 *
....20-02                          1 *
....20-03                          2 **
..23
```



```
...23                              2 **
....23-04                          1 *
...25                              1 *
...26                              2 **
....26-06                          1 *
...28                              2 **
....28-01                          2 **
..31                              14 **********
...32                              2 **
....32-02                          2 **
...33                              9 *******
....33-03                          1 *
....33-05                          2 **
....33-08                          8 ******
...34                              1 *
....34-06                          1 *
...37                              2 **
....37-01                          1 *
..42                               2 **
...44                              1 *
....44-03                          1 *
...45                              1 *
....45-01                          1 *
...47                              1 *
....47-03                          2 **
....47-04                          1 *
..51                               2 **
...52                              1 *
....52-01                          1 *
..59                               3 ***
...59                              1 *
....59-01                          1 *
...61                              3 ***
....61-01                          2 **
....61-02                          1 *
...64                              1 *
....64-01                          2 **
...65                              1 *
....65-01                          1 *
..70                               1 *
...71                              1 *
....71-01                          1 *
...75                              1 *
....75-02                          1 *
..80                               1 *
...84                              1 *
....84-02                          1 *
..88                              11 ********
...89                              1 *
....89-03                          1 *
...91                              4 ***
....91-01                          2 **
....91-02                          1 *
...93                              5 ****
....93-01                          1 *
....93-02                          9 *******
....93-03                          4 ***
```

TOTALS FOR SUMMARY PERIOD Mon Mar 8 1993 TO Wed Sep 8 1993

Files Transmitted During Summary Period	225
Bytes Transmitted During Summary Period	8427572
Systems Using Archives	0
Average Files Transmitted Daily	8
Average Bytes Transmitted Daily	290606

Daily Transmission Statistics

Date	Number Of Files Sent	Number of Bytes Sent	Average Xmit Rate	Percent Of Files Sent	Percent Of Bytes Sent
Mon Mar 8 1993	5	588762	5.5 KB/s	2.22	6.99
Tue Mar 9 1993	2	52280	26.1 KB/s	0.89	0.62
Wed Mar 10 1993	1	50330	50.3 KB/s	0.44	0.60
Tue Mar 30 1993	4	107696	4.3 KB/s	1.78	1.28
Thu Apr 29 1993	2	4323	2.2 KB/s	0.89	0.05
Wed May 5 1993	1	158	0.2 KB/s	0.44	0.00
Wed May 26 1993	2	9593	3.2 KB/s	0.89	0.11
Fri Jun 4 1993	4	170948	3.1 KB/s	1.78	2.03
Wed Jun 16 1993	4	70496	2.4 KB/s	1.78	0.84
Thu Jun 17 1993	8	16498	1.8 KB/s	3.56	0.20
Tue Jul 13 1993	4	14857	3.0 KB/s	1.78	0.18
Fri Jul 16 1993	2	11413	5.7 KB/s	0.89	0.14
Mon Jul 19 1993	3	12410	3.1 KB/s	1.33	0.15
Tue Jul 20 1993	1	8542	4.3 KB/s	0.44	0.10
Wed Jul 21 1993	8	687804	0.9 KB/s	3.56	8.16
Fri Jul 23 1993	2	2170	1.1 KB/s	0.89	0.03
Tue Aug 3 1993	3	521	0.2 KB/s	1.33	0.01
Mon Aug 9 1993	3	12145	3.0 KB/s	1.33	0.14
Mon Aug 23 1993	1	9439	9.4 KB/s	0.44	0.11
Tue Aug 24 1993	8	104424	4.2 KB/s	3.56	1.24
Wed Aug 25 1993	20	592520	1.8 KB/s	8.89	7.03
Fri Aug 27 1993	1	1974	2.0 KB/s	0.44	0.02
Mon Aug 30 1993	10	239155	5.2 KB/s	4.44	2.84
Tue Aug 31 1993	2	58266	5.8 KB/s	0.89	0.69
Wed Sep 1 1993	7	229789	5.9 KB/s	3.11	2.73
Thu Sep 2 1993	2	10291	5.1 KB/s	0.89	0.12
Fri Sep 3 1993	31	1223217	4.7 KB/s	13.78	14.51
Tue Sep 7 1993	69	3255968	5.9 KB/s	30.67	38.63
Wed Sep 8 1993	15	881583	5.6 KB/s	6.67	10.46

15

Total Transfers from each Archive Section

Archive Section	Files Sent	Bytes Sent	Percent Of Files Sent	Percent Of Bytes Sent
scan	67	212875	29.78	2.53
scan/archive	49	3634900	21.78	43.13
scan/current	109	4579797	48.44	54.34

1

Total Transfer Amount By Domain

Domain Name	Number Of Files Sent	Number of Bytes Sent	Average Xmit Rate	Percent Of Files Sent	Percent Of Bytes Sent
it	8	16498	1.8 KB/s	3.56	0.20
uk	2	4323	2.2 KB/s	0.89	0.05
com	8	85353	2.5 KB/s	3.56	1.01
edu	42	982091	2.3 KB/s	18.67	11.65
mil	1	18079	4.5 KB/s	0.44	0.21
net	8	687804	0.9 KB/s	3.56	8.16
nasa.gov	145	6482040	5.4 KB/s	64.44	76.91
unresolved	11	151384	9.5 KB/s	4.89	1.80

These figures only reflect ANONYMOUS FTP transfers. There are many sites which mount the archives via NFS, and those transfers are not logged and reported by this program.

17

Top 15 Most Popular Archive Sections By Bytes Transferred

Archive Section	Files Sent	Bytes Sent	---- Percent of ---- Files Sent	Bytes Sent
scan/current	109	4579797	48.44	54.34
scan/archive	49	3634900	21.78	43.13
scan	67	212875	29.78	2.53
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00
			0.00	0.00

Machine Translation Project

Katie Bajis

Existing Translation Systems

We were looking at existing machine translation systems that are real-world; that is, they're in use, not under development (Viewgraphs 1-4). These systems at the top of the chart are two of the systems that we looked at. This is SYSTRAN. It was developed by the U.S. Air Force, and has been around for about 30 years. GLOBALINK: this has come on the scene in the last four or five years. STYLUS was developed in Moscow; it's questionable by which organization. PC-TRANSLATOR is a small PC system. Of the systems we looked at, only the first four were really in the running. SYSTRAN was at the top of the list, essentially because of the dictionary subjects that were available. SYSTRAN has, on the current version that we have at CASI, 16 subject dictionaries. None of the other systems, PC or mainframe have the same range of subjects or the same size dictionaries. The dictionaries for SYSTRAN for Russian alone are about 240,000 words. French and German versions have their own dictionaries.

Critical Factors

Denise Bedford and I started to look at the critical factors: size, nature of the subject dictionaries, and user friendliness. Well, SYSTRAN isn't all that user friendly. Even in its more user friendly versions, because of the size of the dictionaries and the complexity of the software itself, it was still not as easy to use as GLOBALINK. GLOBALINK has what I would call the maximum amount of user friendliness for the average person. Because of its user friendliness, GLOBALINK 2 was considered an excellent option for procurement as a supplement to SYSTRAN. Some of the subject dictionaries available from GLOBALINK are not covered by SYSTRAN at all. There are some business, legal, and finance dictionaries that are not available from SYSTRAN. So, Denise and I figured that some of these supplemental dictionaries that are available under GLOBALINK and not available under SYSTRAN would be useful to NASA right now, particularly because of the Russian space initiatives and efforts. Additionally, GLOBALINK is coming out with a reverse capability for Russian; it's due out in the next month or two. We figured that reverse capability would be useful in some cases, again because of the Russian initiatives. STYLUS was a package that was offered to us, essentially free from a Russian group that was visiting here. They briefed us on some of the capabilities.

Translation Tests

Denise and I tried a couple of sentences that we had already produced through SYSTRAN. The test case was a sentence from an astrophysics text. The translation came back nearly

identical, which sort of surprised us because, if the dictionary was so small, we were kind of wondering why it would have a lot of this technical vocabulary. In discussing this and talking to a couple of people who developed SYSTRAN, we came back with some information on the possibility that maybe it wasn't all developed in Moscow. We still don't know. It was offered to us for free. It has some very interesting capabilities. For one, it has the reverse capability of English to Russian. It also does Russian to Italian, German, French, and Spanish and vice versa. We may use it as a chain capability to go from Italian to Russian to English because all of the software is stored in the same PC configuration. With PC-Translator, we decided to procure only Italian because it was not covered by GLOBAL, SYSTRAN, or STYLUS, and we thought that we might need some Italian capability. So, those are the four basic systems that we were recommending for purchase or acquisition. With SYSTRAN, we don't have to purchase the software, only the hardware.

Question: What platform will you use?

Answer: SYSTRAN is going to be on the PS2, OS2; the others will run on PC's. But they're all going to be installed on the same hardware. The SYSTRAN system in the configuration we've decided on is, for the most part, very, very similar to what the Air Force is using at FASTC; they developed the system. They are putting theirs on PS2's. It will be possible for four or five simultaneous users to get into the system for SYSTRAN. GLOBAL, STYLUS, and PC-Translator will be available to one person at a time.

System Features

I'll quickly go over some of the system features that we looked at. Dictionary characteristics that we looked at were whether they could handle phrases or idioms, abbreviations, acronyms, or glossary creation; that was considered to be very, very important. GLOBALINK allows you to create your own dictionaries, to build them as you go along, to save some of the information from doing translations with corrections. You can save some of the information and store it in a file. With input file formats, we needed to be able to take machine readable text from various foreign languages and run it through the system; so, the input file system was also a consideration.

Basically, this chart was meant to show, in a graphic way, the capabilities we would get if we purchased or acquired four different systems and put them on the same configuration of workstations (Viewgraph 5). These are the subject dictionaries that are available, the primary languages, and the reverse languages. The subject dictionaries and which language pairs were needed were primary considerations. If we got all of the language pairs available for SYSTRAN, Globalink, and PC-Translator, these shaded areas would be all the languages, pretty much, that we can cover.

Procurement Recommendations

Now, it was our recommendation that we not necessarily get all of these languages because some of these, like Norwegian, Swedish, Korean, Portuguese, and Greek are not all that

important to us at this time. So, the basic languages that we are going to offer are Russian, French, German, and Japanese, when it becomes available. None of the systems we've looked at can do Japanese, although SYSTRAN has a Japanese-English, English-Japanese system that, I believe, is under development now. Spanish is also covered under the SYSTRAN system that we're going to get. So, basically, we're only going to have four languages covered under the configuration and language pairs that we've selected so far. In the first couple of months of project implementation, we're going to collect information on what other language pairs users would be interested in. We'll probably decide in another year or two which of these language pairs would be most important for our program to have.

Work Flow

Question: What is the work flow that you anticipate?

Answer: It could be tremendous. There has been a lot of interest. We know that a lot of the users take the attitude that, if there's a magic box available, they'll use it. The system is configured so that a user will send us a fax. It will go to a machine here, and a computer will pick it up from a board so that we don't have to scan it. In most cases, we're going to have some capability to convert this into some sort of foreign language ASCII, and then that machine-readable text is going to be processed through in a raw format, and we're not going to do any editing. We're going to give it back to the user either by fax or by e-mail in a raw, unedited form so that there will be the least amount of processing to do here. One reason we want to do that is to see whether users will accept the raw, unedited stuff. We want to get them exposed to that because it's the least expensive and will only require the services of an operator who doesn't necessarily know the foreign language. We're going for the lowest common denominator at the beginning to see what the user wants and whether we can make the system run with someone who doesn't know a foreign language. In most cases, we are not going to be able to have a competent translator doing the editing and proof reading at every site. So, the volume could be extremely high. We're going to determine, probably on the basis of cost, what we're going to do in the future.

Factor	Weight	SYSTRAN	GLOBALINK	STYLUS	PC-TRANSLATOR	M(A)ICROCAT	LEXITRANS	WINTER 92	MICROTAC ASSISTANTS	SOCRATA/XLT	TOLTRAN	INTERGRAPH	ALPS	LOGOS	TOVNA/MTS
User Friendly Interface	2		O	O					O	O	O	O			
Vendor Experience & Support	3	O	O	d	O										
Multiple Standard Platforms	3	O	PC	PC	PC	O	PC	PC	PC	PC	PC	O	PC	PC	PC
Primary Languages															
Chinese	2		d								d				
French	3	O	O		O	O			O	O	O	O	O	O	O
German	3	O	O		O	O			O	d		O			d
Italian	2	d			O	O				O	d				d
Japanese	3	d				O									
Russian	3	O	O	O			O	d		d	O				O
Spanish	2	O	O		O	O		O	O	O		O			
Danish	1				O			O							
Norwegian	1				O						O				
Swedish	1				O										
Korean	1	O													
Reverse Language Pairs															
French	2	O	O		O	O			O	O			O	O	
German	2	O	O			O			O				O	O	
Japanese	1	O													
Russian	2	O	O	O						d	d				
Chinese	1		d												
Portuguese	1	O			O	O									
Norwegian	1	O													
Korean	1	O													
Greek	1	O													
Dutch	1	O													
Finnish	1	O													
Swedish	1	O			O										
Danish	1	O			O										
Italian	1	O			O	O				O	d		O	O	
Spanish	1	O	O		O	O			O	O	d		O	O	
Dictionary Subjects															
General	3	O	O	d	O	O		O	O	O	O	O	O	O	O
Aeronautics	3	O	O												
Astronautics	3	O		O											
Astronomy	3	O													
Physics	3	O													
Chemistry & Materials	3	O													
Life Sciences	3	O													
Mathematics	3	O													
Computer Science	3	O	O												
Engineering	3	O	d												
Geosciences	3	O													
Space Sciences	3	O													
Social Sciences	3	O													

Table 1. Characteristics and Capabilities of All Machine Translation Products Evaluated

Factor	Weight	SYSTRAN	GLOBALINK	STYLUS	PC-TRANSLATOR	M(A)ICROCAT	LEXITRANS	WINTER 92	MICROTAC ASSISTANTS	SOCRATA/XLT	TOLTRAN	INTERGRAPH	ALPS	LOGOS	TOVNA/MTS
User Friendly Interface	2		O	O					O	O	O	O			
Vendor Experience & Support	3	O	O	d	O										
Multiple Standard Platforms	3	O	PC	PC	PC	O	PC	PC	PC	PC	PC	O	PC	PC	PC
Primary Languages															
Chinese	2		d								d				
French	3	O	O		O	O			O	O	O	O	O	O	O
German	3	O	O		O	O			O	d		O			d
Italian	2	d			O	O				O	d				d
Japanese	3	d				O									
Russian	3	O	O	O				d		d	O				O
Spanish	2	O	O		O	O		O	O	O		O			
Danish	1				O			O							
Norwegian	1				O						O				
Swedish	1				O										
Korean	1	O													
Reverse Language Pairs															
French	2	O	O		O	O			O	O			O	O	
German	2	O	O			O			O				O	O	
Japanese	1	O													
Russian	2	O	O	O						d	d				
Chinese	1		d												
Portuguese	1	O			O	O									
Norwegian	1	O													
Korean	1	O													
Greek	1	O													
Dutch	1	O													
Finnish	1	O													
Swedish	1	O			O										
Danish	1	O			O										
Italian	1	O			O	O				O	d		O	O	
Spanish	1	O	O		O	O			O	O	d		O	O	
Dictionary Subjects															
General	3	O	O	d	O	O		O	O	O	O	O	O	O	O
Aeronautics	3	O	O												
Astronautics	3	O		O											
Astronomy	3	O													
Physics	3	O													
Chemistry & Materials	3	O													
Life Sciences	3	O													
Mathematics	3	O													
Computer Science	3	O	O												
Engineering	3	O	d												
Geosciences	3	O													
Space Sciences	3	O													
Social Sciences	3	O													

Table 1. Characteristics and Capabilities of All Machine Translation Products Evaluated

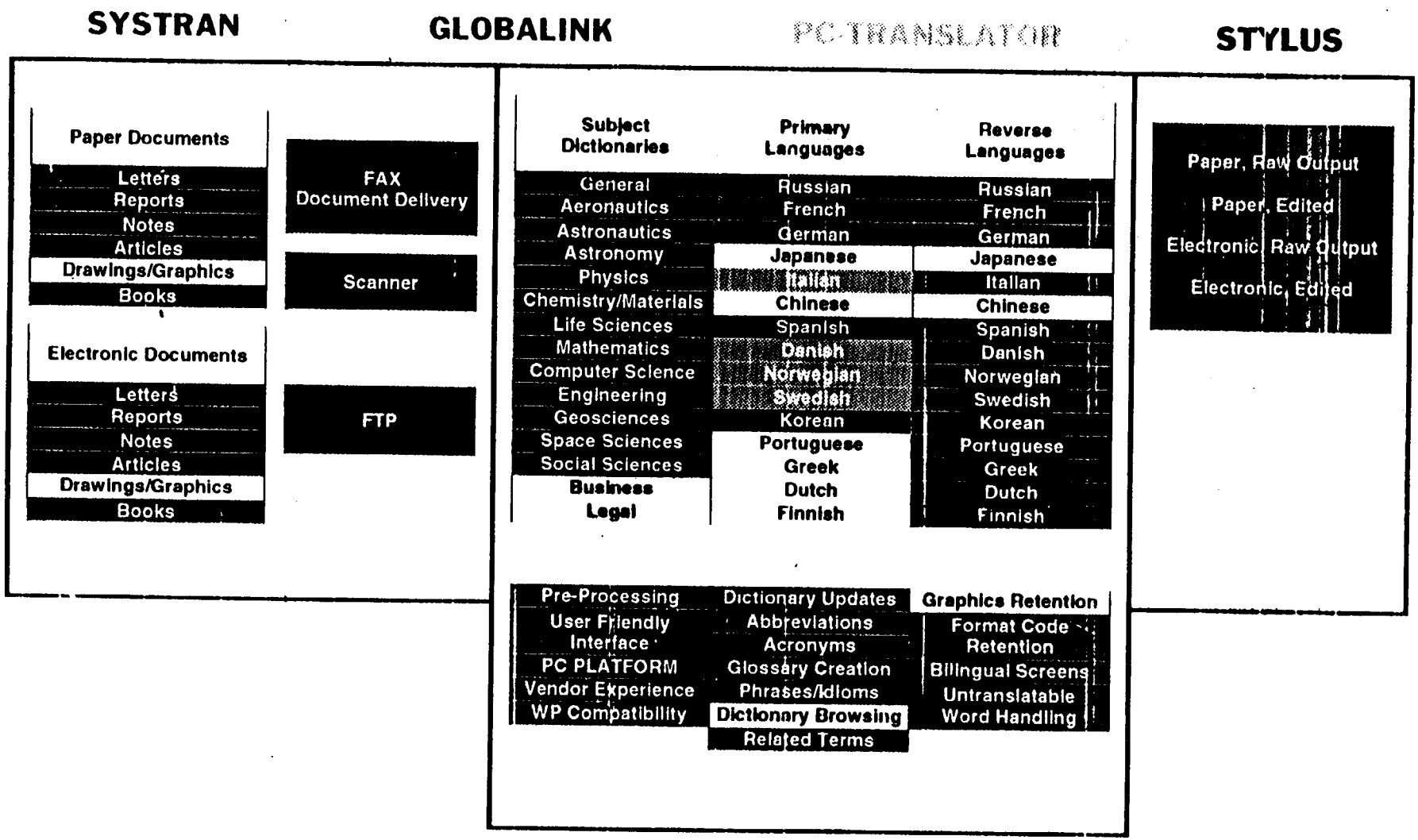
Factor	Weight	SYSTRAN	GLOBALINK	STYLUS	PC-TRANSLATOR	M(A)ICROCAT	LEXITRANS	WINGER 92	MICROTAC ASSISTANTS	SOCRATA/XLT	TOLTRAN	INTERGRAPH	ALPS	LOGOS	TOVNA/MTS
Other Dictionary Features															
Phrases/Idioms	3	O	O	O	O		O								
Abbreviations	3	O		O	O		O								
Acronyms	3	O		O			O								
Glossary Creation	3	O	O	O	O		O	O							
Update Rates	3	O	d	d											
User Modifications	3	O	O	O	O		O	O							
Input File Formats															
English ASCII	3		O					O		O					
English WordPerfect	3		O					O	O		O				
Non-Cyrillic ASCII	3		O		O	O									
Foreign WordPerfect	3	d	O	O		O									
Russian ASCII	3		O	O											
Graphics Capabilities	3		O												
OCR Capabilities	3	o	O	O			O								
Object Language Handling															
Spell Checkers	1		O	O		O				O					
Grammar Checkers	1		O	O		O				O					
Processing Features															
Related Terms	1	O	O	O				O							
Dictionary Browsing	1			O											
Batch/Interactive	1	O	O	O		O		O	I	O			O	B	B
Bilingual Screens	1	O	O	O		O		O	O						
Untranslatable Word Han	1		O	O		O		O		O	O				
Target Language Handling															
Spell Checkers	1		O			O									
Grammar Checkers	1		O			O									
Product and Delivery Options															
English ASCII	3	O	O	O				O							
English WP	2		O	O				O							
Graphics Format Retentic	2														
TOTAL SCORE	151	98	75	54	42	39	24	24	23	18	15	14	13	12	7

3

Table 1. Characteristics and Capabilities of All Machine Translation Products Evaluated (continued)

Factor	Weight	SYSTRAN	GLOBALINK	STYLUS	PC-TRANSLATOR	M(A)ICROCAT	LEXITRANS	WINGER 92	MICROTAC ASSISTANTS	SOCRATA/XLT	TOLTRAN	INTERGRAPH	ALP3	LOGOS	TOVNA/MTS
Other Dictionary Features															
Phrases/Idioms	3	O	O	O	O		O								
Abbreviations	3	O		O	O		O								
Acronyms	3	O		O			O								
Glossary Creation	3	O	O	O	O		O	O							
Update Rates	3	O	d	d											
User Modifications	3	O	O	O	O		O	O							
Input File Formats															
English ASCII	3		O					O		O					
English WordPerfect	3		O					O	O		O				
Non-Cyrillic ASCII	3		O		O	O									
Foreign WordPerfect	3	d	O	O		O									
Russian ASCII	3		O	O											
Graphics Capabilities	3		O												
OCR Capabilities	3	o	O	O			O								
Object Language Handling															
Spell Checkers	1		O	O		O				O					
Grammar Checkers	1		O	O		O				O					
Processing Features															
Related Terms	1	O	O	O				O							
Dictionary Browsing	1			O											
Batch/Interactive	1	O	O	O		O		O	I	O			O	B	B
Bilingual Screens	1	O	O	O		O		O	O						
Untranslatable Word Han	1		O	O		O		O		O	O				
Target Language Handling															
Spell Checkers	1		O			O									
Grammar Checkers	1		O			O									
Product and Delivery Options															
English ASCII	3	O	O	O				O							
English WP	2		O	O				O							
Graphics Format Retentic	2														
TOTAL SCORE	151	98	75	54	42	39	24	24	23	18	15	14	13	12	7

Table 1. Characteristics and Capabilities of All Machine Translation Products Evaluated (continued)


SYSTRAN
GLOBALINK
PC-TRANSLATOR
STYLUS
Paper Documents
Letters
Reports
Notes
Articles
Drawings/Graphics
Books
FAX
Document Delivery
Scanner
Electronic Documents
Letters
Reports
Notes
Articles
Drawings/Graphics
Books
FTP
Subject Dictionaries
General
Aeronautics
Astronautics
Astronomy
Physics
Chemistry/Materials
Life Sciences
Mathematics
Computer Science
Engineering
Geosciences
Space Sciences
Social Sciences
Business
Legal
Primary Languages
Russian
French
German
Japanese
Italian
Chinese
Spanish
Danish
Norwegian
Swedish
Korean
Portuguese
Greek
Dutch
Finnish
Reverse Languages
Russian
French
German
Japanese
Italian
Chinese
Spanish
Danish
Norwegian
Swedish
Korean
Portuguese
Greek
Dutch
Finnish
Paper, Raw Output
Paper, Edited
Electronic, Raw Output
Electronic, Edited
Pre-Processing
User Friendly Interface
PC PLATFORM
Vendor Experience
WP Compatibility
Dictionary Updates
Abbreviations
Acronyms
Glossary Creation
Phrases/Idioms
Dictionary Browsing
Related Terms
Graphics Retention
Format Code Retention
Bilingual Screens
Untranslatable Word Handling
= SYSTRAN Capabilities Only
= GLOBALINK Capabilities Only
= PC-TRANSLATOR Capabilities Only
= STYLUS Capabilities Only
= Capabilities Common to Two or More

N94- 36847

Multimedia

Karen Kaye

Definition

Multimedia has been defined to mean applications that include graphics, text, sound, video, and animation (Viewgraphs 1 and 2). It need not necessarily include all of the above. In fact, it may only include a single item. We want to emphasize the things that we can do with multimedia within the program; that is, the interactive learning, training kind of uses, the audio-visual uses, the presentation-display uses, and multimedia publications.

Classification

If you try to define a classification scheme for multimedia, you essentially can define it in terms of three application types (Viewgraph 3). Initially, this was done by someone at Apple. First of all, we have the narrative - the watch and listen type of multimedia. Now, this is what we're all used to. Everyone watches television; anyone who doesn't is very strong-willed. Additionally, we have interactive multimedia in which the user of the application can choose and do. I, as the user, essentially guide the application along different pathways. Some of you here have seen the Columbus video which we show, which was, essentially, done commercially. There is also participative multimedia, which isn't out there too much now. What it does is allow the end user to contribute and create additional multimedia applications. There is a prototype being done by the British Film Institute which provides that capability.

Multimedia Initiative Objectives

Now, what are the multimedia initiative objectives that we have within our project? (Viewgraphs 4 and 5) First of all, we want to verify the economic and technical feasibility of delivering multimedia within the STI Program. We also want to make a positive educational and informational impact and develop an exploitable capability that can be used by others and in other applications. Why did we get started in multimedia? Not because it's the buzzword of the 90's, but rather because we realized, as a result of our user studies and talking to people at the Centers, that the user population out there is producing multimedia now. There are virtual reality applications being done at some of the NASA Centers, and there are traditional multimedia applications. Many of you have seen the kiosk applications at the Air and Space Museum and the interactive kiosks at Goddard's Visitor's Center - these are all multimedia applications. We also recognize a need, in talking to the public affairs people and NASA video producers, to provide a union catalog, first of all for videos, then other types of multimedia, to facilitate dissemination. We have talked to scientists and engineers who know of multimedia applications that have been developed. They know something was done about a year ago, and maybe they know who did it, but the person retires, and that person took the

application with him. So, it becomes, essentially, lost STI, lost information. Thus, we realize that there's a need to capture this information and make it available to the users. Additionally, we want to provide enhanced user support, management support, training, etcetera. What's being done at the NASA Centers now? As part of this project, we conducted a survey of the different types of non-print media currently used (Viewgraph 6). Videos, motion pictures, and especially CD-ROMs and laser disks are often used for multimedia applications. And, as you'll note, there are a couple of Centers that are leading in production of multimedia applications.

Non-Print Project

Now, within the multimedia initiative, the first big project we have is to acquire NASA produced non-print material and handle and disseminate it practically. We call that our Non-Print Project. Within that project, we have an initial focus on videos. Over a year ago, we completed a project plan (Viewgraph 7). We participated in the CENDI Working Group that worked to define non-print submission guidelines. We identified interim CASI procedures for immediate handling and dissemination of non-print. Essentially, we found that there was a backlog of non-print there that needed to be input into the system and made available for dissemination. We also formed an STI Program video guidelines advisory group and held several video teleconference meetings. We had a lot of interesting NASA Center participation there. We've essentially been told that, once we get our services in this area up and running, that we will be saving NASA, as an agency, millions of dollars.

Non-Print Project Procedures

We also have been discussing the non-print procedures that are already in place in similar organizations. Johnson Space Center has a big video archive, and we have been working very closely with them. We also produced an initial print product, the *NASA Headquarters Public Affairs Video Catalog*, which lists public affairs videos available to the community (Viewgraph 8). We have been addressing some of the particular issues related to the video portion of non-print, such as how to package videos, how to label videos, whether we should have a kit where we could include material along with the video - questions of that nature. The CENDI cataloging working group has been dealing with the non-print media cataloging issues; for instance, the report documentation page annotation that will eventually become standard that will allow for the proper entry of non-print materials. We are now working to determine the Center for AeroSpace Information final prototype procedures for non-print, and we have been extremely busy. Lots of people that have been involved in that area have been helpful. Our video facility, initially, will be at the CG4; it's very small, but once we get past the prototype stage, the video facility will be at the Center for AeroSpace Information. We are also nearing completion of a video catalog that will be disseminated far more widely than the first and will go out to the Air and Space Museum, Johnson Space Center's Space Center Houston and some of the other facilities that cater to the public and provide NASA information. The point is to make this information available as widely as possible.

Video Projects

Just a little more on the video projects. Specifically, the point, of course, is to have a central repository for the dissemination of NASA produced video (Viewgraph 9). When we first started out, we thought that we would handle all of the NASA videos, but soon realized that this is beyond reasonable scope. There were Centers that held 20,000 videos, like Kennedy Space Center. Archiving all of these is really beyond our scope. There would not be much of a demand for some of the information, for example, raw video footage of a plant growing at zero gravity. We would not want to view 48 hours of plant growth; that is very slow. Additionally, we are trying to build our expertise. Several people here have acquired some expertise in this area, and we also have Patrick Curran, who has experience with actually shooting production videos. We are trying to develop more experience in-house in this area. In terms of our other multimedia projects, one we are considering will be a global change project that will deal with a subset of information within a single global change domain.

In-House Support System

We are also looking at doing an in-house support system (Viewgraph 10). What this would provide essentially would be an interactive capability to provide performance support. What support information do I need to make my job better and more effective relative to everyone within the program? This includes training as well as access to information. Because NASA is such a large agency and there is so much being done, it is always a challenge to have available the information that is needed. So this is what we are going to be looking at in this project. Now, we also went out and got equipment needed to support these projects, piece by piece. We don't have all of it yet. We have some that we just ordered and some ordered months ago. The software chosen was essentially selected to facilitate application development while providing cross-platform playback portability.

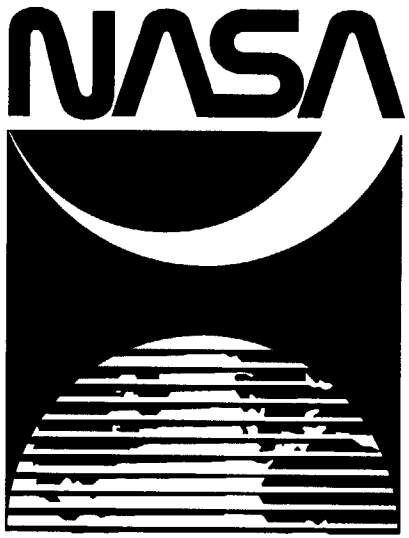
Hardware and Software Testing

We are also currently testing hardware and software that has just come in, and one of the problems there is that we really don't have the staff we need. We are looking at bringing in more staff, and when that happens I think we will get more done. In terms of testing, we recently completed testing of video compression algorithms included in our system. Our hardware compression board is a New Video Eye Q, and the reason for choosing that one was essentially their promise of supporting multiple codecs, multiple compression-decompression algorithms. In fact, the board does that quite successfully. It supports six different codecs, and the vendor is planning on adding a new one. It also has very good quality. What I would like to do is show you a NASA video. This is digital video which will be shown at the full NTSC rate, which is 30 frames per second and gives you an example of the type of quality that is capable with our system. (*The video demonstration follows.*) What I am trying to show you

here is that the technology is essentially here. There are networking questions that needed to be addressed, but we are seeing more and more in the way of available technology that can be used to deliver multimedia information.

NASA Scientific and Technical Information Program

– Multimedia Initiative –



STI PROGRAM
SCIENTIFIC &
TECHNICAL
INFORMATION

Karen Kaye
September 9, 1993

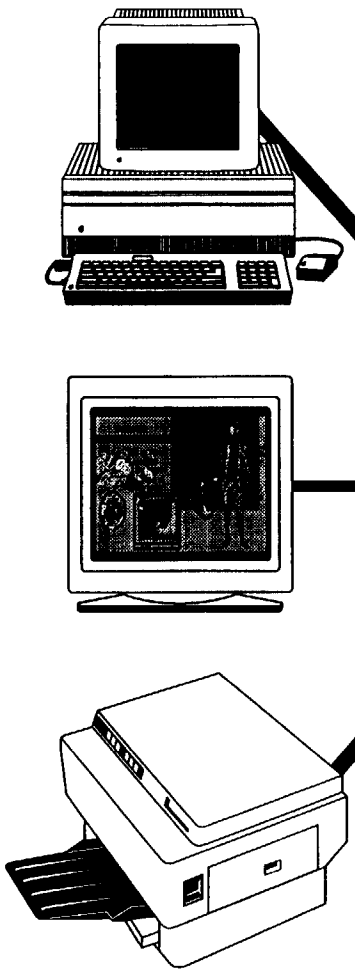
NASA STI PROGRAM
SCIENTIFIC & TECHNICAL INFORMATION

Multimedia - Graphics, Text, Sound, Video, and Animation

Interactive Power

Audiovisual Power

Publishing Power



Multimedia

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Multimedia Initiative Objectives

- Verify the economic and technical feasibility of delivering multimedia information within the STI Program framework
- Make a positive educational & informational impact
- Develop an exploitable capability

71

NASA STI PROGRAM
SCIENTIFIC & TECHNICAL INFORMATION

A Multimedia Classification Scheme

Application Type	User Interaction	Example
Narrative	Watch & Listen	Television
Interactive	Choose & Do	Columbus
Participative	Contribute & Create	British Film Institute

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

STI Program Multimedia Initiative

- Driven by changing user population/requirements
  - Publications
  - Central Catalog/Dissemination
  - User Support
  - Management Support

73

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Types of Non-Pring Media Currently Used

Ames	**Videos, Motion Picture, Diskette, CD-ROM, Laserdisk, Photos**
Dryden	**Videos, Motion Pictures, Diskette, CD-ROM, Photos**
Kennedy	**Video, Motion Pictures, Photos**
Goddard	**Videos, Photos**
JPL	**Videos, Audiocassettes**
Langley	**Videos, Diskette, Laserdisk, Magnetic Tape, CD-ROM**
Johnson	**Videos, Motion Picture, Audiocassettes**
Marshall	**Videos, Films, Photos**
Stennis	**Videos, Motion Picture, Diskette, Laserdisk, Photos**
Lewis	**Videos, Motion Picture, Audiocassettes, Magnetic Tape, Photos**

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Non-Print Project Status

- Completed detailed project plan
- Participated in CENDI Working Group to define non-print submission guidelines
- Identified interim CASI procedures for immediate handling and dissemination of non-print
- Formed STI Program Video Guidelines Advisory Group and held ViTS meetings
- Completed survey of Centers regarding current non-print production and future non-print publication plans
- Discussed non-print procedures with organizations currently handling non-pring (DTIC, COSMIC, National A-V Center, JSC Archive, etc.

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Non-Print Project Status (cont.)

- Responded to request to produce NASA HQ Public Affairs video catalog
- Completed first video catalog
- Addressed packaging, labeling issues for videos
- Worked with Centers to acquire copies of videos listed in HQ Public Affairs video catalog
- CENDI Cataloging Working Group has made progress on non-print RDP, but it is not final
- Currently working to determine CASI final prototype procedures for non-print
- Completing setup of video dubbing facility
- Near completion of second video cataloging

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Video Project

- **Central Repository / Dissemination of NASA produced videos**
- **Provides entry to field**
- **Builds on existing expertise**
- **Develops experience**

77

NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Other Multimedia Initiative Projects

- **Two are just getting underway**
  - **Global Change**
  - **Performance Support System**
- **Equipment needed to support these projects has been arriving in a piecemeal fashion**
  - **Software chosen was selected to facilitate application development for both MAC and PC**
  - **Currently testing hardware and software received**

N94-36848

Electronic Document Interchange

Dick Tuey

DocuTech

Many of you here today are already aware of the Scientific and Technical Information Office's goal in the evaluation of a networked DocuTech at the Center for AeroSpace Information (Viewgraphs 1 and 2). To bring you up to date, the following phasing schedule shows the progress completed to date on the evaluation project. Specficially, the diamonds portrayed on the schedule are milestones that have been completed. I might add that progress is being made and that we will meet the target date of having a completed Evaluation Report by early January 1994 (Viewgraph 3). One of the objectives of the evaluation is the transmittal of a desktop publication such as produced by WordPerfect with figures and graphics imbedded in the text to the Evaluation DocuTech located at CASI.

Document Transmission Tests

I might also add that, from the same workstation at NASA Headquarters, we are using the same document and sending it to the NASA Headquarters mainframe computer's laser printer, the Xerox 4090 (Viewgraph 4). Using this same scenario, the identical document is being sent to the networked DocuTech located at the Lewis Research Center and to an Apple Laser Writer II connected via the Code J LAN. The Apple Laser Writer II is located within the STI office. The purpose of this exercise is to identify any transmission speed and communication protocol problem associated with the sending of documents to various laser printers located throughout the Headquarters site. At this time, the advantage of sending a publication to the networked DocuTech at LeRC is that the publication comes back to you as a finished publication. At this time, job ticket software which spells out what the user wants as an end product is still lacking within the Wide Area Network environment. As part of the evaluation, we are exploring the best means to handle this problem.

Mail Merge and Bar Coding for Postage

Another objective that the evaluation is to demonstrate is the capability to do mail-merge and bar coding for postage. The idea here is to enable the delivery of the publication to recipients designated by the publisher of the document as it would be defined and identified on the job ticket. To ensure that the evaluation DocuTech at CASI will be able to demonstrate all the advertised functionality, an extensive benchmark demonstration test is planned to be performed by mid-December 1993. Results of the evaluation will be documented and presented in a written draft report within one week of the completed benchmark date.

Document Size

Question: What is maximum document size?

Answer: I don't really know. So far there does not appear to be a limitation, but I am sure that one exists. The largest single file that I have sent to LeRC has been around 20 million bytes; what I am finding out is that the Internet is pretty fast. At this time, I won't attempt to go into detail about the specific steps that one needs to go through to get a document printed. However, all the specifics will be available in the Evaluation Report which will be available for general distribution in March 1994.

Job Size Limitations

Question: I am trying to translate a document 180 pages long on 8 x 11.5 paper. Would this be practical for issuing 200 copies to area centers?

Answer: No. You probably want to go ahead electronically, and they would print it out themselves. In other words, if you have it available electronically and it is under the GPO thresholds, then you can bypass the GPO. For those publications that are high TMs and fall within the 25,000 page threshold, they could do them themselves - print their publications; send them out; do them on the DocuTech instead of sending them to GPO. We can electronically send the same publication to CASI, and, as long as we stay under the 25,000, CASI could then reprint the document and send it out. The objective is that we electronically send and try to minimize the print products at the local site. The goal here is to provide the user print-on-demand; that is, only print what is needed as a finished publication. To ensure that only high quality publications are printed, it is recommended that each publisher have an technical editor review the publication before it is sent out. Essentially, this the procedure that is followed by LeRC. At LeRC, a user cannot have technical documents printed unless a technical editor concurs. Therefore, for all print jobs, a job ticket or order must be signed by the appropriate authority plus the editor's signature. The printing office supervisor doesn't even allow you to send a job through unless they see the editor's signature on the job order.

Question: What software are they using on the 4090?

Answer: The Xerox 4090 accepts postscript files from the user workstation such as a PC and MacIntosh.

Print-On-Demand

On Electronic Document Interchange, or more specifically, the print-on-demand through the retrieval of electronically stored documents, I would like to give an update of where we are. I

think a lot of you have heard about the project. Just to give you a little bit of information: this project started about a year and a half ago and basically started as a result of the shutdown of the printing facilities throughout all of NASA Centers (Viewgraph 5). Except for Kennedy Space Center, all of NASA's printing facilities have been closed and, as a result of that, what we are looking for is a cost effective alternative for each Center to do their duplicating (Viewgraph 6). I don't think people are aware that the joint committee on printing has very stringent rules about publications. With the STI information services, it means anything over 25,000 total in terms of groups, and 5,000 for a single print. For example, with a flyer or brochure, you have to send the material to the General Printing Office. When we talk about electronic publishing per se, these systems are high production systems, so it's not like running off print on the laser printer located on your desk. These machines print at 135 pages per minute, and the type of resolution is 600 dpi versus 300 dpi on your laser printer. With this capability, you have high quality graphics in terms of your publications, and the results match up with sending it to a print shop. Right now, what I would like to do is address what is happening over the next four months or so. We are going ahead and putting what I refer to as network DocuTech at the Center for AeroSpace Information. We hope to go ahead and have the system installed by late October. (*Editor's note: The DocuTech was installed and demonstrated as operational on October 29, 1993.*)

Network DocuTech at CASI

The network DocuTech for CASI consists of several components. The network publishing system prints at 135 copies per minute and has a bypass transport to enable it to connect to the signature booklet maker. The system also contains a cover insertion module which enables the booklet maker to provide for 17 x 11 saddle stitch documents or 8.5 x 11 saddle stitch documents with a hard cover. Additional components of the system are extended storage, print server, scanning station, and network server. The primary purpose of the extended storage is to provide the capacity to store the rasterized file that the DocuTech generates so that we can retrieve for future print-on-demand requirements. The extended storage capability will be used extensively by the Technology Transfer Office in response to user requests for their TSP's. The Evaluation Report will cover all the specifics concerning the use of the extended storage, costing algorithms in determining the cost per copy to the user and any other issues which might arise. As stated earlier, the report is due for release in March 1994.

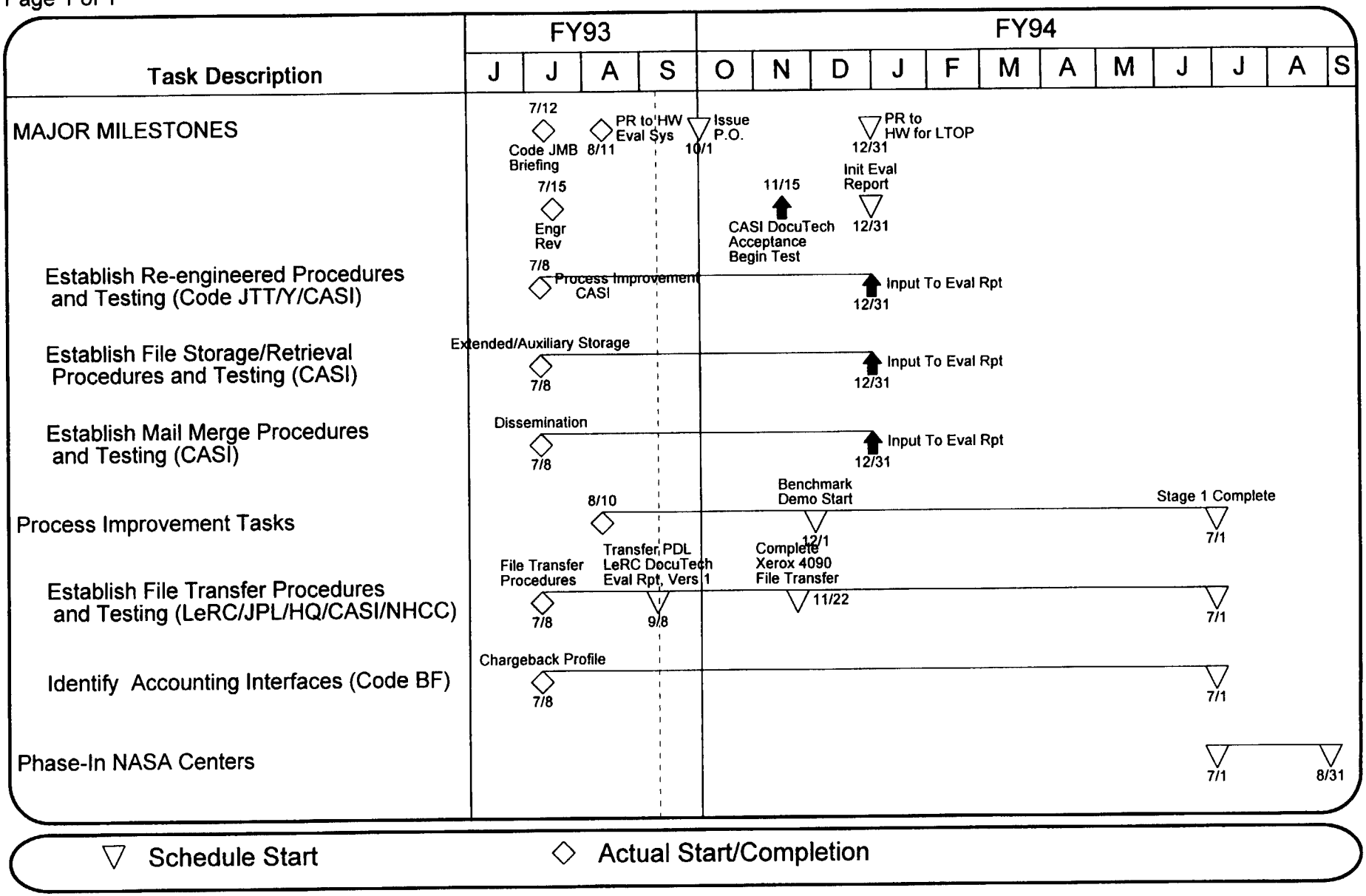
NASA-WIDE ELECTRONIC PUBLISHING SYSTEM

Update

September 9, 1993

NASA - WIDE ELECTRONIC PUBLISHING SYSTEM
EVALUATION AND IMPLEMENTATION TASKS
HQ/CASI/LERC/JPL

9/8/1993

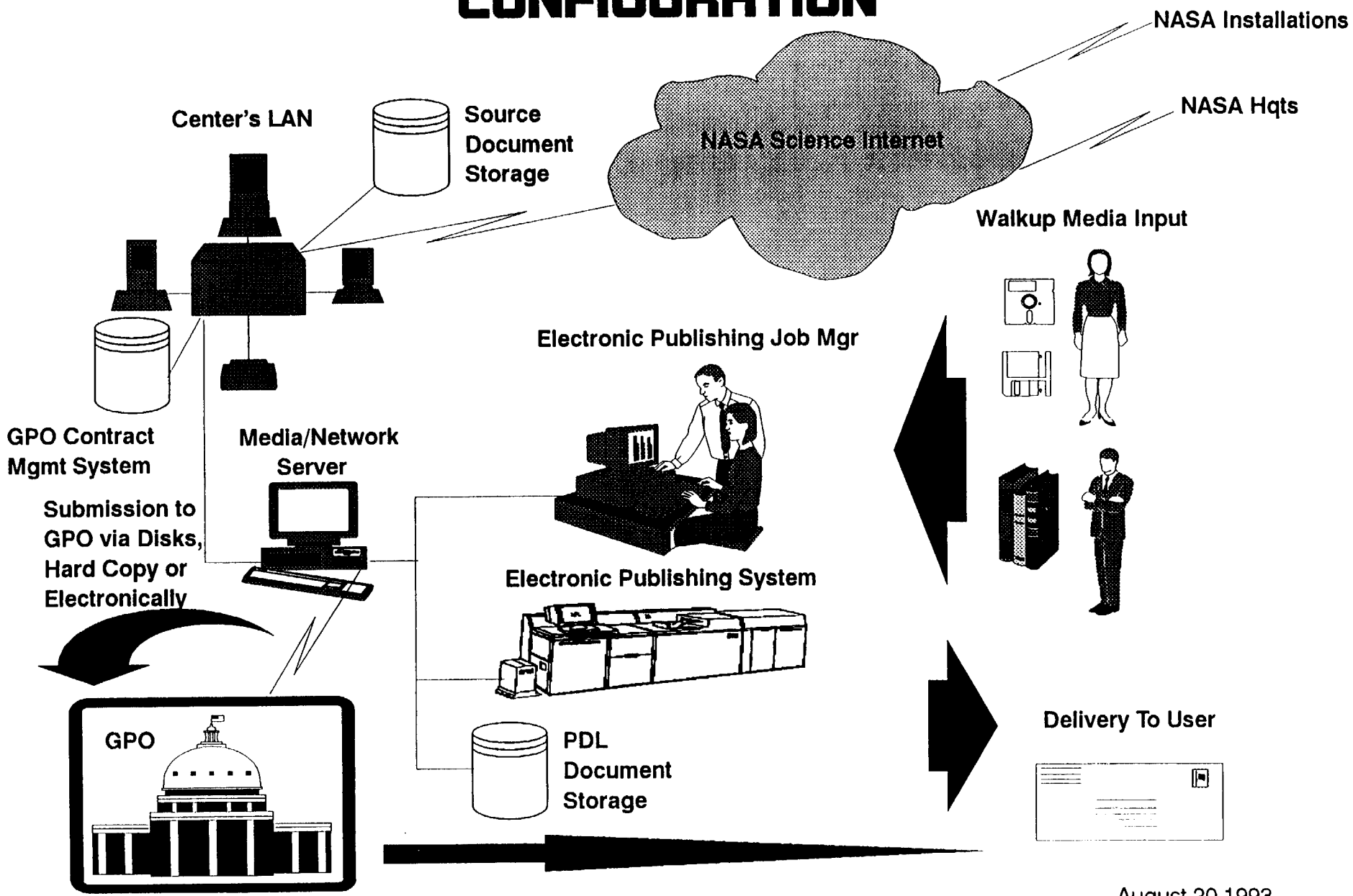


NASA-WIDE ELECTRONIC PUBLISHING SYSTEM

CURRENT EVALUATION GOALS

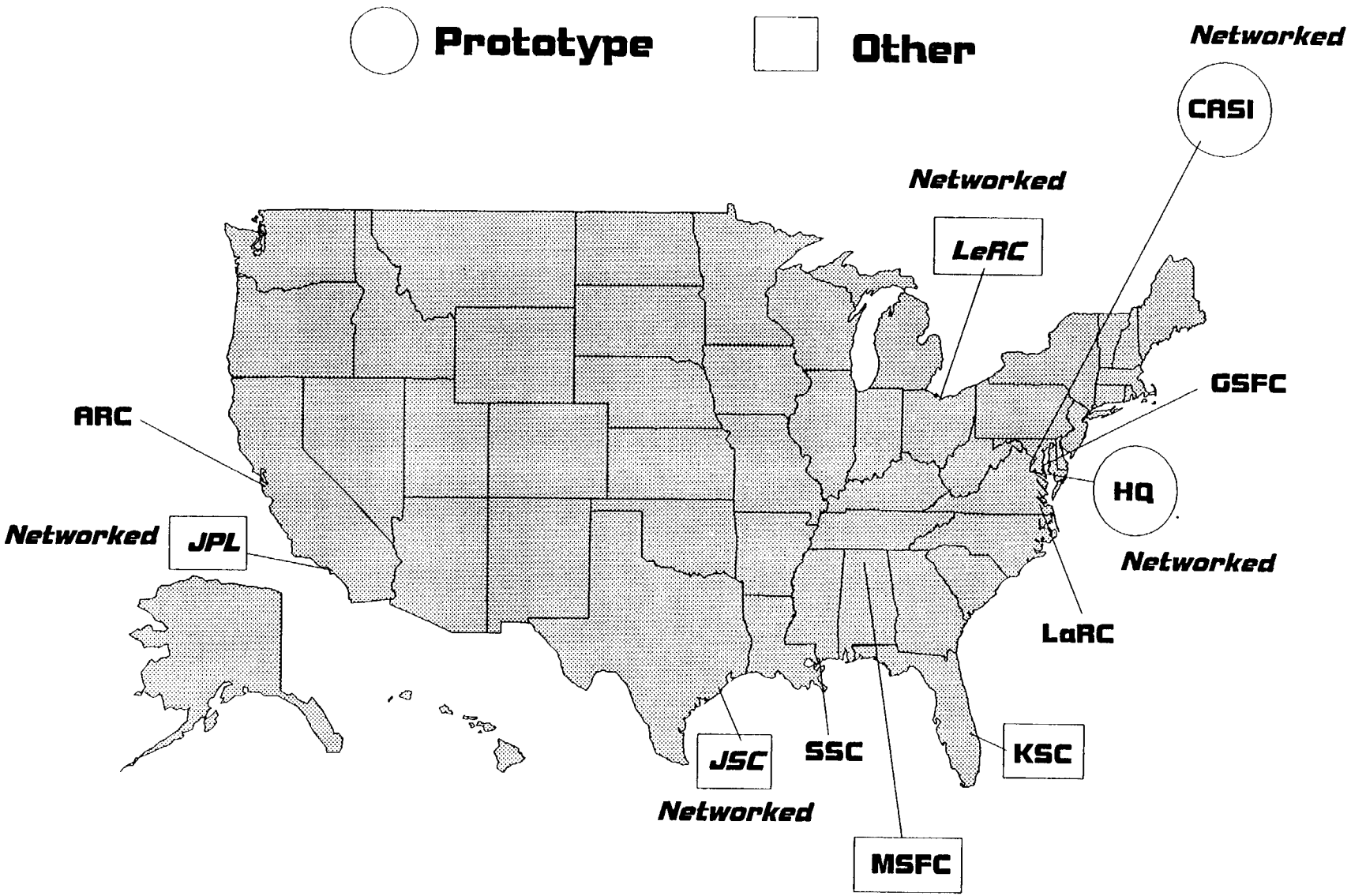
- EVALUATION OF:
  - NETWORKED DOCUTECH - CASI
  - PRINT-ON-DEMAND - (ON-SITE/REMOTE)
  - ACCOUNTING (CHARGE-BACK)
  - MAIL-MERGE
  - BAR CODE (USPS)
  - EXTENDED STORAGE (DOCUMENT RETRIEVAL)
  - PDL FILE TRANSFER FROM LERC/JPL/HQ (HTMs/LTMs/REPORTS)
  - DOCUMENT MANAGEMENT (E.G., Documetrix 1200 Functionality)
- DOCUMENTATION OF:
  - COST BENEFITS (COST PER COPY COMPARISONS)
  - PRODUCTIVITY INCREASES THRU REENGINEERING (TSPs/INDEXES/BIBLIOGRAPHIES/LISTINGS)
  - RESPONSE TIME TO USER COMMUNITY
  - IMPROVED PRODUCT QUALITY

NASA - ELECTRONIC PUBLISHING SYSTEM CONFIGURATION



August 20,1993

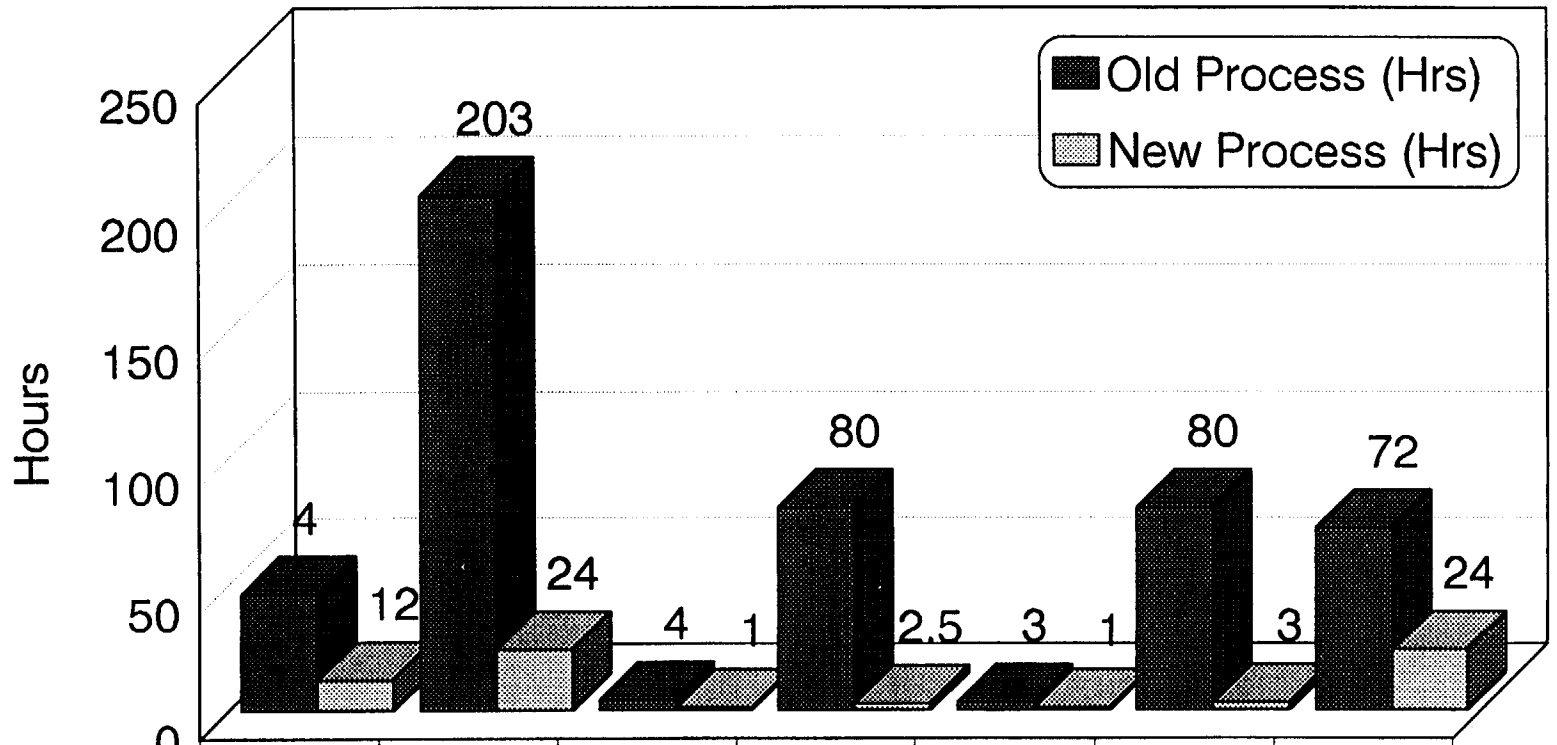
DOCUTECH LOCATIONS



August 20,1993

NASA - ELECTRONIC PUBLISHING SYSTEM

Productivity Improvements



	JSC (1)	KSC (2)	KSC (3)	MSFC (4)	JPL (5)	LeRC (6)	CASI (7)
Old Process (Hrs)	46	203	4	80	3	80	72
New Process (Hrs)	12	24	1	2.5	1	3	24
Productivity Gain (%)	74	88	74	90	33	90	34

Applications

Applications:

1. NASA Blue Book
2. Facilities Control Book
3. Analysis Reports
4. SSL Preprints
5. GPO Contract Mgmt System
6. Visitor Center Program Brochure
7. Tech Support Packages

August 20,1993

Electronic Document Interchange

Jim Erwin

User Demand for Electronic Documents

I am going to take a different tact. This is a project that we know we have to do, and it's scheduled in the modernization plan for FY94. What I am going to present today are really some of the issues that are associated with this project because, with the RECON retrieval effort, there's a lot of issues that need to be addressed, and it's a whole lot bigger effort than it might look to be on the surface. At NASA, as well as at DTIC and all of the other STI providers, our users are calling for electronic documents. We get an E-MAIL message: "Can you provide a whole list of documents in electronic format?" I have to reply, "No. We can provide the bibliographic information, but we can't provide the documents." In terms of cost, we haven't done a strict cost analysis, but I think there are some cost savings that could be involved in going to the electronic document approach.

Need for a Concept of Operations

However, I feel that we need a concept of operations. We have to understand what kinds of services we are going to provide and how we are going to be able to get all these various documents into the system. We get some documents from DTIC on microfiche; we get documents from NASA; we get them in electronic form, but we can also get them in hardcopy. We receive documents from ESA, from Israel. So, we have a lot of different document providers who provide these documents in a lot of different formats, and we aren't necessarily going to have total control over them. So, we have to look at how the documents are going to come in and what services and capabilities we want to provide on the output side. Do we want to merely provide the documents in electronic form, or provide a print version of the document like Dick is going to be able to do with DocuTech? Or do we want to be able use the documents to provide bibliographic information? Do we want to bring the documents saved in a tagged format and use those to create the surrogate records? Or do we just want to take the information in its entirety and put it up on the machine and let the people be go through it in a full-text mode and actually be able to look at the document online?

Format Issues

And again, because of the different formats that we receive, if we decided to go with, say, a full-text approach, then when we get documents in what I will call *analog format*, either hardcopy or in microfiche, do we leave them in that format, or do we want a hybrid system in which we want only some of the documents to be available in a full-text mode? Or do we

convert those documents that are coming in microfiche or hardcopy to the full-text version? To me, these are the questions. In a way, that is why I hesitated to give this presentation, because all of the other ones said, "This is all the good stuff that we are doing." Here, at least, I am saying that I have a lot of questions. But again, it's something we need to do.

Operational Impact

When we talk about the concept of operations, I need to reiterate the operational impact. Depending on what we are going to do, it is going to impact the storage format or multiple storage formats. Are we going to image full-text? Are we going to allow some PDL? Are we going to SGML files? Are we going to have some other kind tagged format? Are we going to have all of those? In terms of cataloging, are we going to continue to have a surrogate record? A bibliographic record? Or are we going to the full-text? Or are we going to go to a hybrid system where we have some full-text, some surrogate records? How are we going to handle the Mac file? If it's strictly a kind of a demand printing, electronic document exchange, then we can just say the Mac files are no problem. We will just convert the documents as they are ordered. However, if we go to a full-text format, we may have to do that document conversion if we want to have a hybrid system. Finally, in terms of distribution, are we going to eliminate the initial distribution? Are we going to continue with the initial distribution? How is it going to impact secondary distribution? Do we really want to go with printing an electronic delivery? Or do we want to provide online document viewing? Off the top of my head, I came up with three possible alternatives: demand distribution, full-text retrieval and kind of an SGML tagged format.

Demand Distribution

Very quickly, if we were going to go with what I am calling *demand distribution*, we have a storage format of image and multiple PDL, and the main idea would be the electronic distribution of the documents for a quick secondary distribution in terms of printing. In terms of the cataloging, we go with the surrogate record; everything would look pretty much like it does now and it would fit with a RECON-like system. Handling the backfile would be on-demand. So, we could take our microfiche or hardcopy and, at order time, scan it in. Once it had been ordered, we could put it up on an optical disk or our mass storage, and then we would have it for the next person who wanted to order it.

Full-Text Retrieval

If we go to the full-text retrieval, then our storage formats again are probably multiple. We have a full-text. You could have multiple PDL postscript - the HP type of format. In terms of the cataloging, we may have this type of hybrid system. We may have our old documents - surrogate records prior to 1994. The documents in the future would be in a full-text. Or it might just be the NASA documents that are in full-text, and the DTIC documents would be in

the surrogate format, or the DTIC documents in the full-text and the NASA documents would be in the surrogate, depending on what moves quicker. Would we handle the backfile on-demand, or would we have to do a NASA conversion? That would depend on whether we wanted this hybrid system approach. In terms of distribution, we could use them for electronic delivery, and we would also have the capability of the online browsing - the SGML approach. I used this to stand for a tagged format. That would allow us to do what full-text did, but in addition we would be able to bring the documents in and actually process them to a great extent unattended (human unattended), pull the bibliographic information, and create index terms based on what we knew were the fields in that document. So, that would provide an additional capability. But, then again, we would probably have to deal with multiple format and multiple processing screens. The SGML would be handled one way, microfiche and hardcopy full-text handled another way.

Services and Products to Guide System Configuration

In conclusion, I think we, as an organization, as a program, have to decide what services and products we want to provide up front, and that's going to determine the configuration. Conceivably, this is all because I am confused, but it does seem to be an issue. I feel that there is a lot of confusion in this area. There are people who say, "We are going to scan the documents" or "We are going to go full-text." That doesn't necessarily follow from the idea of scanning the documents in. So, you really have to look at cost, tradeoff, what kind of system you want to have, and what kind of control problems you may have for all of these scenarios. Now that I have covered the negative issues, we will go back to the positive accomplishments.

N94-36850

NASA Access Mechanism (NAM)

Judy Hunter

Background

In her presentation this morning, Karen Kaye described NASA's vision for the future of information management. Part of this vision involves providing the NASA user community with a set of tools to assist in identifying sources of information and to navigate the networking infrastructure to connect to the sources in order to extract the relevant information (Viewgraph 1). The project was initiated in March 1990 to demonstrate the concept of using a Graphical User Interface (GUI) and Intelligent Gateway Processor (IGP) to provide the users with the semblance of a one stop shopping environment for information management.

User Requirements Study

A user requirements study was conducted at five of the NASA Centers from which it was determined that the NASA users want 1) access to diverse sources of information; 2) an intuitive approach to using the system in order to decrease the learning curve; 3) to avoid the requirements to learn the system query languages; 4) access to peers and other informal sources of information; and 5) simplified and enhanced presentation of search results. This study was completed in May 1991.

Intelligent Gateway Processor (IGP)

At the same time the user requirements were being evaluated, past applications of the IGP technology and the networking infrastructure at NASA were being evaluated. The user requirements and the IGP and networking studies were used to complete the initial NAM design in November 1991. Computer programming began in December 1991. Four months later, the alpha version was demonstrated at the annual STI Conference in April 1992. The beta version was completed in December 1992 and was deployed to 60 user desktops for six months of user testing. The testing period formally ended on May 31, 1993.

Lessons Learned Document

A Lessons Learned document is being prepared that will include everything NASA learned from the prototype from the user perspective and from the technical perspective. This document will include recommendations about what applications the NASA STI Program may see for this technology in the future. It will be presented to the internal NASA STI Engineering Review Board in October. When I show you the NAM screen, I'll get into that a little further.

Peer Locator Services

Peer locator services: a totally unrelated survey. We had to go out and survey people who do not use our system. How do you figure out who the people are if they don't use our system? Well, we went out to the NASA Centers; they had all of their phonebooks in digital format, different formats, but digital format. We loaded them onto a database and made them accessible through NAM. We used Finger, which is an Internet utility which goes out and looks at all the UNIX boxes if they are marked for this on any of those systems that will give you information: name, address, phone number, and Internet address. We needed to provide some sort of e-mail. We happened to use e-mail. Under Miscellaneous Utilities: a lot of our scientists already are on systems where they can download information that is not bibliographic and use some model. Our intention here was to show a graphics capability. We were not actually doing modeling at this point, but we wanted to show that you could bring it in and do some graphics and be able manipulate data later.

NASA Phonebook

Question: A quick question on the NASA phonebook. You put them into NAM? Are they updated every six months?

Answer: For the prototype, we do not update them (Viewgraph 2). In the Lessons Learned document, we'll figure out what we want to do about it. There are some things in the prototype that would lead you to really step back and say, "Okay, we loaded 54,000 NASA scientists and engineers in digital format. Over time, if we decide to go up in operation, do we really want to be in a position where we have to update these phonebooks?" That's a good question that I don't have the answer for at this second.

NAM Menu

The next view slides are actual copies of NAM screens (Viewgraphs 3-8). I have four people and they all have terminals in their offices and are all ready and able to show you NAM if you would actually like to see it sometime before you leave today. The main NAM menu shows a basic functionality, helps them find sources of information that are available, gives them E-MAIL capabilities, helps them locate/communicate with their peers, locate others utilities that are available, and things like that. In this particular thing we are showing graphics capability. It's a point and click and Window based. I stripped all of the technical stuff about NAM out to make this a high level discussion, so if you have technical questions, just ask. If you know the source, you just point and click; if you know the source of the information, then you can select it and it will tell you what file collections are available. For some reason, I like to use RECON. So, if you know the file collection, you select it; at that point NAM goes out automatically and connects to it. The user is just sitting there and it

comes back with the screen. You have three choices: novice, intermediate, or expert and you choose that. This is a user configuration that you set up beforehand. In the intermediate screen, my illustration is just to show you that the user can just go in the boxes, and in this example, fill in author, title, some keywords. He hits search and goes out and makes the connection to RECON. Actually, I switched to STN in this example. It gives you a list of citations; you tell it how many you want it to display. Each of these is a button. You select it and it displays the full citation. The presentation here is a little different than if you actually dialed into RECON; we tried to make this simpler and easier to read.

Question: You get the same presentation no matter the source?

Answer: Correct. If you would like to save what you downloaded to a file you can do that. If you want to e-mail it to someone or to yourself, you can do that. If you would like to order the full document online, you can do that.

Database Management Systems

Question: How many different database management systems are we sitting on right now? What is the user transferring to?

Answer: Right now, he translates it to the post computer's query language: RECON and STN for the prototype. That's another question we are addressing right now in writing the Lessons Learned document. It's one thing to have the ability to hide the query languages from the users; however, if you decide to go into operation and a user has five databases that he searches regularly, do you really want to be in a business of keeping those? Every time a host system changes a query you would have to go back and change the system. How do you really handle that? Do you do this kind of bridge translation until everybody is SQL or something? I don't know.

Question: Are you still displaying the translating query?

Answer: The user can decide. It will show you what is actually getting sent to the system and the theory here is, if you use it enough, you can actually use it to learn the query language from a part of your system.

E-Mail and the Peer Locator

E-mail: that we are using right now (Viewgraphs 9 and 10). The nice thing is, if you want to send your downloaded search to yourself, and you hit the mail button, it automatically pops up on an e-mail window for you. How do I know who to send it to? Well, we handle that problem also with our peer locator. The last time I looked at the NASA phonebook, it was something like 54,000 names just from the NASA phonebook. We added the last year as we

were going along. NASA has their own implementation of X.500. Doesn't everyone? So, now we are making an X.500 available. The idea here is that you are sitting at your desk; you are working on a project and you say, "I need to find Cumber or maybe it's Cumbly or somebody." Maybe he is at Johnson or he has been doing work in this area; you knew he was a NASA person. You might select the NASA phonebook; put his name in. It will give you an index of everything that comes before or after. If you don't exactly know his last name, it will help you to determine which one is the right one. If you click on it, it will go out and search the digital phonebook and give you whatever information is available. The Center's phonebooks have different information sometimes, but it will have his name, address, phone number, and, if available, his electronic e-mail. At that point, you can go immediately into e-mail and send that person a message.

NAM, Front End to Internet

Amazingly enough, NAM has been written up in a couple of magazines, *Government Computer News*, and, most recently, *Computer World.* And amazingly enough, this is what most users find to be one of the most exciting things about NAM. They see this as a front end to the Internet. It has a point and click access, things such as Usenet News (that's the read news on Internet), WAIS, wide information area servers (Viewgraph 11). You can go for things, go out there, and search servers that are available on the Internet. You can come back with all kinds of information. This has been the thing that a lot of people are very excited about. The Internet's out there, and there's a lot of information on the Internet.

Question: Are those servers running on the workstation or on the server?

Answer: They are working on our Sun Server.

Graphics

For the graphics: just to show the ability to use graphics, we loaded up a weather map. I think it runs on the University of Michigan's servers (Viewgraphs 12 and 13). It actually shows the weather changing. You can come in, figure our what modeling packages you might need to use if you are downloading a lot of data, maybe at Goddard, and load that immediately into a modeling package of some sort. The future of NAM, at this point, is a little undetermined. We are in the process of finalizing the Lessons Learned document.

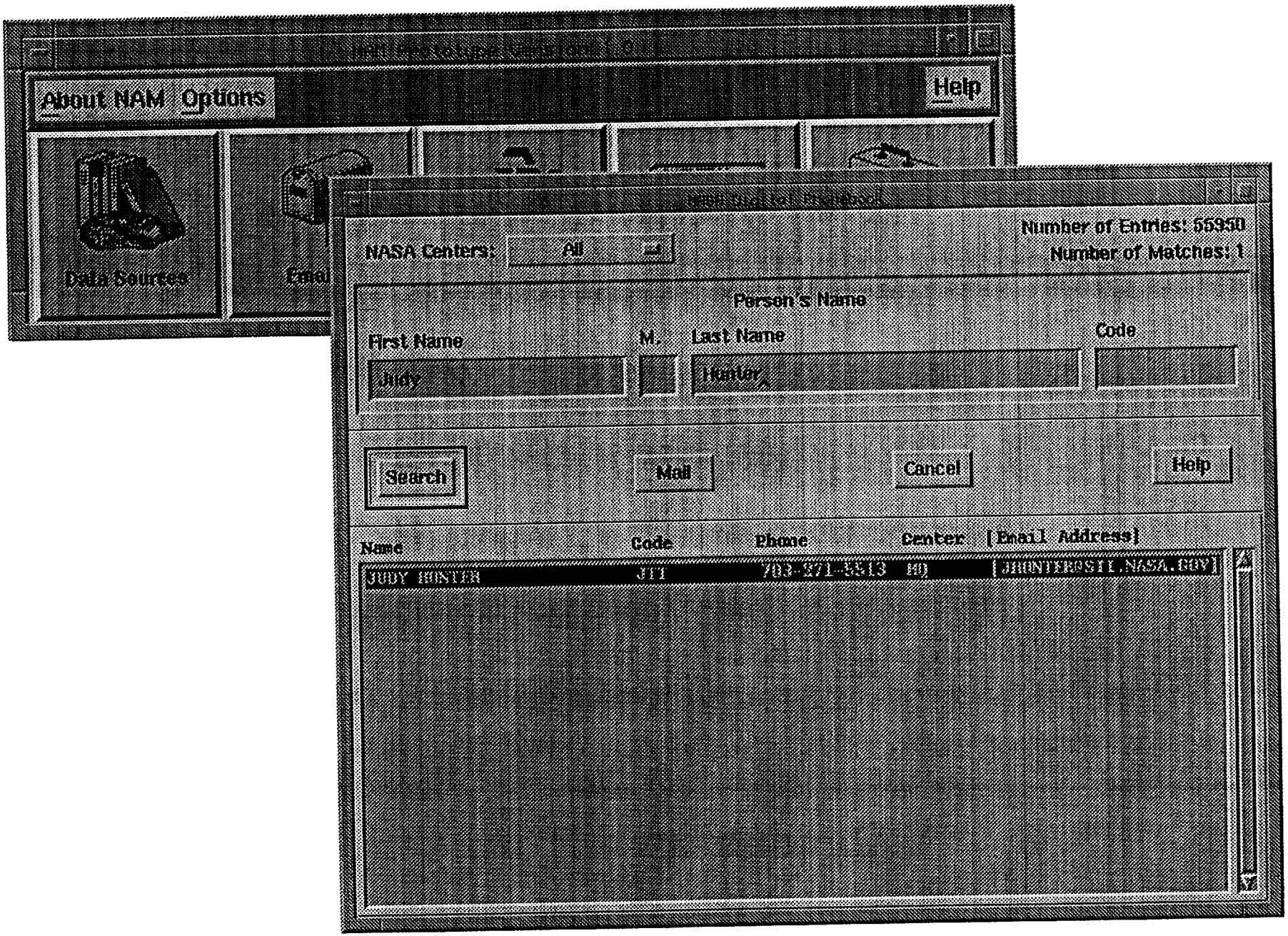
NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

Chronology

- Project Initiated in March 1990
- User Requirements Study Completed May 1991
- High Level Design Completed November 1991
- Coding Began in December 1991
- Alpha Version Demonstrated in April 1992
- Beta Version Deployed in December 1992
- Prototype Testing Completed in May 1993
- Lessons Learned Document Completed in October 1993

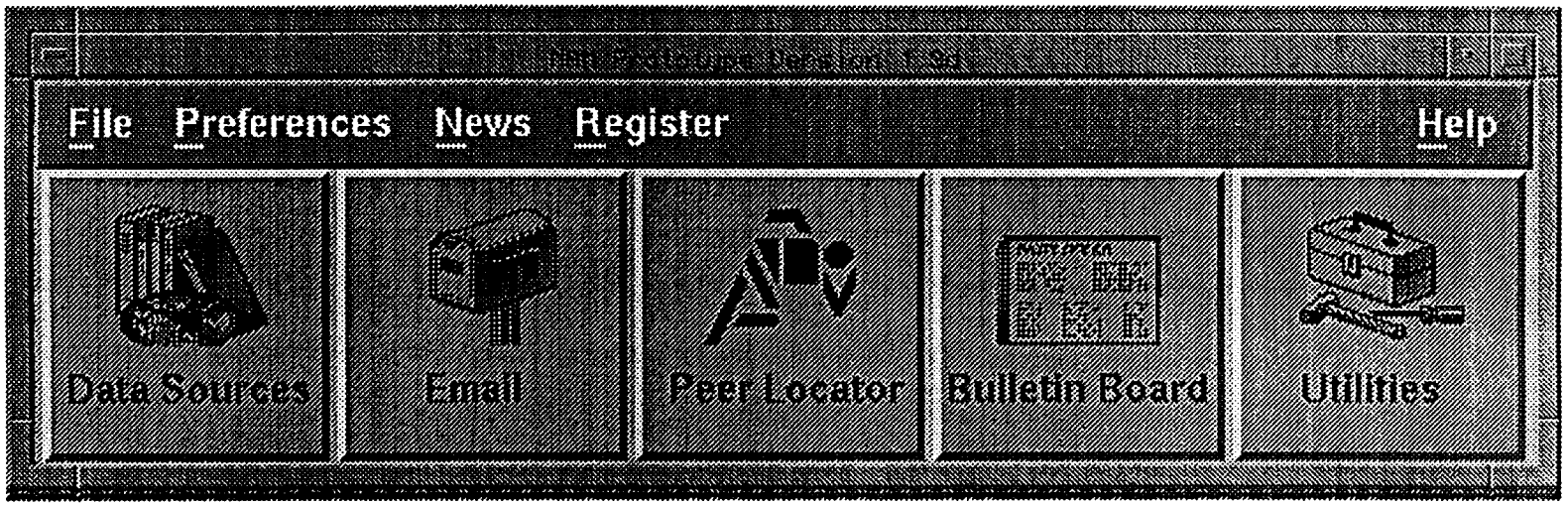
97

JH 1313 - 10/04/93


About NAM Options
Help
Data Sources
NASA Centers:
All
Number of Entries: 66950
Number of Matches: 1
Person's Name
First Name
M.
Last Name
Code
Judy
Hunter
Search
Mail
Cancel
Help
Name
Code
Phone
Center
[Email Address]
JUDY HUNTER
JJ1
703-621-5515
HQ
JHUNTER@SIL.NASA.GOV

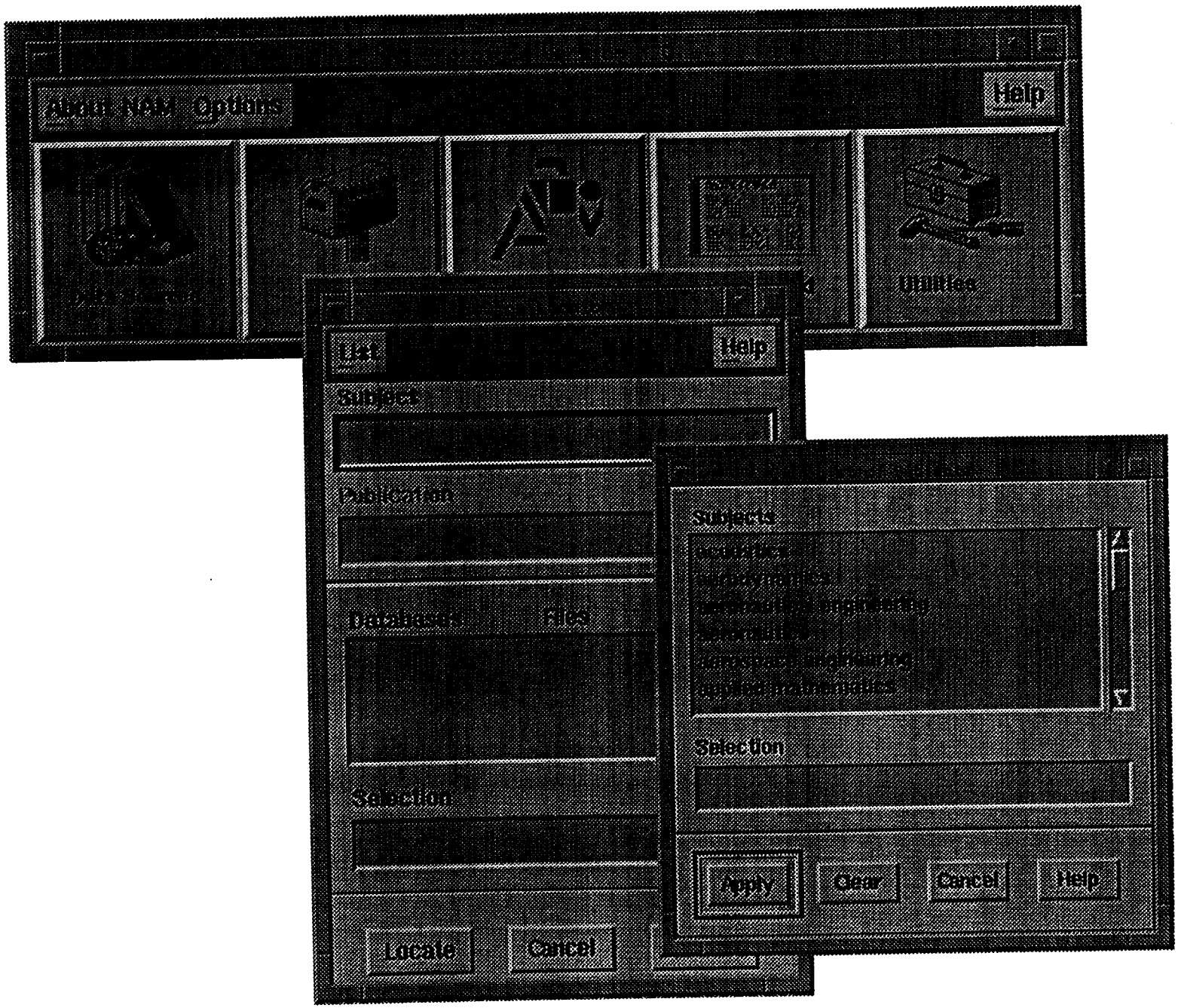
NASA STI PROGRAM SCIENTIFIC & TECHNICAL INFORMATION

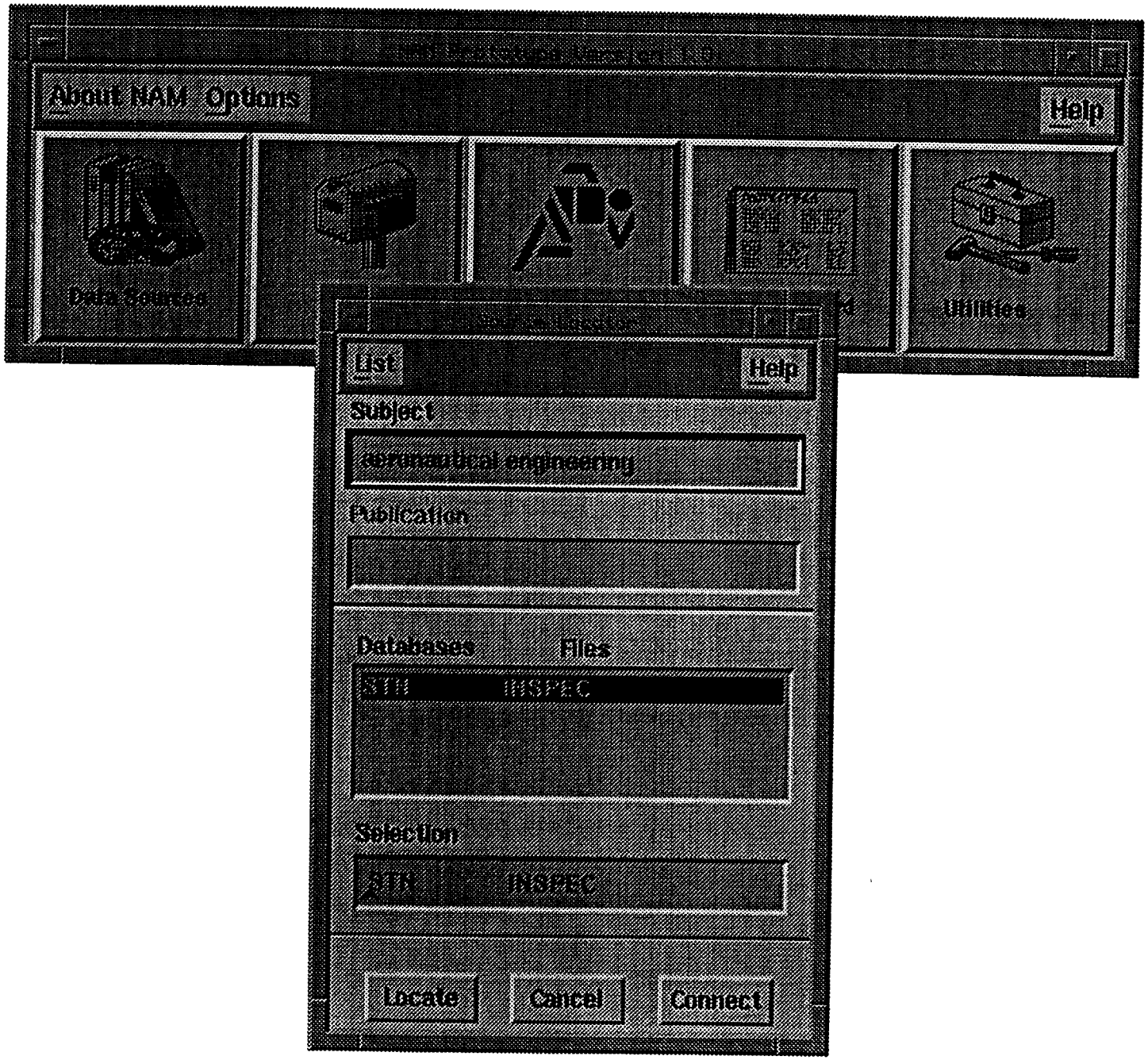
One Stop Shopping



66

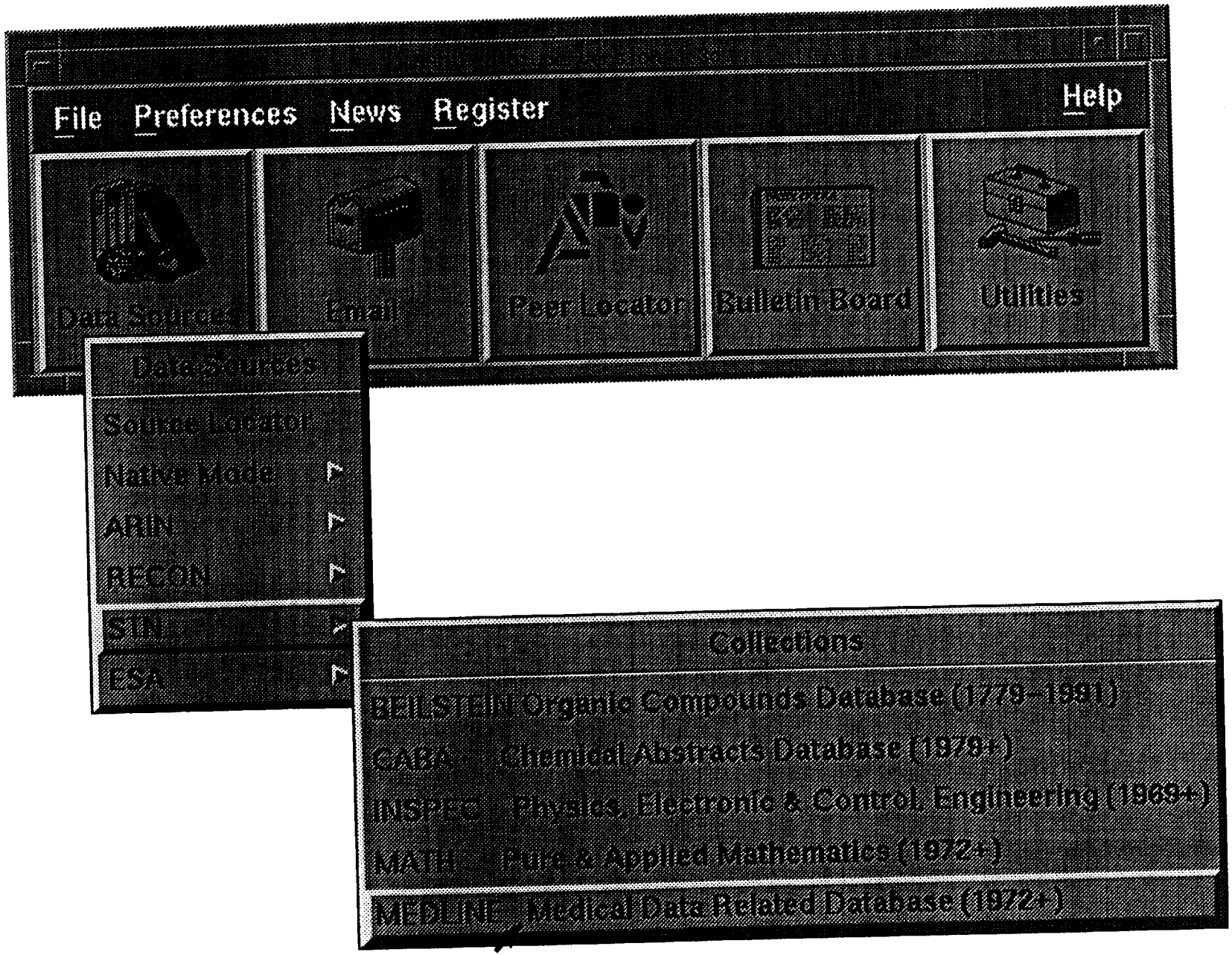
JH 1312 - 10/04/93

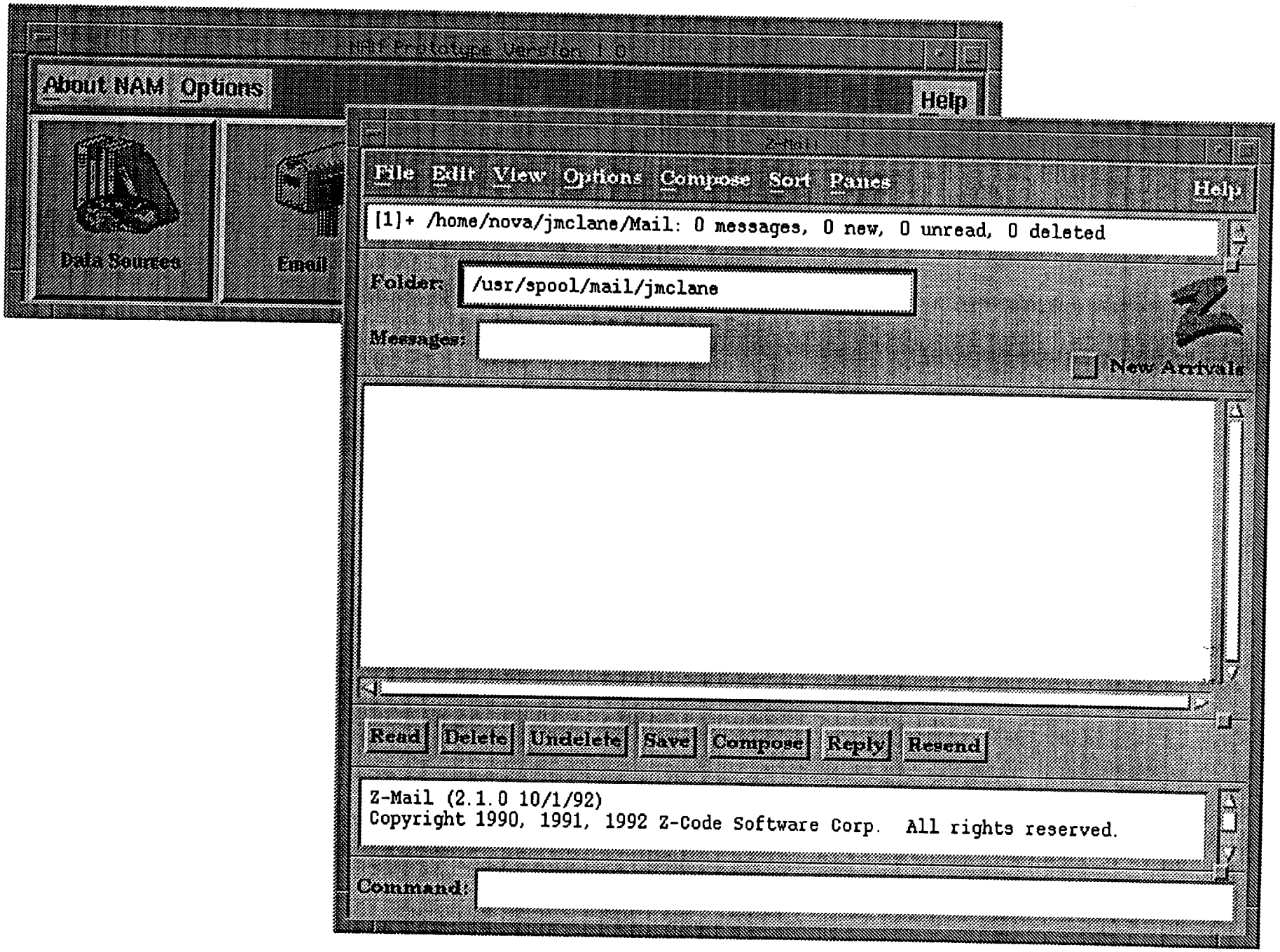

Help
Utilities
Help
Subject
Publication
Databases
Files
Selection
Locate
Cancel
Subjects
Selection
Apply
Clear
Cancel
Help



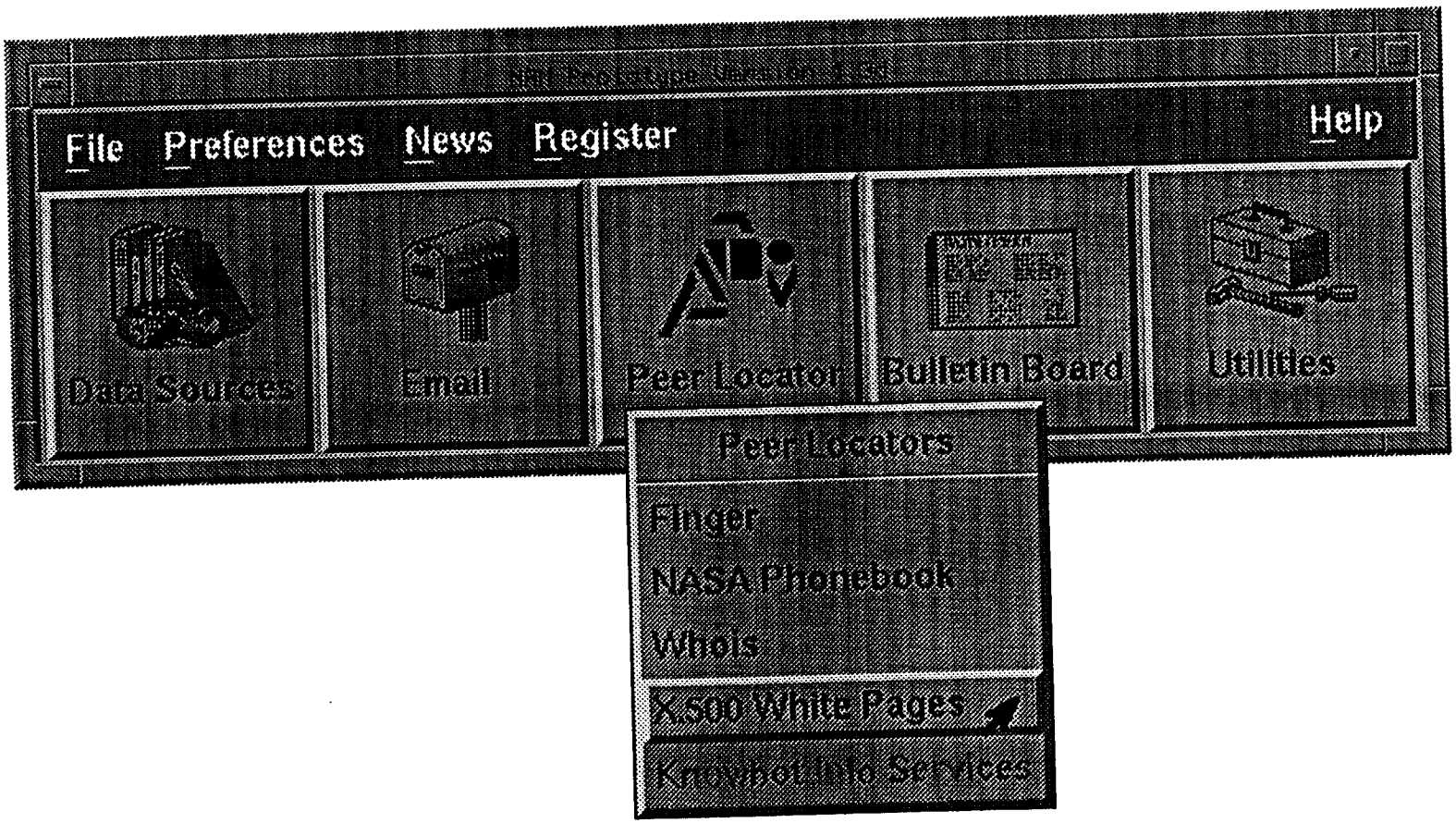
101

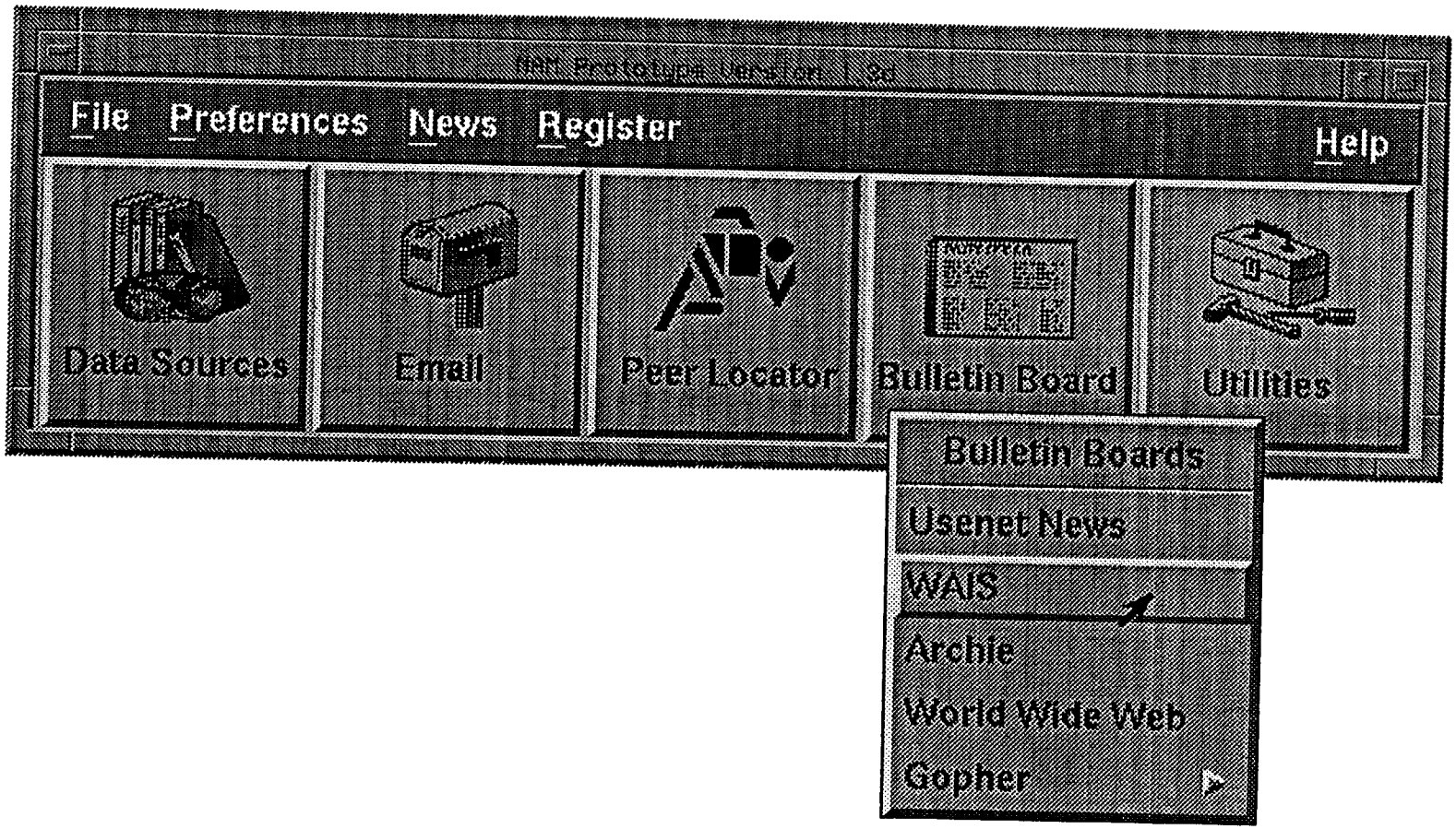
JH-1315 – 10/13/93


File Preferences News Register Help
Data Sources
Email
Peer Locator
Bulletin Board
Utilities
Data Sources
Source Locator
Native Mode
ARIN
RECON
STN
ESA
Collections
BEILSTEIN Organic Compounds Database (1779-1991)
CABA Chemical Abstracts Database (1979+)
INSPEC Physics, Electronic & Control, Engineering (1969+)
MATH Pure & Applied Mathematics (1972+)
MEDLINE Medical Data Related Database (1972+)



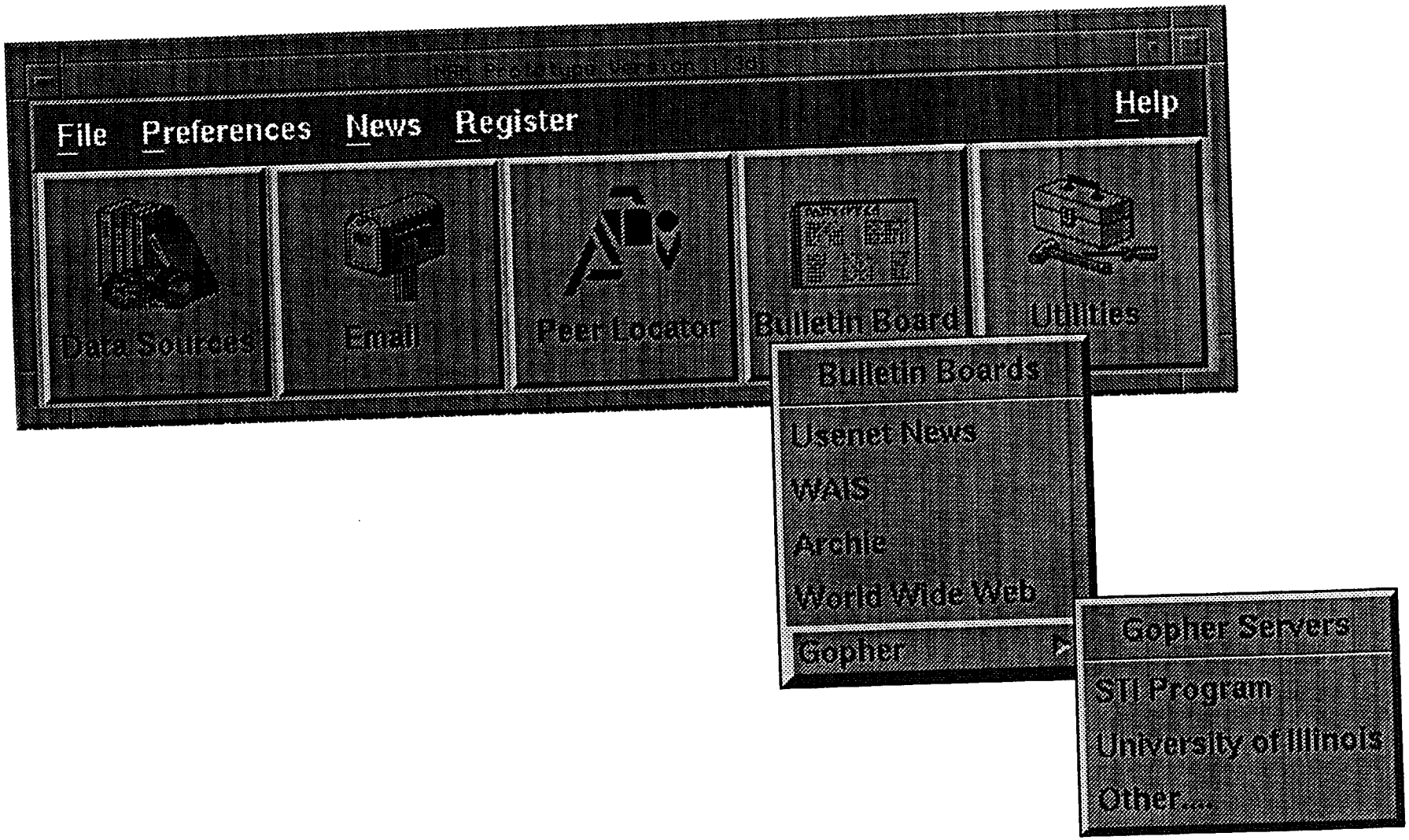
JH-1320 – 10/13/93


File
Preferences
News
Register
Help
Data Sources
Email
Peer Locator
Bulletin Board
Utilities
Peer Locators
Finger
NASA Phonebook
Whois
X.500 White Pages
Knowbot Info Services



105

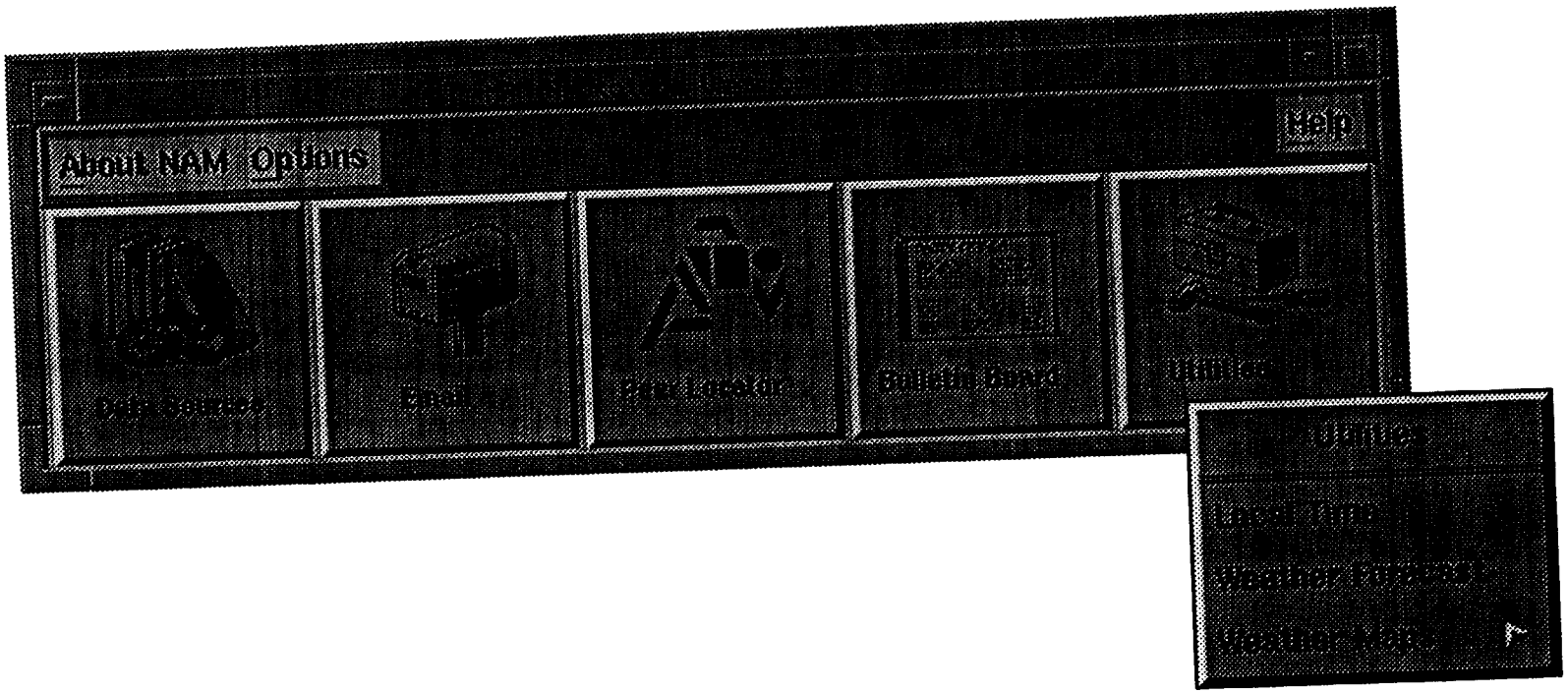
JH-1325 – 10/13/93



JH-1323 – 10/13/93

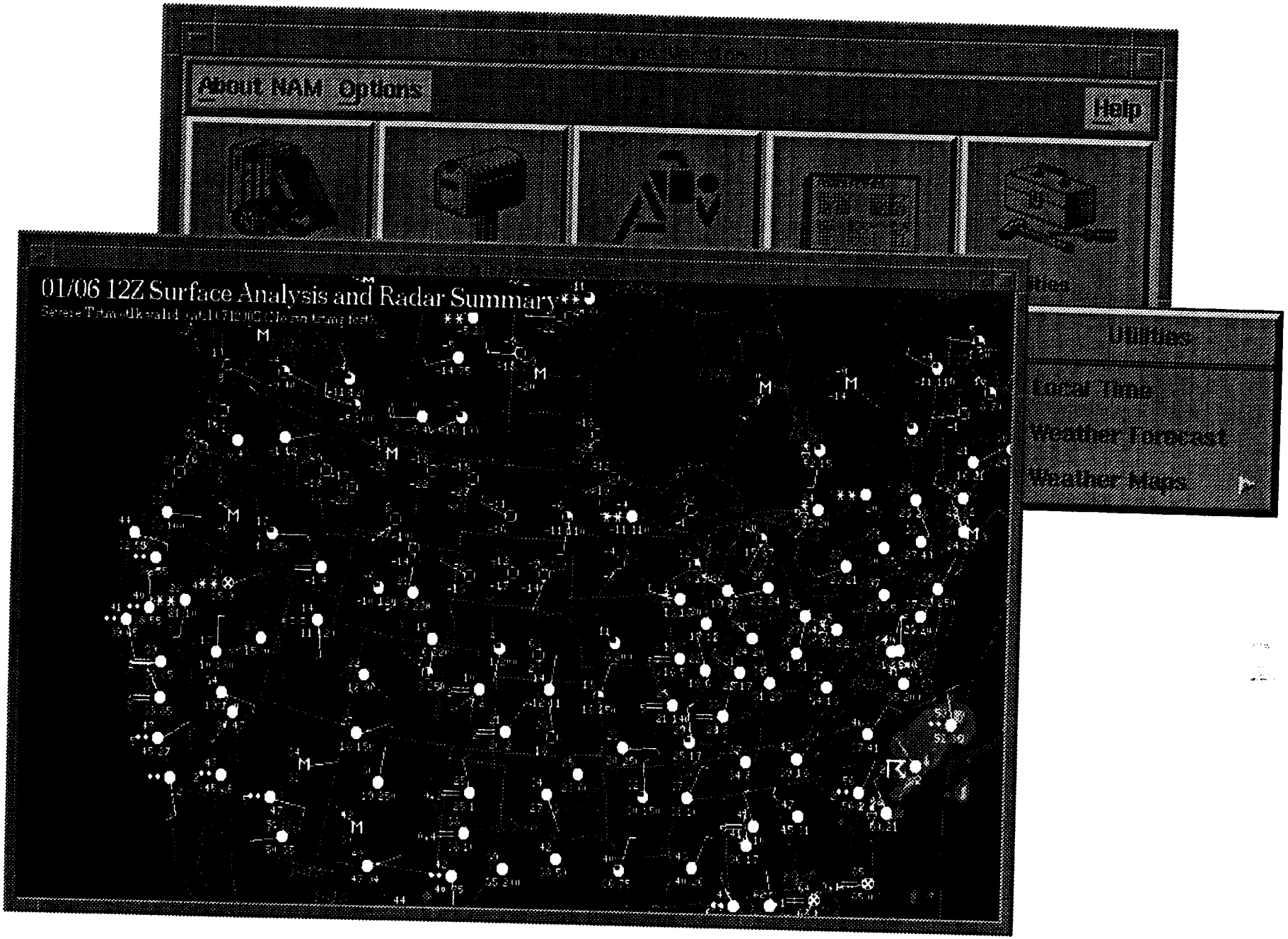
WAIS

- Wide Area Information Server
- Based on Z39.50 application layer protocol
- Locate and retrieve information using English like queries and relevance feedback mechanism
- Supports text, image, voice, etc . . .
- Public Domain Graphical User Interface clients available for PC, Mac, and others



108

JH-1331 – 10/13/93



109

JH-1332 – 10/13/93

Network Upgrades

Roland Ridgeway

CASI Mainframes

Network Upgrades, as part of the modernization plan, cover equipment and software for both of the STI Program Local Area Networks (LAN) (Viewgraph 1). Both the CASI and CG4 LANs are part of the plan. The client/server environment requirement is also part of the modernization; it is a way of rightsizing our information systems, of taking advantage of the open systems, improving them, and providing more flexibility. A lot of our current environment in our production operation at CASI by BWI Airport is mainframe based. There are two IBM-4381s at CASI. They are used to gather information, acquire it, prepare it for the databases and load it onto the databases, utilizing the mainframes' terminal connections. These computers don't give you a lot of flexibility for some of the new things we want to do, such as the easy electronic transfer of data, or some of the imaging we want to be able to do. So, we need to take the current mainframes that we have and move them, upgrade them, make them more available to our support staff so that they can provide more services to the customers. That's what this particular line item is in the modernization plan.

CG4 LAN

The CG4 LAN was put in as the program was moving ahead (Viewgraph 2). Gladys came on board and brought us additional management staff that had some ideas for modernizing. They put in a better LAN than was in CG4. They provided a basic office automation functionality that wasn't here at that time. The mainframe and computer network system accesses were provided, as well as anonymous Gopher access and the SCAN product file that was mentioned earlier. The LAN was brought in and was a pretty good start. Most of the staff had good equipment but needed some additional upgrading of the memory and storage capabilities, additional boards for graphics or other multimedia-type requirements. Some of the money for those items will come out of this modernization line item, but most of the money will be going to the CASI LAN upgrade requirements.

Office Automation Upgrade

For a number of years, the CASI staff were utilizing terminals and the mainframes for their work (Viewgraph 3). We started putting a LAN in about two years ago; we provided some initial office automation capability and access to the mainframe. We received some money last year and were able to add to what we've started. So, it's about half complete. Many CASI staff still operate without any desktop equipment or with PCs that are outdated and underpowered. This modernization money will enable us to complete the LAN and provide additional capabilities. We are going to replace the terminals that are out there and modernize the systems. We are going to migrate from a mainframe terminal environment to the

client/server LAN environment. We are going to add additional PCs; we are bringing in 486s with 16 megabyte RAM. We have different PC configuration requirements at CASI because of the different functionalities of the staff - the database development group, the database processing personnel, and publications group. The LAN is being developed with this upgrade to the point that all the CASI staff that needs a PC or MAC to perform their assignments will have a machine on their desk or have access to a machine near their work area. Network laser printers are being installed that will be shared by the CASI staff so that everyone will have access to a quality hardcopy output device. Additional disk storage devices and services are also being purchased and installed as part of this upgrade to support client/server functions.

Flexibility in Customer Services

This equipment is part of the process to establish a redefined platform and environment which will allow us to be more flexible in supporting the STI Program's customers. So, with these additional PCs, we are going to provide more basic office automation functionalities, mainframe and computer network systems accesses, and client/server functions. The LAN upgrades will provide the staff the capability to utilize the LAN to access some of the other information systems that are around so that we can provide the information to our customers when responding to search requests. We will have the capability of doing anonymous FTP and providing Gopher access file creation. The two technical staffs that are supporting the CG4 gateway and CASI gateway are beginning to work together more closely and utilize each others' knowledge. We have done some prototyping in CG4 with the Gopher and the SCAN product. If it's appropriate, we can move SCAN in the future to the CASI operation were it will be fully supported by the Help Desk if that is needed, and it will meet the demands of our production environment. So, if we do a lot of prototyping at CG4, we can also do prototyping at CASI and move the products into the CASI production environment. The products would be supported by the Help Desk, for user information, for ordering documents, video tapes, or whatever. That's part of the CASI operations that we are running.

Personal Computers

Also, we are going to be able to provide additional client/server functions with the modernization money and establish a redefined platform environment. We are moving towards more open architecture. We are getting off the mainframes. The 4341 is a very old technology, which is not flexible. It's very expensive for the maintenance of the software and the equipment. As we move to the client/server with a more open architecture, we will have more flexibility and more power for the systems. That's another reason why we are doing a lot of this upgrading. We need to bring in printers, have network laser printers to share so everyone can have quality output. We are going to have PCs on everyone's desk or access in their work area. We are purchasing the equipment for the CASI upgrade in quantities that will allow us to implement them in an efficient manner. After the shipment is received, all the equipment is checked out, moved to the user's site, installed, software loaded, connected to the LAN, and the user walked through an introduction, if necessary. The CASI staff have been attending noon time training session on Windows, WordPerfect, E-Mail, Harvard

Graphics, etc. to prepare for the new environment. Because we only have two staff persons doing this installation, we are having the items delivered in groups and installing them as they are received. We are experiencing a good situation with the procurements because the prices of all this PC equipment keeps coming down, allowing us to buy really good equipment at very reasonable prices. This pricing will allow us to buy more items for the LAN, such as boards to provide FA capabilities to send and receive messages at the person's own workstations, than we may have been able to buy earlier.

Off-The-Shelf Software

We will be using commercial off-the-shelf software to meet some of the requirements to improve our services and customer support (Viewgraph 4). The RECON replacement system will be an off-the-shelf product. The ARIN System is supported by an off-the-shelf package called NOTIS. NOTIS is rewriting and improving their product, making it more platform independent. Almost 1,000 patrons have been registered now for access to the ARIN system from their workstations, PCs, or MACs. The NOTIS product was developed for mainframes, but NOTIS has been rewriting their product in the last couple of years, and they are moving towards the client server environment. We are now looking into some of their products to provide more user services such as document delivery. They redesigned or re-engineered their product to provide additional capabilities to take advantage of the servers and the client user interfaces which they can build to run under Windows and other software like that. They are now in the position of lining up customers to test the new software - beta testing. By the end of 1994, or the beginning of 1995, they are going to have their client/server product available. So, we are going to be in a position, I hope by that time, to move the ARIN system from the mainframe to the client/server environment.

RECON/STIMS Replacement

Our replacement for RECON/STIMS should be in the final phase of implementation on the in-house client/server environment. We are working under that assumption and moving to finalize the parallel testing environment and data conversion for RECON/STIMS. There will be a number of servers on this LAN that will have external connection to them for our users, so they do not go through the internal LAN. Our internal staff will have access to the LAN and servers for internal processing requirements such as dupe checking of documents and processing data for inclusion to the databases. The staff will be able to help the users in ordering information, providing information, responding to users' concerns and questions.

Password Requirements for Access

We are going to start looking at the ID and password requirements for access. As a government agency, we have to worry about IDs and passwords to provide access so we can isolate these products and servers in order to safeguard the data and the operation. We can open up this a little broader so that maybe you only need an ID for access. We are also looking at some of our inhouse products and services to improve access.

Rightsizing Project

As far as the rightsizing project is concerned, we are looking at how we are going to move off the mainframe and improve systems, along with taking advantage of the PCs, clients and servers. We have already made some decisions inside of CASI to use MS Windows as our main user interface. We are developing applications to run under Windows. In fact, we have already started by developing a Help Desk system that runs under MS Windows and was developed with PC software development tools using an object orientated programming approach. This will allow objects to be reused from a library for new development requirements and easier maintenance. The tools were used to develop screens and generate code so that we didn't have do everything from scratch. We bought products that will allow us to generate systems - again, commercial off-the-shelf products. We use the object library and see what's available to extend the existing system or to add additional capabilities.

User Interfaces

We are looking at developing user interfaces to integrate application software. For example, the software system we end up procuring to replace RECON/STIMS will probably require the development of interfaces; we don't know exactly what kind of user interface it's going to have. We hope it's going to have a very nice interface, but since we have to market to a customer base that has vastly different machines - PCs, with different capabilities and different levels - we are going to have to be able to develop user interfaces. There are some good user interfaces that are on PCs now, but to provide flexible interfaces to our user community for all our services, we are going to have to develop improved ones. The development tools mentioned before will be used along with off-the-shelf software packages, changing our development role to one of integrating.

DocuTech

We are moving towards a staff that can understand how these different software packages need to be put together and integrated to utilize some of the things that we were talking about earlier, for example, the DocuTech. If the prototype works out, the DocuTech will be hooked to the LAN at CASI so that our publications group, the people who develop our publications graphics capability, can send what they created electronically to the DocuTech. The images

are all saved and the information is created to utilize that capability. When we have the DocuTech installed, CASI will have a lot more capabilities to test and work with.

Redesign of Mainframe Applications

The mainframe applications will all be redesigned. We talked earlier about replacing the mainframes. What's really happening there is that, as we get the software in, we see what it will run on, what it needs to be efficient. That will tell us what kind of hardware we need to replace the mainframes with. They may not be what you think of as normal IBM mainframes or Amdahls or one of those kinds of machines; it will probably be a server with a lot of capability. We are redesigning our systems, and will be integrating the application software. Network upgrades will support the development in the use of the new retrieval system. The mainframe replacement will support an improved level of customer service, providing access to the servers. We will try to build some backup capabilities with equipment for this whole configuration. We will be able to move to different servers if we have to; we will be able to move the application to another server while we are taking care of a server that we are having problems with so that our systems are always up for our user community.

Online Systems Availability

We want to increase the time that our online systems are available to the community; we've done that with the mainframe in the last year or so. RECON and ARIN are available from 7 a.m. to 12 p.m. eastern time now. This has helped the West Coast out a lot; they would like to see 24-hour availability so they can come and go when they feel like it. Their researchers and scientists come in all hours of the day and night. That is what we are doing with the modernization access upgrade. We are not necessarily talking about bringing in new backbones and new bandwidth, but we could if that's a requirement. What we are looking to do is identify the requirements and move towards open systems, creating situations where we have products and services to help our user community. Our customer is the main concern, and we want to help that customer become better, quicker, and more efficient in his or her job as we improve our systems hardware and software.

NETWORK UPGRADES

- Crystal Gateway 4 (CG4) LAN
- CASI LAN

CG4 LAN UPGRADES

- Basic Office Automation Functions
- Mainframe & Computer Networks/Systems Access
- Anonymous FTP & Gopher Access/File Creation
- Client/server Functions

CASI LAN UPGRADES

- Terminal Replacement
- Additional PCs
- Basic Office Automation Functions
- Mainframe & Computer Networks/Systems Access
- Anonymous FTP & Gopher Access/File Creation
- Client/Server Functions
- Establish Redefined Platform Environment

APPLICATION NETWORK UPGRADES

- Commercial Off-The-Shelf Software
- User Interface & Client Development Tools
- Redesign Of Application Systems
- Integration Of Application Software

Standards in the Architectural Effort

Howard Markham

STI Architecture Group

You can think of this as the beginning of a brainstorming session because that's probably the state of the organization of this presentation (Viewgraph 1). Looking at the title of the presentation, you might say, "What STI architecture? What standards?" Those are both very big subjects and I'll try to address them in a brainstorming fashion, where, in an early stage, you are looking at both questions from an STI world view. Who is looking at them? Karen Kaye heads an architecture group, and I am sitting on it along with several people from here. This is the purpose of the STI architecture group:

- To build a framework for modernization (reinvent the STI Program) acquisition,

- To provide guidance and direction for STI Program standardization and integration efforts, and

- To use a reference model--a set of concepts, interfaces, entities, that provides a basis for specification of standards.

So, I thought I'd try to address these topics, maybe not in that order, as a way to give a setting for this effort (Viewgraph 2). Since we just started, I am not going to answer the questions shown here, or even propose any, but outline the issues that we are dealing with. A lot of these questions are far easier for many of you to answer than they are for me since I haven't been around STI very much.

What Is the STI Data Processing Architecture?

Here's a quick statement of what I think of when I hear the phrase *STI Data Processing Architecture*: it is a set of diagrams and descriptions that characterize the principal functions and services and the hardware and software components used for the functions (Viewgraph 3).

What Is STI Infrastructure?

A first answer to this question might be "the complex of facilities, equipment, processes, and staff that operate behind the scenes to provide services to STI users" (Viewgraph 4). I wanted to show here a few ways of looking at it. From the point of view of a user, STI is merely the

(virtual) library that has the information that the user wants or needs for his work as a NASA scientist, or whatever. And that's all he cares about - a place to go and get the information. However, as depicted in the lower part of the diagram, these issues that ideally are transparent to users are the focus of the people who run STI: Where is the data stored? How is it organized? What technologies are we using? What applications do we have to write to make it work? What does the user interface look like? All of those questions are immaterial to the user if there's a good user interface that allows him to specify the kinds of information he wants. That's really all he cares about - as long as what's behind the interface, namely, the infrastructure, works in a way that allows him to find the information and retrieve it. Just as much of the technology should be transparent to users, within STI there is also a level of technology that is more or less transparent. These days that would be something like communication networks, computer operating systems. You can generally go out in the market and buy interoperative pieces and built a pretty complex computing architecture and network architecture in a fairly straightforward way in the 1990s. But, you still have to write the applications, and you still have to be concerned with the data organization that does the delivery to the user of what he wants. That's what I grouped under this heading called *applications*. Here in the STI Program, all of these things that we hope are transparent to users are highly visible and are actually nettlesome issues in most cases.

What is the STI Data Processing Architecture?
Logical View

This is a very simple, graphical depiction of the idea that here is an infrastructure, conceptually, to a user (Viewgraph 5). To an author who might want to contribute to this body of information, it is just a catalog which tells him what is in the library; he can access through these functions. So, when we try to draw a picture of what is the STI architecture, we might start with something like that and work our way down. I missed a couple of the architecture group meetings, but I believe they have made some attempts to draw more detailed pictures. We have that conceptual vision picture that was shown earlier today, about what is the STI future, and in the back of the upgrade infrastructure document there is a detailed wiring guide of the computer installation at CASI. What we need is something in between those pictures, something deeper than this that also identifies the functional components.

Question: Can authors be users?

Answer: Authors are also users, most of the time.

Example of Technical View

I have been assisting the AIM program at NASA Headquarters, where they design the business applications that are used NASA-wide in managing their computing architecture. This is a picture we currently draw of what it would look like (Viewgraph 6). Today it's still

IBM host, 3270 terminal access. They are about to start buying servers and workstations and client-server software products. Ultimately, the architecture discussion has started looking something like this, where you show the key layers, the systems software, the application software in the various kinds of platforms that you have wired together to build your computing architecture. I don't want to step through that; it just shows the layers of communication that connect the pieces of a distributed application. So far, standardization has not gotten to the point where we can just buy entire packages of this off-the-shelf together. We have to be careful which products we choose in each category. So, that's a little bit about architecture. The STI Architecture group is meeting biweekly and is trying to describe the STI architecture in a way that is useful for managing future upgrades, buying products, that kind of thing. When we will finish, Karen knows; I'm not sure I do.

Role of Standards in an Architecture

A few words about the role of standards (Viewgraph 7). I'll mention the NIST application portability profile for an open systems environment in a couple of minutes. You will see there that the focus of the standards that the industry likes to discuss and, to some degree, implement, is on interfaces between components; for example, a POSIX-compliant operating system has a certain specified form for calling the functions of the operating system. It's the calling interface from the application program that is standardized, but the vendor decides precisely how he is going to implement it. Every vendor's operating system that supplies a POSIX interface is different, even if it's a UNIX operating system. Every vendor's UNIX operating system is implemented in a different way; the operating system itself is a proprietary product. But to the extent that the POSIX interface exists on the product, and to the extent that POSIX is a standard, then if I write an application that does only POSIX calls to the operating system, it should run on any of those vendors' POSIX-compliant operating systems products. That's one example of the idea that it's the interface; if you get the interfaces to mesh, you don't really care how the engine works inside the black box - as long as it gives the performance you want.

Scalability and Interchangeability

Having standard interfaces promotes these kinds of things that most people have heard about for several years - scalability and interchangeability. So, if I have a server that has 20 Specint92 units of power and I want to have a server rated at 40 Specint92 and the server is a POSIX server that talks to the networks, then all I have to do is replace the 20 Specint92 server with the 40 Specint92 server. That larger one may be from a different vendor; it may have a lot of other different features, but as long as it has those standard interfaces you can just unplug one and plug in the other. It gives you scalability; it allows you to operate the technology in the similar way. It should be more economical; you have vendors competing to provide the same set of services in a standard fashion, with a standard appearance. If there is somebody that wants to commercialize a product for his own business interest, then the fact

that there is a standard market out there makes it a more attractive kind of market to enter - a bigger market having a set of standards that you could lean on can then be used to guide procurements.

Architecture vs Standards

The application people have to know what the environment is for which their application is being built (Viewgraphs 8 and 9). The architecture is a description of the system environment. One way of describing an information system architecture is to talk about three components: data architecture, application architecture, and what some people called technical architecture, which is the hardware and the system software. It may be useful to draw pictures of each of those areas for STI in the architecture group.

Need for STI Standards for Inherently Diverse Environment and the Need for STI Standards for Modernization of Technology

These two charts summarize aspects of the NASA STI environment that seem to emphasize the potential value of standards (Viewgraphs 10 and 11). One highlights the inherent diversity of the environment, to include points of origin and use that span disciplines and national boundaries, and a wide range of relevant technologies. The second draws attention to the modernization of technology that is now becoming possible, as is reflected in the recent STI Infrastructure Upgrade Plan.

NASA STI Standards Proposal - Principals

The only item that is new here is the idea that in STI it would be beneficial in the future to have a standards program that is formally recognized as such and identifies people who are responsible for certain areas of standards, certain processes by which the standards might get changed (Viewgraph 12). Certainly, the ERB would be part of the standards superstructure because they would presumably apply the standards guidance for STI in evaluating the various proposals that come before them.

Focus on Interfaces

Viewgraph 12 re-emphasizes the idea that the focus is on interfaces. This is the reason that standards take so long to develop. A standard is essentially an agreement, and if it's of any use, it's an agreement of a lot of people about how to do something. It's important to get that participation, and it's important to go through the consensus building process. It takes time but, in the end, we will get a lot of gain from it. The standards within CASI will be easier to manage than the standards that you are going to agree to with external partners.

Local Standards

I have heard mention of several instances where you do have internal standards in STI or CASI. I think Roland mentioned they are mainly using Windows to build interfaces into products within their architecture at CASI, and that's a standard. That excludes certain other kinds of things. It is advisable to set standards within an organization. You might also say for now we are only going to support TCP/IP communications. Anyway, that's the idea there. Local standards should be created where needed. You need them if there is potential for diversity that would get you into trouble. If that potential exists, then it's probably advisable to agree on some standard approach. It may not be an international/national industry or anybody else standard; it's your standard. As a more elaborate example, for ten years in the AIM program in NASA, the standard has been IBM mainframes running proprietary IBM system software and proprietary Software AG database products. That is the standard. If you are in the Ames Research Center and you want to run applications that are built by the AIM Program, you to have to buy IBM compatible equipment and proprietary Software AG database products. That was a legitimate standard at the time it was chosen (1984) because there was no such thing as international or national standards that would cover all of the kinds of connectivity they had to be able to assure.

STI Standards Proposal - Scope

The standards apply at many levels (Viewgraphs 13 through 16). At the machinery and electronics level, we don't worry about that much anymore because that's been sorted out in the past 20 years in the computing industry, and the standard plug configurations, and standard electrical characteristics, all of that has been pretty well dealt with. This is a big one for STI - storage formats for all kinds of media and information. The communication protocols, the software at the software level above the machinery and electronics level are still issues. User access is an issue and the way you construct queries, the way you format, the way you build interface screens - all those things are areas that need to be addressed.

The Library Function

This might be a newer idea and it may not actually be appropriate; it needs to be worked out. The idea of the STI function, or the library function, is of interest to many different levels within the organization, down to the project level. I'm sure it happens within each NASA scientific research project: the project has certain interests in information and they want to share it with the project members. It would be nice if there were some standard, automated way to build that local library and, as papers are written within that arena, within that project, they could be fed back into the wider repository. There would be some easy way to just say, "Okay, this paper is ready to be sent to CASI" or wherever. Within a project, a person might have a subset of topics that have been listed and access it. That would be a nice system.

What about the source of standards? How do you know when you have sufficiently specified a set of standards? This is one criterion: the standards need to be complete enough to insure that all NASA STI functions have been included.

NIST APP Services

The Federal Government has published the application portability profile; they issued the second edition in May 1993 (Viewgraph 17). It has to do with application portability, or at least that's the way they originally began thinking about it. They said an application would be portable if it has standard ways of accessing the operating system, the human computer interface, (let's skip software engineering for a moment) the data that it wants to access and manage, the way it interchanges data with external agents, the way it manages graphics, and the way it interfaces with the network. So, they have specified in the application portability profile a set of recommended standards for Government agencies to consider when they are preparing to buy computer or data communications equipment. In addition to these areas, the topic of software engineering has to do with the tools for developing applications. It's not really quite the same kind of animal as these other six are, but they are looking at standards that would guide the buying of CASI products and repositories.

Security Functions and System Management Functions

Then there are two areas that span all of these, namely, security functions and system management functions. System management is a very important subject when you decide to distribute your architecture and start to have multiple servers and hundreds of workstations, all of them general purpose computers to start with that have various mixtures of software on them. You want to keep that software safe to some degree; automated management methods are needed.

OSE Reference Model

Some of you may have seen this - an open systems environment reference model (Viewgraph 18). This comes out of the IEEE, and NIST uses it as an overall picture of these service areas that I just mentioned. But the idea is that this is your platform; you decide what you want this thing to be. Up here are your applications, which are software which you have written or bought off-the-shelf, but they are top-layer software. They operate on this platform and they acquire support from the platform in the form of operating services, network, data management, human computer interfaces and things like that. So, the standards that NIST recommends are interface standards between the application and the application platform. That is one segment. The other area is between the application platform and the external environment. As it happens, a lot of the same services are needed to communicate with the outside environment. There are 35 standards in the current version of the application portability profile, some of them conflicting. You would never use all of them; it's up to a

user to decide which ones he is going to use. Some of them don't have any products available today, so you wouldn't use them, but at least it's a direction. They update this things every year and a half or so.

Long Term Effects of Standards

We can already see the long term effect of earlier standards in 1993 if we look back to 1980s or 1970s. At one time, application developers had to build a lot of this function by hand into their own applications. Communications: definitely the user interface; menu managers; screen managers; database management functions; all of these things used to be built by every application developer. Or, if you were a smart shop, there might be a library that somebody had built so that you could call from that library. It was being done everywhere. There were computers, but you couldn't go out and buy a DBMS off-the-shelf. The communications industry had not agreed on what the protocols were. Today, all of that is settled and we are here. We still have to worry about different operating systems, different network protocols, different DBMS. The DBMS all support SQL, but SQL isn't a strong enough standard yet.

Technology Advance

Some of the ideas we are going come to grips with in the STI architecture effort are a description of the STI architecture and an outline of what an STI standards program might look like and how the standards would fit into the architecture. I might just mention, although it's obvious: the kinds of technology advance we see happening now are going to keep happening during the modernization period of five years. How useful will it be to have something like an architecture, a set of standard guidelines? Obviously, there is some consensus among you on what it is already, but it's sort of in people's heads. It has been sufficient to get you this far. Maybe this business of formalizing it a bit will streamline future efforts.

Standards for STI Architecture

H. Markham, MITRE

September 9, 1993

MITRE

Outline

- What is the STI Data Processing Architecture?
- What is the role of standards?
- Overview of NIST Application Portability Profile
- Standards for STI
- A Standards Program for STI

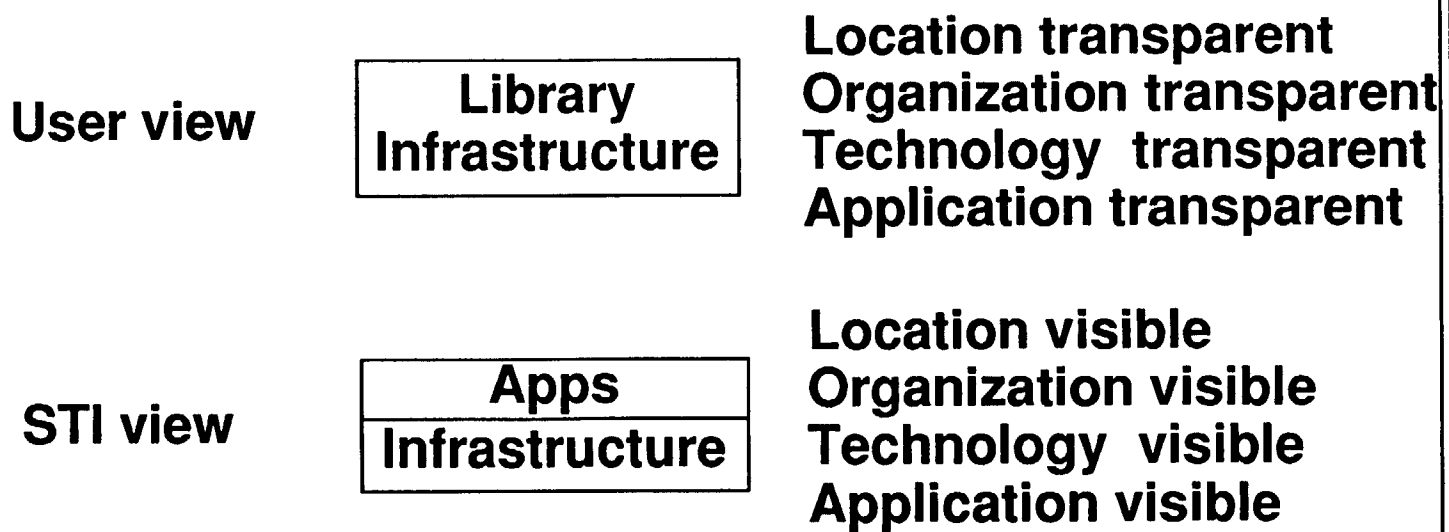
What is the STI Data Processing Architecture?

- Schematic of data processing systems and equipment for performing the following functions for STI documents and other sources
  - Acquire
  - Catalog
  - Store
  - Search
  - Retrieve
  - Present

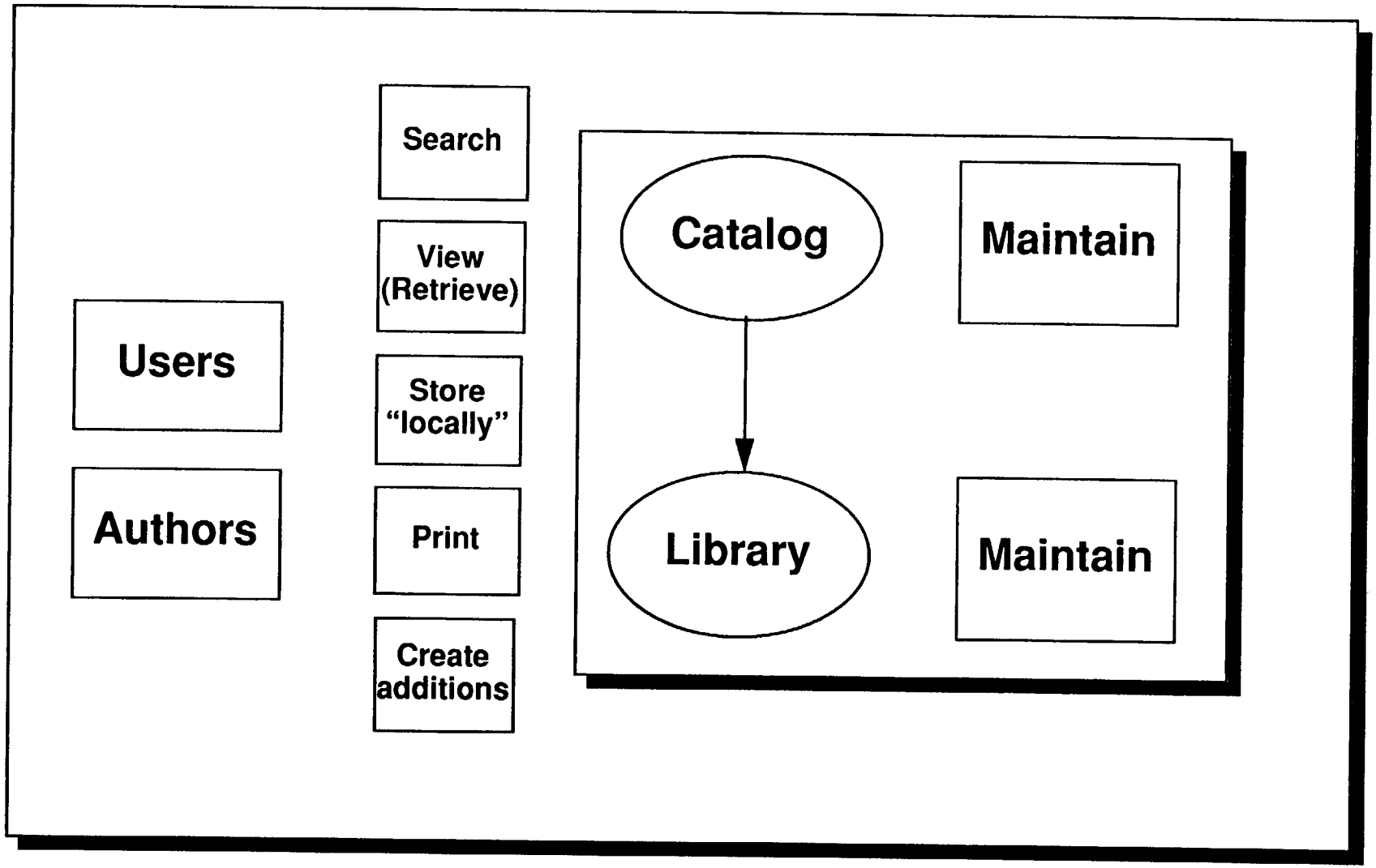
MITRE

What is STI Infrastructure?

- User view: hardware, software, and data resources used to make available the *search, retrieve,* and *present* functions
- STI view: hardware and software resources used to create a data processing environment for applications that support all STI functions

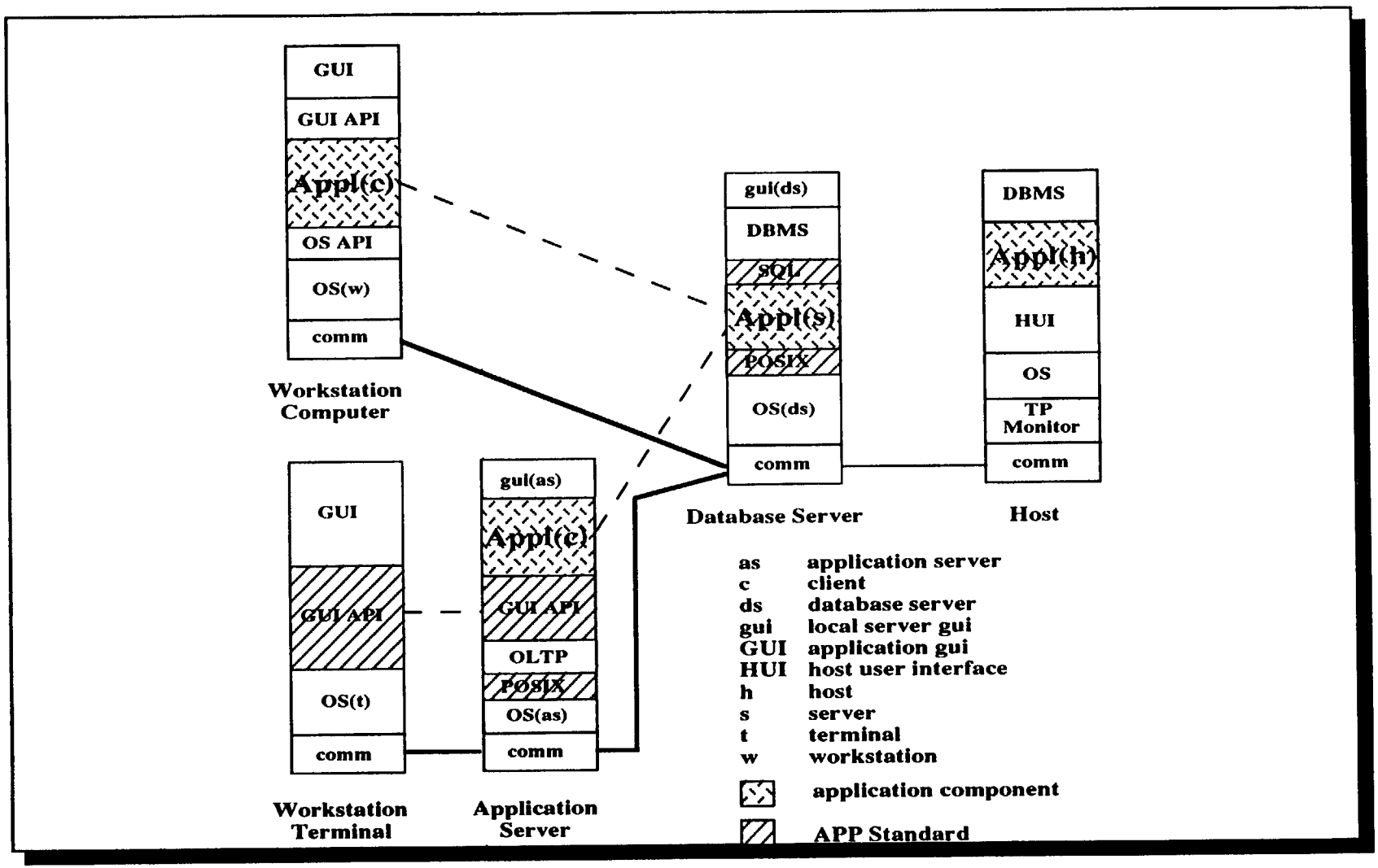


What is the STI Data Processing Architecture? Logical View



MITRE

Example of Technical View



MITRE

Role of Standards in an Architecture

- Focus is interfaces between components
- Standard interfaces promote
  - Scalability
  - Technology "refreshment"
  - Economy
    - Commercialization
    - Competition
  - Interoperability
- Standards guide procurements
- Standards guide application developers
- Standards guide management and planning

MITRE

Architecture vs Standards

- Architecture describes major components of system
- Information system architecture
  - Data architecture (ER diagrams, data dictionary, databases and files)
  - Application architecture (programs, COTS software, data flows)
  - Technical architecture (workstations, mainframes, servers, networks)
- Standards define characteristics that facilitate
  - Interoperation between components
  - Interchange of components

Architecture vs Standards (concluded)

- Information system standards define
  - Interfaces between architecture components
  - Formats of data elements and files
  - Syntaxes accessing services and data
- Architecture will guide
  - Procurements
  - Application developers
  - System operators
  - Future planning (including standards)
- STI standards program will encompass
  - Organization, roles and responsibilities
  - Activities (sponsorship, monitoring, review, waivers)
  - Architecture as framework for standards

MITRE

Need for STI Standards
Inherently Diverse Environment

- Many sources of STI
  - Points of origin
    - NASA and other US Government agencies
    - Scientific community and industry
    - Foreign countries
  - Accumulators and disseminators
    - NASA center STI programs
    - Other US Government agency STI programs
    - Scientific and professional organizations
    - Commercial organizations, foreign organizations
- Many media and technologies
- Many users
- User interests cross organizational and national boundaries

Need for STI Standards
Modernization of Technology

- Increasing automation of information
  - Word processing
  - Document scanning
  - Compound documents: text, images, audio, video
  - Retrieval and display
  - Printing and publishing
  - Searching and indexing
  - Translation
- Competing technologies
- Competing products and formats
- Competing standards (industry, national, international)

MITRE

NASA STI Standards Proposal Principles

- Standards are required primarily for interfaces between distinct processes, organizations, and machines
- All parties affected by standards should participate in their development or selection
- Standards within an organization can be managed more completely than standards for interaction with external partners and users
- Standards intended for broad use require consensus among many competing parties and evolve slowly
- Local standards should be created where needed if industry, national, and international standards do not exist

NASA STI Standards
Scope

- Standards apply at several levels
  - Machinery and electronics
  - Information storage formats
    - Text
    - Catalog entries
    - Drawings
    - Images
    - Audio and video
  - Communications protocols
  - User access conventions
  - Search query syntax

MITRE

NASA STI STandards
Scope (continued)

- Standards in NASA STI should facilitate the following broad purposes throughout the agency
  - Acquisition of STI
    - Printed reports
    - Electronic sources
    - Audio and video media
  - Storing and cataloging
  - Dissemination of STI
    - Search facilities
    - Retrieval facilities
    - Access from within agency
    - Access by external partners

NASA STI STandards
Scope (continued)

- Standards in NASA STI should facilitate the following broad purposes throughout the agency (continued)
  - Creation of local STI activities down to projects and persons
    - Possibly of limited duration
    - Initial population of local repository
    - Access control by document class
    - Accommodate documents in preparation
    - Close out and selectively assimilate material into continuing STI repositories

MITRE

NASA STI STandards Scope (concluded)

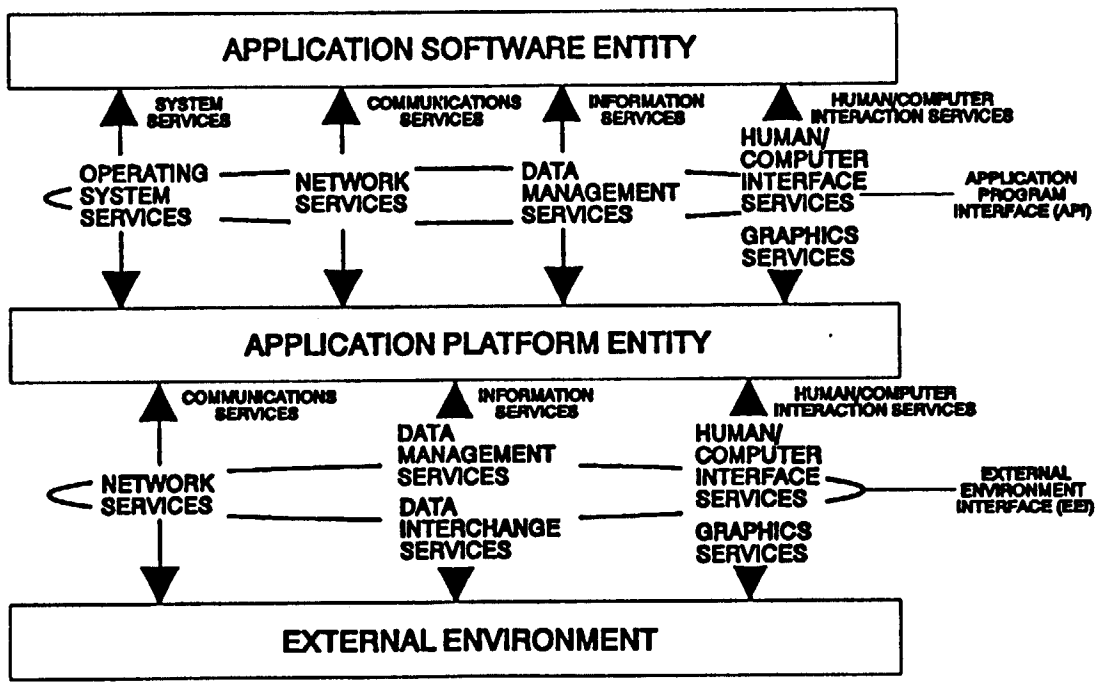
- NASA STI Program should establish STI standards for agency based on
  - NIST standards
  - Industry standards
  - Other
- Standards should be complete enough to ensure that all NASA STI functions are included
- NASA STI should identify preferred standards for information exchange with external partners
  - Access to external organizations and sources should be negotiatied with each partner
- Participate in committees and other organizations that develop standards

NIST APP Services

- Operating system
- Human/computer interface
- Software engineering
- Data management
- Data interchange
- Graphics (*not applicable to AIM*)
- Network
- Integral supporting services
  - Security
  - System management

MITRE

OSE Reference Model



Long Term Effects of Standards

- Application developers can focus on higher level functions
- It is no longer necessary to develop
  - Communications functions
  - User interface processors
  - Database management functions
  - Format translators
  - Character set translators
- Eventually, it will not be necessary to develop
  - Different versions for different operating systems
  - Different versions for different GUIs
  - Different versions for different network protocols
  - Different versions for different DBMSs

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 1993	Technical Memorandum

4. TITLE AND SUBTITLE
NASA STI Program Coordinating Council,
Eleventh Meeting: NASA STI Modernization Plan

5. FUNDING NUMBERS

6. AUTHOR(S)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Headquarters

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
TM-109415

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 82

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

The theme of this NASA Scientific and Technical Information Program Coordinating Council Meeting was the modernization of the STI Program. Topics covered included the activities of the Engineering Review Board in the creation of the Infrastructure Upgrade Plan, the progress of the RECON Replacement Project, the use and status of Electronic SCAN (Selected Current Aerospace Notices), the Machine Translation Project, multimedia, electronic document interchange, the NASA Access Mechanism, computer network upgrades, and standards in the architectural effort.

14. SUBJECT TERMS	15. NUMBER OF PAGES
NASA programs, conferences, information systems, data bases, computer networks, architecture (computers), standards, electronic mail, multimedia	147
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASS	UNCLASS	UNCLASS	UNLIMITED

Available from NASA Center for AeroSpace Information
800 Elkridge Landing Road
Linthicum Heights, MD 21090