/IN-82

/50549

P. 10

NASA Access Mechanism - Graphical User Interface Information Retrieval System

J. Hunter, D. Duncan, and C. Generous

**NASA STI PROGRAM
SCIENTIFIC &
TECHNICAL
INFORMATION**

NASA ACCESS MECHANISM — GRAPHICAL USER INTERFACE INFORMATION RETRIEVAL SYSTEM

Judy Hunter
Curtis Generous
National Aeronautics and Space Administration
Scientific and Technical Information Program
Washington, DC 20546
jhunter@sti.nasa.gov — cgenerous@sti.nasa.gov

Denise Duncan
Logistics Management Institute
Bethesda, MD 20817
dduncan@sti.nasa.gov

Abstract

Access to online information sources of aerospace, scientific, and engineering data, a mission focus for NASA's Scientific and Technical Information Program, has always been limited by factors such as telecommunications, query language syntax, lack of standardization in the information, and the lack of adequate tools to assist in searching. Today, the NASA STI Program's NASA Access Mechanism (NAM) prototype offers a solution to these problems by providing the user with a set of tools that provide a graphical interface to remote, heterogeneous, and distributed information in a manner adaptable to both casual and expert users. Additionally, the NAM provides access to many Internet-based services such as Electronic Mail (Email), the Wide Area Information Servers (WAIS) system, Peer Locating tools, and electronic bulletin boards (USENET). The NAM is constructed using a tool kit approach, making use of standards (UNIX, TCP/IP, X Windows, OSF/Motif) and tools that make it adaptable to various host and workstation platforms.

Background

The NASA Scientific and Technical Information (STI) Program has a requirement to provide its user community with easy access to relevant scientific and technical information (STI) worldwide. This information requirement covers research and development activities ranging from basic research to applied engineering and maintenance, and involves nearly every discipline. The kind of STI needed by this community includes complete reports of basic research, observation data, graphic representations of observation data, and information available through personal contacts.

Some of this information is available in several online database systems such as the NASA STI Database or bulletin board systems, but the information from peer contacts is a vital part of solving the information requirement. Peer contacts are made using the telephone, correspondence, and (increasingly) electronic mail.

In 1990, the STI Program gained new resources for meeting its burgeoning user requirements, a new management staff that has experience in networked access to multiple databases and knowledge about major technical advances such as the Intelligent Gateway Processor (IGP), which was developed at Lawrence Livermore National Laboratories (LLNL) in 1983 under sponsorship of NASA, the Department of Energy (DoE), and the Department of Defense (DoD).

With these new resources, the STI Program quickly initiated a study of user requirements to determine the feasibility of providing an intelligent gateway system to assist users with access to multiple sources of STI. The report resulting from this study *(NASA Gateway Requirements Analysis Report, #NS001R1)* stated that the user community utilized diverse sources of STI, and that a significant number of STI users wanted to access information without leaving their offices. Another finding was that the Internet is a powerful research tool, but that knowledge and use of the Internet was found only randomly throughout the community. The report's recommendation was to provide a prototype system to the user community, to demonstrate facilitated desktop access to needed sources of STI.

Design Considerations

NASA is using a rapid prototyping approach to allow its users to participate in the process of designing the final system. Users are the best evaluators of the relative value of the information available from the various sources; only users can decide the appropriate level of assistance and best 'look and feel' to be

provided by the interface. Once the community has had some experience with the prototype, they will be better able to articulate what must be provided by a production version of the system.

This prototype builds on the work of many other organizations. Many of its facilities are a synthesis of ideas and existing capabilities from other projects, presented in an intuitive, easy-to-use manner. For example, the IGP software can use the communications and networking capabilities of a host system to communicate over both synchronous and asynchronous dial-up and leased lines, and to use TCP/IP to access other hosts located on the Internet. It also provides a capability to communicate simultaneously with multiple connections under the control of intelligent 'scripts' to automate logins, transfer files, control user access to shared I/O (input/output) resources, select optimal routing, and initiate tasks on the target machine(s). However, the IGP was written for the type of terminal common in the 1980s, which transmitted information a line at a time to the host system; it needed to be brought into the technology of the '90s and beyond.

Many technological improvements have become available since the IGP was developed. For example, today bit mapped workstations and graphical user interface (GUI) software, make it possible to provide an icon-driven interface. This type of interface provides the user an intuitive approach to performing more complex tasks in a rich information environment. Distributed, client-server architectures have replaced the centralized resource control of the '80s and make it possible to take advantage of the information retrieval applications and standards of the '90s.

In addition to management and technological changes, NASA itself has undergone a demographic change. It now has a workforce that includes more recent graduates who are accustomed to using computers and databases themselves. Many of these users are eager to conduct online research themselves from their desktop workstations.

In the fall of 1991, NASA began to design the NAM, based on the users' requirements stated in NS001R1 as well as the lessons learned from earlier gateway efforts at other agencies. For the first prototype, the following functional requirements were deemed to be essential for the NASA community:

- A system which could be used by scientists or engineers from their desktops, around the clock

- A tool to assist users in locating an appropriate source of information for their subject

- Simple access to the sources of information commonly used in the STI Program user community

- Assistance in using the information sources to perform the most common queries

- Different levels of assistance for users who are inexperienced in using these sources of information, and expert users of these systems

- Assistance in viewing, extracting, and saving retrieved data

- Tools to make users aware of Internet facilities, and which simplify their use.

NAM Architecture

The NAM prototype system consists of eight separate modules:

- The NAM Server
- Communications Tool Kit
- Intelligent Resources Locator
- Database of Databases
- Email tool
- Utilities
- X11 R4 Client and OSF/Motif tool kit
- X11 R4 Server.

The interaction of each of these modules is depicted in Figure 1:

Each module communicates with other modules using a clearly defined protocol. The underlying mechanism used to provide this support is based upon the Inter Process Communications (IPC) abstractions provided by the UNIX Socket Interface (the BSD socket library is a widely accepted communications abstraction, supporting many platforms other than UNIX), which supports peer-to-peer communications primitives well suited for client/server and distributed application architectures.

The design of the NAM prototype facilitates distributing each module to different processors. There are several reasons for this design. First, to improve performance and reliability by distributing portions of the application to processors at the users' sites. Second,
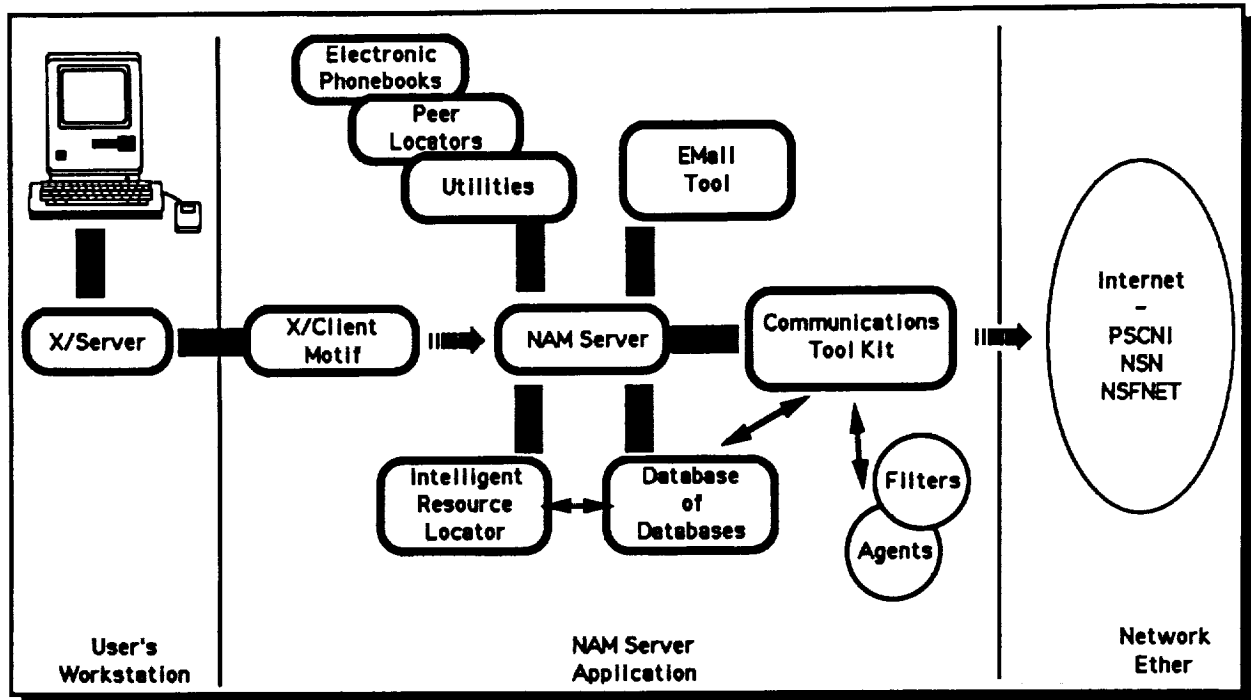
Figure 1. NAM Prototype System Modules

to allow for customization at the module level for the geographically dispersed users. Third, to accommodate a multi-vendor, multi-platform environment. Fourth, to allow for dedicated machines for certain tasks, such as database machines, where needed.

The NAM Server includes the Communications Tool Kit, which in itself provides a modular approach to providing communications capability to information retrieval systems external to the NAM. The Communications Tool Kit modules provide separation of the user interface from the code, to make NAM a true server with no direct user interface. This facilitates modification of the user interface in response to user testing, while not affecting the server code. Another module provides batch and delayed operations. This allows certain time intensive tasks to be scheduled for execution at off-peak hours when communications costs are lower or when contention for network bandwidth is not as severe.

Another module of the NAM Server is the Intelligent Resource Locator. This module provides information about the various remote sources and their coverage of subjects of interest to the users. It provides a mechanism to update a source's relevance score, for the chosen subject based on the users' disposition of search results from that source. This module works

closely with the Database of Databases, which maintains the records of the remote sources, their subject coverage, and relevance scores for each subject.

The final server module is an electronic mail tool, which provides access to Internet mail via an X-based user agent. It includes gateway capabilities to many other mail systems, averting the need for users to know the routing syntax from the NAM mail system to the other systems.

The NAM client software is a custom built windowing application, tailored to the requirements of the STI Program user community. This GUI is built using the X-Windows system and the OSF/Motif toolkit. Use of standard software for the client ensures that the application will be portable to the variety of platforms found in the STI Program user community, such as UNIX workstations, DOS-based personal computers, and Macintosh systems, while providing a unified look-and-feel. The prototype system is being developed in a distributed, multi-vendor platform environment, with the use of standards emphasized to maximize the potential for migration to other platforms in future versions.

The NAM includes several components that can be distributed among geographically dispersed platforms, linked by a network running the TCP/IP protocols. The

3

system requirements for each module are as follows:

- NAM server module requires C on a UNIX platform (currently a Sun SPARC 4/470)

- Communications Tool Kit module requires PERL on a UNIX platform

- Windowing system requires X11 Release 4 of the MIT X-Windows system

- Look-and-Feel requires OSF/Motif tool kit

- Database service requires INGRES with SQL/C

- User workstations must be X11 Release 4 servers. Current user workstations supported include Sun, Macintosh running Exodus, and personal computers running Microsoft Windows and Exodus.

The communication path between the various components is provided by the TCP/IP protocols-based Inter Process Communications facilities, supported on Local Area Networks and Wide Area Networks. The NAM communication tool kit accesses the NASA Science Internet (NSI) and the National Science Foundation Network (NSFNet), which are used as the access paths to the remote information retrieval services.

Status of the Prototype

The NAM was rapidly developed and was first demonstrated to the user community in April of 1992. This version of the prototype provides the following:

- A source locator that searches a database of the available sources to select those sources which cover the subject of interest

- A query formulation window that assists users in performing the most common types of searches

- A brief display of the retrieved information that allows the user to use the mouse to select a display of the full citation

- A window that presents the retrieved information in its entirety and allows the user to mail, print, fax, or save the retrieved information, or order a copy of a report cited in the bibliographic database

- A peer locator function that provides several ways of locating contact information for NASA employees and contractors: an X.500 directory user agent; electronic telephone books for most of the NASA centers; Knowbots, which search the user files of a number of mail systems; and Internet Whois and Finger utilities.

- Bulletin board systems that assist users in communicating with other Internet users (USENET), or searching large numbers of remote hosts (WAIS, Archie, and Gopher).

These facilities are all provided through an iconic, form-based interface, where the user fills in a form with the keyboard or selects options with the mouse to prepare a query. The NAM handles connection to the remote system, selection of the appropriate application, and processes the query at the remote host. The NAM also presents the results of that remote session to the user in windows that allow the user to perform further action on the retrieved information.

Figures 2-6, which go through the steps to query a remote bibliographic database, illustrate the functionality and look-and-feel of the NAM user interface:
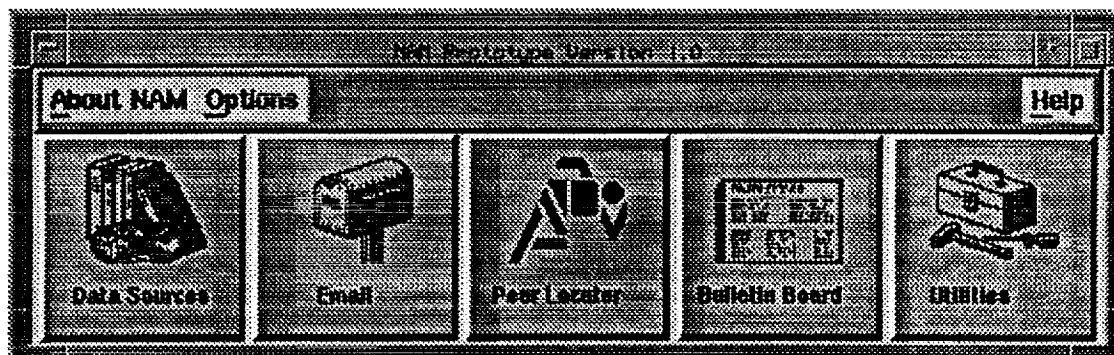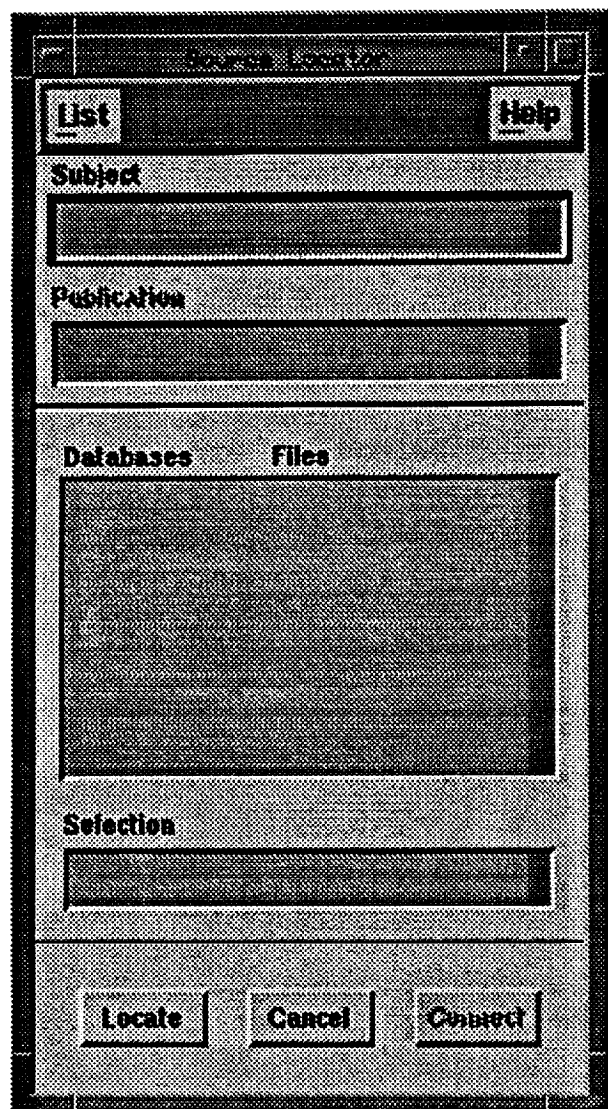

Figure 2. NAM Main Menu

Figure 3. Source Locator Screen

1. The user clicks on the Data Sources icon from the NAM Main Menu (see Figure 2) with the mouse.

2. The NAM then displays the Source Locator window (see Figure 3). The user enters the subject of interest by typing it in the Subject area, or by selecting a subject from the List of subjects under the List pulldown. When the Subject area is complete, the user clicks on the Locate button and the NAM displays the suggested source(s) for that subject in the Databases and Files area. The user then selects a source with the mouse and clicks on the Connect button, and the NAM connects the user to that remote system and to the appropriate file in its database.

3. Once the connection is made, a query form at the selected level of expertise (Novice, Intermediate, or Expert) is displayed (see Figure 4 for an example of an Intermediate screen). The user fills in at least one of the possible parameters for the query. When the Search button is selected with the mouse, the NAM parses the search parameters and translates the query into the appropriate syntax, displays the translated query, and transmits the query string to the remote system. The user can search iteratively by adding new search parameters and applying them to the most recently retrieved set. When each search is completed, the number of records found is displayed in the Number of Hits: area. At that point, the View button is active; if the user clicks on it, the NAM displays the Range Selection Menu, where the user selects the number of citations to view, and clicks on its View button.

4. The NAM then displays the selected range of records in their short format (see Figure 5), where each citation is itself a large button.

5. The user selects those citations of interest with the mouse and clicks on View to see the selected records in their entirety (see Figure 6). At this point the user may perform the operations shown at the right of the full citation display.

A similar visual, form-based interface is provided for the other NAM facilities.

Future Plans

The NAM prototype will be evaluated by users at NASA Centers, the European Space Agency (ESA), and some NASA Program offices. This evaluation will focus on several issues: 1) What proportion of scientists and engineers prefer to conduct on-line research themselves?, 2) Which of the offered search mechanisms is preferred?, 3) Are most of the frequently-asked queries supported?, 4) What are the priorities in providing other forms of information besides text?, 5) What are the priorities in providing other sources of information?, and 6) Which Internet tools are needed most?

One requirement of the future NAM is that the user interface be easily customized for specific user communities. Another requirement for the future is to provide for two- and three- dimensional visualization and rendering of data sets, both for user input and results output. This can, for example, provide searching of a chemical database by structure, and visualization of a retrieved dataset. Both of these capabilities exist, either

5

Figure 4. Intermediate Screen

as a function of the remote source, or as commercial or public domain software that could be incorporated into the NAM. A third improvement would provide mail gateway access. This would allow users with no direct network connection to access NAM by mail. The user would fill in a form on his/her local processor and mail it to a specified address. The form would be decoded by the mail gateway and fed to NAM, which would execute the canned operation or query. Results would be mailed back to the originator of the message.

Using the evaluation information collected during prototype testing, the NASA STI Program will develop a report of the lessons learned from the prototype, and begin the design of the future NASA Access Mechanism.
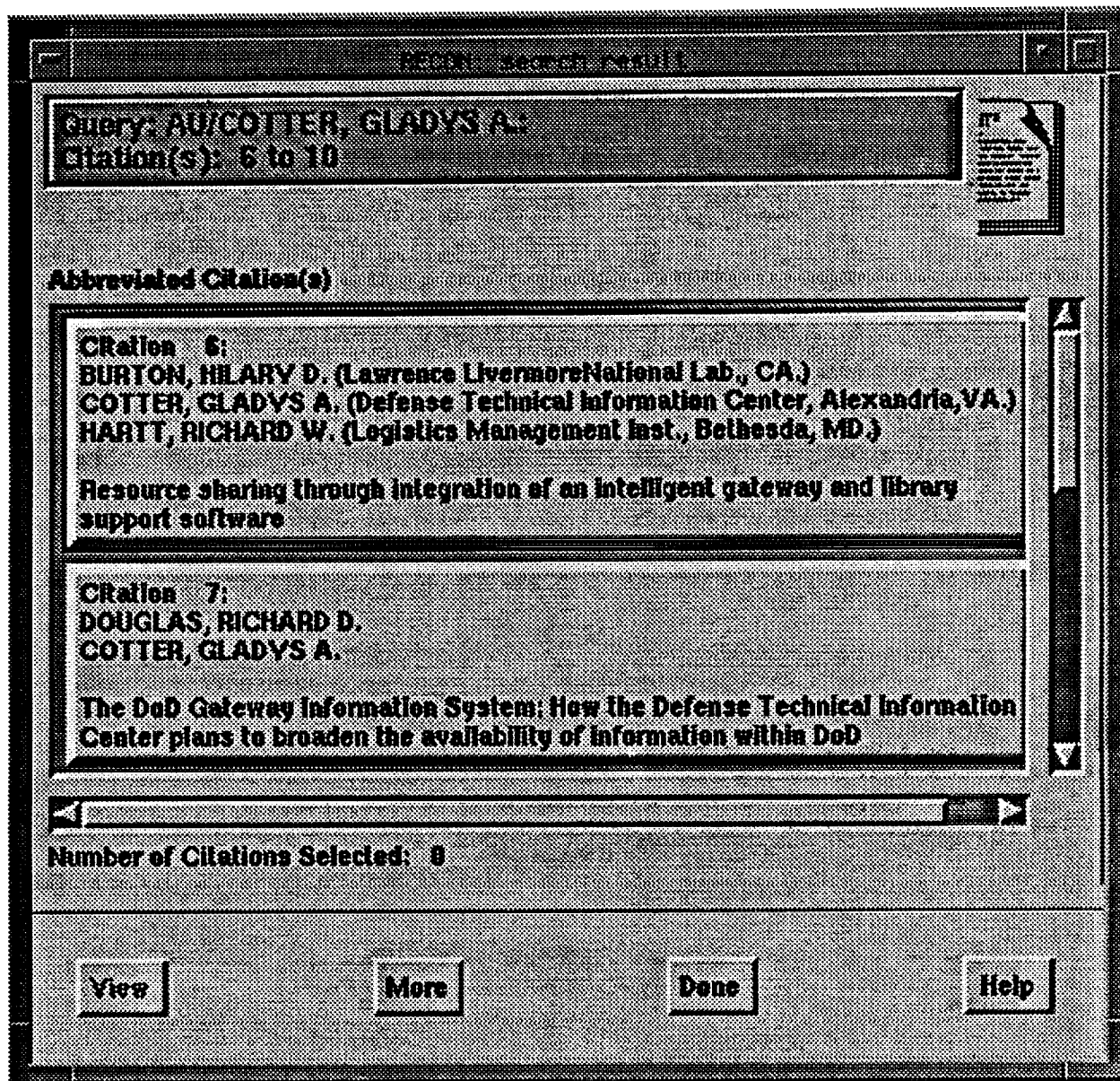
6

Figure 5. Abbreviated Citations Screen

Figure 6. Long Citations Screen

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE January 1993	3. REPORT TYPE AND DATES COVERED Technical Memorandum

4. TITLE AND SUBTITLE
NASA Access Mechanism - Graphical User Interface Information
Retrieval System

5. FUNDING NUMBERS

6. AUTHOR(S)
J. Hunter, D. Duncan, and C. Generous

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-108981

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
UNCLASSIFIED - UNLIMITED
SUBJECT CATEGORY 82

12b. DISTRIBUTION CODE

13. ABSTRACT *(maximum 200 words)*

Access to online information sources of aerospace, scientific, and engineering data, a mission focus
for NASA's Scientific and Technical Information Program, has always been limited by factors
such as telecommunications, query language syntax, lack of standardization in the information,
and the lack of adequate tools to assist in searching. Today, the NASA STI Program's
NASA Access Mechanism (NAM) prototype offers a solution to these problems by providing
the user with a set of tools that provide a graphical interface to remote, heterogeneous, and distributed
information in a manner adaptable to both casual and expert users. Additionally,
the NAM provides access to many internet-based services such as Electronic Mail (EMail), the Wide
Area Information Servers (WAIS) system, Peer Locating tools, and electronic bulletin boards (USENET).
The NAM is constructed using a tool kit approach, making use of standards (UNIX, TCP/IP, X Windows, OSF/Motif)
and tools that make it adaptable to various host and workstation platforms.

14. SUBJECT TERMS
EMail, NAM, tool kit

15. NUMBER OF PAGES

16. PRICE CODE

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

Available from NASA Center for AeroSpace Information
(301) 621-0390