THE WIDEST PRACTICABLE DISSEMINATION: THE NASA TECHNICAL REPORT SERVER

Michael L. Nelson and Gretchen L. Gottlich
NASA Langley Research Center, Hampton, Virginia

David J. Bianco
Computer Sciences Corporation/NASA Langley Research Center, Hampton, Virginia

Robert L. Binkley and Yvonne D. Kellogg
NASA Dryden Flight Research Center, Edwards, California

Sharon S. Paulson
Institute for Computer Applications in Science and Engineering, Hampton, Virginia

Chris J. Beaumont
Computer Sciences Corporation/NASA Ames Research Center, Moffett Field, California

Robert B. Schmunk
NASA Goddard Institute of Space Studies/National Research Council, New York, New York

Michael J. Kurtz, Alberto Accomazzi
Smithsonian Astrophysical Observatory, Cambridge, Massachusetts

Abstract

The National Aeronautics and Space Act of 1958 established NASA and charged it to "provide for the widest practicable and appropriate dissemination of information concerning [...] its activities and the results thereof." The search for innovative methods to distribute NASA's information lead a grass-roots team to create the NASA Technical Report Server (NTRS), which uses the World Wide Web and other popular Internet-based information systems as search engines. The NTRS is an inter-center effort which provides uniform access to various distributed publication servers residing on the Internet. Users have immediate desktop access to technical publications from NASA centers and institutes. The NTRS is comprised of several units, some constructed especially for inclusion in NTRS, and others that are existing NASA publication services that NTRS reuses. This paper presents the NTRS architecture, usage metrics, and the lessons learned while implementing and maintaining the services over the initial 6-month period. The NTRS is largely constructed with freely available software running on existing hardware. NTRS builds upon existing hardware and software, and the resulting additional exposure for the body of literature contained will allow NASA to ensure that its institutional knowledge base will continue to receive the widest practicable and appropriate dissemination.

Nomenclature

ADS	Astrophysics Data System, Smithsonian Astrophysical Observatory, Cambridge, MA
CASI	NASA Center for Aerospace Information, Linthicum Heights, MD
CGI	Common Gateway Interface
CNIDR	Clearinghouse for Networked Information, Discovery, and Retrieval, Research Triangle Park, NC
DFRC	NASA Dryden Flight Research Center, Edwards, CA
DTRS	Dryden Technical Report Server
FTP	file transfer protocol
GISS	Goddard Institute for Space Studies, New York, NY
GPOL	GISS Publications On-Line

GSFC	NASA Goddard Space Flight Center, Greenbelt, MD
HTML	Hyper Text Markup Language
HTTP	Hyper Text Transfer Protocol
ICASE	Institute for Computer Applications in Science and Engineering, LaRC, Hampton, VA
ITRS	ICASE Technical Report Server
LaRC	NASA Langley Research Center, Hampton, VA
LTRS	Langley Technical Report Server
MARC	Machine Readable Cataloging
NAS	Numerical Aerodynamic Simulation Division, ARC, Moffett Field, CA
NCSA	National Center for Supercomputing Applications, University of Illinois Urbana-Champaign
NTIS	National Technical Information Service
NTRS	NASA Technical Report Server
SCAN	Selected Current Aerospace Notices
STELAR	Study of Electronic Literature in Astronomical Research
STI	NASA Scientific and Technical Information
URL	Uniform Resource Locator
WAIS	Wide Area Information Server
WWW	World Wide Web

Introduction

The National Aeronautics and Space Act of 1958 established NASA and charged it to "provide for the widest practicable and appropriate dissemination of information concerning [...] its activities and the results thereof." To meet this goal, researchers at various NASA installations have developed several new methods of distributing information to the nation's research and industrial sectors. One key method is the NASA Technical Report Server (NTRS). The NTRS is an inter-center effort to provide uniform access to various distributed publication servers residing on the Internet. It currently provides access to documents from nine different NASA organizations spanning the United States: Langley Research Center (LaRC), Dryden Flight Research Center (DFRC), Numerical Aerodynamic Simulation Division (NAS) of NASA Ames Research Center (ARC), Goddard Institute for Space Studies (GISS), Institute for Computer Applications in Science and Engineering (ICASE) at LaRC, the Selected Current Aerospace Notices (SCAN) and RECON databases maintained by the NASA Scientific and Technical Information (STI) Program, the Study of Electronic Literature in Astronomical Research (STELAR) Project from Goddard Space Flight Center (GSFC), and the Astrophysics Data System (ADS) Abstract Service at the Smithsonian Astrophysical Observatory.

The NTRS is accessible via the World Wide Web (WWW),[1] a multi-protocol Internet information system, using software such as the freely available and highly popular NCSA Mosaic,[2] developed at the National Center for Supercomputing Applications (NCSA) at the University of Illinois Urbana-Champaign. There are many different WWW "browsers" available, most of them free of charge for all popular platforms (UNIX, Mac, PC, VMS). Implementing NTRS with the WWW reduced the development time necessary. WWW is used for many services, thus NTRS is built on the tools and lessons learned from many other WWW projects. From the users' perspective, NTRS is one of many available WWW services, allowing access with a consistent and well-known interface. NTRS is available at the following uniform resource locator (URL):

http://techreports.larc.nasa.gov/cgi-bin/NTRS

Project Goal

The goal of NTRS is to provide "one-stop-shopping" for NASA technical publications. The intended audiences are researchers and scientists, not information specialists or librarians. NTRS is not the final word in searching and indexing; its intent is to provide maximum connectivity and exposure to the already existing body of NASA electronic literature.

NTRS is a grass-roots electronic document reuse effort. Many researchers prepare their conference papers, technical reports, and journal articles using sophisticated word processing and desktop publishing tools. High quality document preparation systems for personal computers and workstations have enabled researchers to produce fully (or mostly) electronic publications. After the paper is printed and sent to the publisher, it is then assembled and preserved in hardcopy format and often included in a proceedings document. While there are advantages to preserving the document in this format, most researchers maintain an electronic copy of the publication. NTRS attempts to gather the diffuse collection of electronic publications, index them, and offer them to the scientific community.

Fig. 1 NTRS as seen through an X Window System ® World Wide Web browser.

NTRS is not a document life cycle management system; there are many commercial products that address those issues. The internal procedures involved in creating and reviewing documents are the focus of such systems. Some of these systems offer interesting and useful functionality, such as remote collaborative editing and annotating. These systems do not necessarily compete with the components of NTRS because NTRS is focused only on the customer side of searching and retrieval. While NTRS could be expanded to provide access to different document types, only "finished products" are currently indexed in NTRS. NTRS provides access only to publications that have passed through the existing approval and review mechanisms; it does not directly address internal approval processes prior to publication.

Initially, attempts were not made to convert existing paper documents to an electronic form, although it is desirable that all significant publications such as NACA reports eventually be accessible via the WWW. The more difficult problem of adding electronic access to legacy collections is the focus of other projects.[3] The results of these other projects will be included in NTRS when they are available.

Services Provided

The NTRS provides the following services to the user:

1. A single standard interface to multiple disparate NASA technical paper databases.

The NTRS is a virtual "wrapper" script for several different Internet database servers NASA-wide. It simplifies the user interface for all servers to a single, common user format. The NTRS script hides the operational differences of each of the servers from the user and submits the properly formatted query to each participating report server. This is a widely distributed database system, allowing each site participating in NTRS to update and maintain data locally, thus eliminating the need for central administration of the system. Figure 2 shows all current NTRS sites.

2. Simple and rapid searches for information on NASA technical reports.

Since the NTRS is available via the WWW, the user interface provided is a common, familiar, easy to use "point-and-click" style. WWW interfaces are available for most UNIX® workstations, Macintosh® computers, Windows®-based PCs, and DEC VAXen®. The information is available to users' desktops 24 hours a day. Currently, searching is limited to only the abstract and bibliographic information, not the full text of the document. At this time, only Astrophysics Data System (ADS) database implements fielded searches. All other databases do not allow restricting search terms to specified fields (author, title, abstract, etc.).

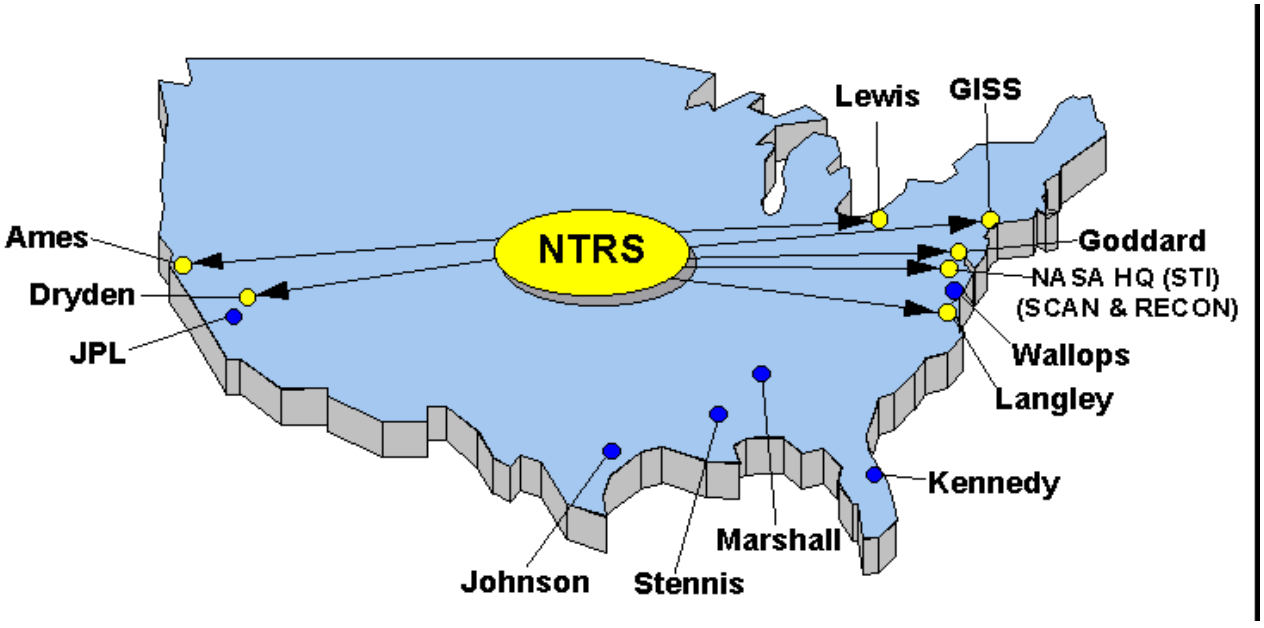


Fig. 2 Map of NTRS sites.

Table 1. NTRS machines.

Service	IP name	HW	Operating system	Dedicated for WWW?	Local WAIS server
NAS	www.nas.nasa.gov	SGI Power Challenge	IRIX 5.2	Yes	No
ADS	adswww.harvard.edu	Sun 690MP	SunOS 4.1.3	No	Yes
DTRS	www.dfrf.nasa.gov	Sun 690MP	SunOS 4.1.3	No	Yes
GPOL	www.giss.nasa.gov	IBM 590	AIX 3.2.5	No	No
ITRS	www.icase.edu	Sun SS10	SunOS 4.1.3	No	Yes
LTRS	www.larc.nasa.gov	Sun IPX	SunOS 4.1.2	No	Yes
RECON	www.sti.nasa.gov	Sun 10/41	Solaris 2.3	No	Yes
SCAN	www.sti.nasa.gov	Sun 10/41	Solaris 2.3	No	Yes
STELAR	hypatia.gsfc.nasa.gov	SGI 4GX340	IRIX 3	No	Yes

Table 2. NTRS contents and usage.

Service	Reports on-line	Citations on-line	Reports served
NAS	95	95	12,100+
ADS	0	165,000	–
DTRS	18	640	167
GPOL	16	363	85
ITRS	92	113	1429
LTRS	429	429	16,500+
RECON	0	270,000+	–
SCAN	0	1,000+	–
STELAR	0	75,000+	–

Note: Contents and distribution for each site not normalized with respect to time available.

3. Rapid delivery of complete copies of technical reports.

If a report is available on the NTRS system, the user may choose to download a copy by a single click. The reports are currently stored in either PostScript® or Hyper Text Markup Language (HTML) format. This is a significant savings in user time and effort. Retrieval time is measured in seconds, not days. If a report is not available on-line, the information is available to allow the user to order a paper copy via more traditional means.

NTRS Architecture and History

Much of the design and toolset for implementing NTRS was developed for the Langley Technical Report Server (LTRS).[4,5] LTRS began in 1993 as an anonymous file transfer protocol (FTP) server with just over 100 LaRC formal technical reports. While anonymous FTP was fine for those familiar with computers, it intimidated some of the more casual users. The transition of LTRS to a WWW interface integrated the keyword searching and document retrieval, and allowed casual computer users to access the publications.

When the LTRS proved useful for researchers at LaRC and beyond, it became obvious that technical publications other than LaRC's should be available as well. Specifically, the implementers of NTRS desired easy electronic access to publications at other NASA facilities. The shell scripts, methods, and other products developed for LTRS were shared, so that LTRS-like nodes could be implemented at other sites.

Thus each participating facility was able to initiate and maintain its own technical report server, and all that was needed was a way to provide a level of integrated access. The NTRS home page was developed as a common gateway interface (CGI) perl script that would present users with a single page from which to perform a unified search. While the individual centers maintain their own publications servers for administrative efficiency, the NTRS page provides users with a single, integrated search facility.

Overview of Related Work

Technical publication servers have developed in a number of scientific communities, most notably physics,[6] computer science,[7,8] and astronomy.[9,10] These specific scientific communities led the way in electronic document exchange, partly because the community is heavily computer oriented and the majority have access to the Internet. The aerospace community is becoming more Internet-capable with time. In addition to servicing the primary aerospace customers, NTRS provides a well-known location for secondary customers to gain access to NASA research activities for potential technology transfer.

NTRS provides a different service than common CD-ROM products. While there are many useful CD-ROM products available (commercial and government), they suffer the same distribution problems of hardcopy. Unless the CD-ROM is available on the Internet, the fact that one site has a given CD-ROM does nothing to assist a site without the CD-ROM. In economic terms, CD-ROM distribution is still a zero-sum game. NTRS is a non-sum game in that although only one copy exists, it is accessible to everyone simultaneously.

NTRS is different from previous NASA "electronic library systems" such as NELS[11] and NAM[12] in several areas. The most obvious is that NTRS is provided in the context of the World Wide Web. NTRS does not introduce a custom client program with specialized features. The other difference is that NTRS is limited to the domain of technical publications, and does not directly address non-publication information and data. While the models developed for NTRS are generally applicable, the current scope is limited to technical publications.

Software Implementation Decisions

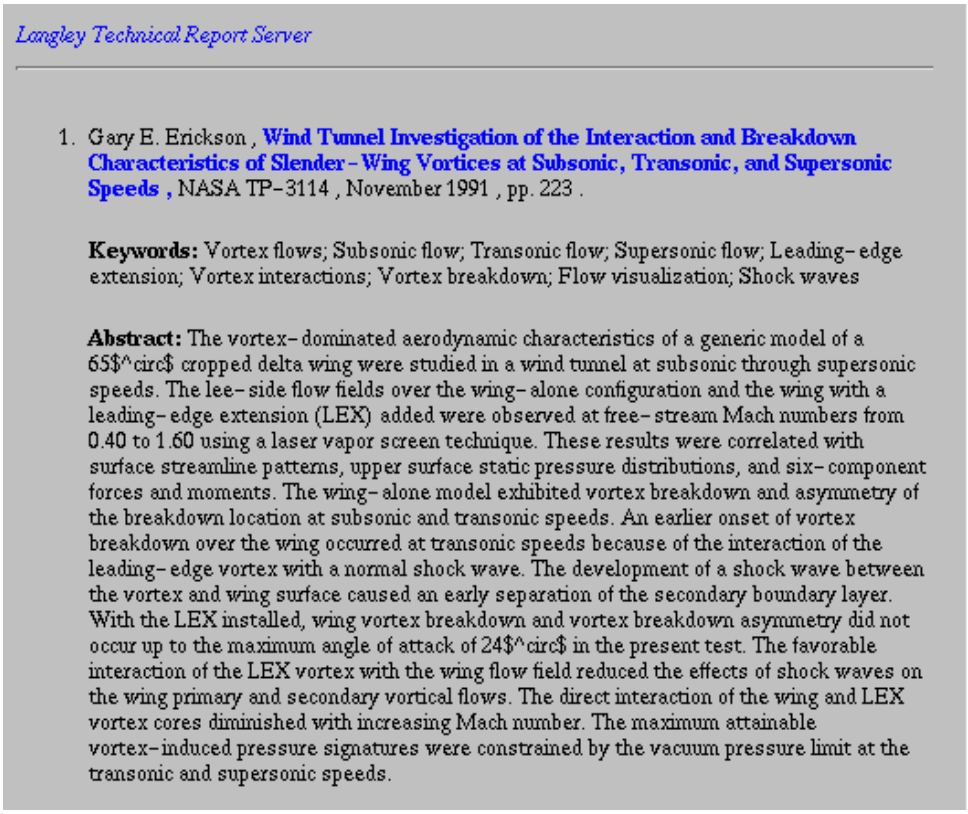
Quick access to technical information is helpful to researchers. Most technical work is built on previous work, and so in the latter half of the 20th century Hyper Text systems were invented to facilitate the intuitive "navigation" of information. The most ambitious and successful of these systems has been the World Wide Web. The Hyper Text Transfer Protocol (http) used in the WWW is efficient in enabling a distributed knowledge network, and the uniquely extensible Mosaic offers an elegant interface. Driven by an ever increasing user community, the WWW offers an ideal vehicle to publish technical data, and professional tools are quickly becoming available for search and retrieval. When published on the WWW, many publication delays will disappear, researchers will be able to add direct cross references in the form of Hyper Text links, and the speed, increasing economy, and convenience of publication and the associated telecommunications technology will facilitate a whole new generation of specialized electronic publications. This process is already beginning. The NTRS is an effort to add value and improve the interface and usability of NASA's on-line technical data publication.

The Wide Area Information Server (WAIS)[13] software was chosen to implement searching for the components of NTRS for several reasons. First, WAIS is freely available. Second, WAIS is a simple, generic search engine that is supported by many different platforms and integrates into a WWW environment easily. The source code is available, allowing site customization where necessary. While WAIS is an attractive implementation for the NTRS searching requirement, NTRS does not directly depend on WAIS. Should a successor to WAIS appear, it would not be difficult to replace WAIS transparently with another search engine, and the user should notice little, if any, difference. The citations for the publications in the various servers are archived in a format different from their presentation. For example, most of the citations are stored in "refer" format, even though the user only sees a properly formatted HTML citation (Fig. 3). Separating the archival format from the presentation format will allow for easier transition to other successor systems and the sharing of data with other non-HTML or even non-WWW systems. Refer format was chosen because it is simple and it is easy to write various translators for refer -> HTML, etc. There is no reason a richer archival format cannot be used, such as bibtex[14], RFC-1357[15] or even MARC.

Tools and Customizations Resulting from NTRS

A number of interesting tools, programs, and methodologies were employed to create NTRS. Where certain tools were lacking, enhancements were developed. Significant pieces include:

- NTRS home page - a perl CGI script
- On-line user-feedback form (perl script)
- CERN WAIS-HTTP gateway 3.0
- NCSA httpd 1.3
- Modified "waisindex" program
- freeWAIS 0.202 server
- Modified refer -> HTML converter (perl script)



Langley Technical Report Server

1. Gary E. Erickson , **Wind Tunnel Investigation of the Interaction and Breakdown Characteristics of Slender-Wing Vortices at Subsonic, Transonic, and Supersonic Speeds ,** NASA TP-3114 , November 1991 , pp. 223 .

Keywords: Vortex flows; Subsonic flow; Transonic flow; Supersonic flow; Leading-edge extension; Vortex interactions; Vortex breakdown; Flow visualization; Shock waves

Abstract: The vortex-dominated aerodynamic characteristics of a generic model of a 65^circ cropped delta wing were studied in a wind tunnel at subsonic through supersonic speeds. The lee-side flow fields over the wing-alone configuration and the wing with a leading-edge extension (LEX) added were observed at free-stream Mach numbers from 0.40 to 1.60 using a laser vapor screen technique. These results were correlated with surface streamline patterns, upper surface static pressure distributions, and six-component forces and moments. The wing-alone model exhibited vortex breakdown and asymmetry of the breakdown location at subsonic and transonic speeds. An earlier onset of vortex breakdown over the wing occurred at transonic speeds because of the interaction of the leading-edge vortex with a normal shock wave. The development of a shock wave between the vortex and wing surface caused an early separation of the secondary boundary layer. With the LEX installed, wing vortex breakdown and vortex breakdown asymmetry did not occur up to the maximum angle of attack of 24^circ in the present test. The favorable interaction of the LEX vortex with the wing flow field reduced the effects of shock waves on the wing primary and secondary vortical flows. The direct interaction of the wing and LEX vortex cores diminished with increasing Mach number. The maximum attainable vortex-induced pressure signatures were constrained by the vacuum pressure limit at the transonic and supersonic speeds.

Fig.3 Example of formatted citation.

Many of these products have been shared with others outside of NASA interested in similar technologies, and others are constantly being developed. While some participants of NTRS may elect to eventually use commercial versions of some software components, there is no requirement to do so; it is possible to be a fully functioning node constructed entirely with freely available software.

Hardware Employed for NTRS

While NTRS can be used by any machine supporting WWW, most machines serving NTRS information are UNIX workstations, many of which are non-dedicated. Additionally, while the index of publications for a given site must be centralized, some sites take advantage of the distributed nature of WWW and maintain full-text copies of the reports in different locations. Table 1 lists the machines that make up NTRS.

The following sections discuss the contents and direction of the various database components of NTRS. Each database is managed locally at its respective site. Sensitive, proprietary, or classified information is not made available through NTRS.

ADS Astrophysics Science Information and Abstract Service

The Astrophysics Science Information and Abstract Service (ASIAS) of the NASA Astrophysics Data System, formerly known as “Abstract Service,” has been very successful in providing researchers and librarians the capability to search astronomical literature. It currently provides access to more than 160,000 astronomical abstracts with a sophisticated search engine.

Usage of the service dramatically increased after it was made available on the WWW, and now averages 30,000 queries and 500,000 retrieved abstracts per month. Its ease of use, flexibility, and data coverage have made it a well known resource in the astronomical community. A WAIS interface to the database was made available in September 1994, and as a result, the service has been integrated into the NTRS.

Queries

The abstract server allows users to specify queries on separate “fields” in the documents. For instance, by defining separate fields for abstract text and paper title, a

query can specify a term to be searched for only in the title but not in the abstract text. Complex queries are composed by searching for different terms in some of the fields and then combining the results according to their relevance with respect to the original query. All query terms are case-insensitive and are compared to a list of synonyms before the term is searched in the database. Query results are ranked by a score determining how they match the input query.

The searchable fields in the database are:

au: author
ti: title
kw: NASA/STI keywords
bc: bibliographic code
at: abstract text

A field search consists of the two letter keyword associated with a particular field, followed by the term(s) to be searched for. If the search text consists of more than one word, it must be enclosed in parentheses. For example:

au = smith

searches for “Smith” in the author index, while

ti = (galaxy cluster)

searches for “galaxy” or “cluster” in the title index.

Search terms may contain (nested) Boolean expressions (specified using the Boolean operators AND, OR, and NOT), partial words (specified by appending an asterisk (’*’) to the root of the search term), and literal expressions (specified by quoting them with a single (‘) or double quote (“)).

Some examples are:

‘black hole’	literal query: the expression “black hole” is considered as a single search term.
m31 AND ‘black hole’	Boolean query: both “m31” and “black hole” must be found in the database.
au=stain*	partial word query: look for authors whose names start with the word “stain”.

Combinations of field queries are allowed:

ti=m31 ’black hole’	find documents having either “m31” in title or “black hole” in them.
ti=m31 AND ‘black hole’	find documents having “m31” in title and “black hole” in them.
ti=(m31 AND ‘black hole’)	find documents having both “m31” and “black hole” in their title.
ti=m31 AND at=’black hole’	find documents having “m31” in the title and “black hole” in the abstract body.

Dryden Technical Report Server

The Dryden Technical Report Server (DTRS) database makes available two sets of information. First, all unrestricted, unclassified, unlimited distribution technical papers written since June 1994 are available on-line in PostScript® form for download via WWW. FTP and gopher access to the on-line papers are currently under development.

Second, the searchable database for Dryden technical papers is being developed to retrieve papers written before June 1994. This will provide another way to find older papers usually obtained via the more traditional hardcopy order methods. Currently, information back to 1979 is on-line and plans are to populate the database with the complete set of bibliographic citations on all technical papers generated at Dryden. This task will be pursued as time permits.

The data files for the technical papers are maintained in directories by year of publication. Every file uses the unique paper number issued to that publication (H-####). Dryden’s “H” number was instituted during the years it was known as the High-Speed Research Station.

New bibliographic citation files and technical papers are periodically added to the DTRS. A WWW CGI program was written to allow the editing and creation of citation files for the database. These “refer” format files generate the HTML files that build the WAIS database for searching. Since the WAIS database is not totally available during a database rebuild, a UNIX cron job is run at midnight each Sunday to re-index the database. This helps minimize the downtime of the service and the impact to the users.

Goddard Institute for Space Studies Publications On-Line

The publications server at the Goddard Institute for Space Studies (GISS) contains abstracts for approximately 360 abstracts of publications (co)written by GISS staff members during the past 20 years. Copies of 16 of the papers are also available on-line in compressed PostScript format. The majority of the documents included in the GPOL database are papers which have appeared in peer-reviewed journals, although documents such as student doctoral theses and NASA conference publications are also included.

Mechanically, papers in GPOL are titled using an NN.XXX method, where NN is the year of publication and XXX is the last name(s) of the (first two) author(s). Using author names in file names provides pre-performed alphabetic sorting when a perl script is run to convert the "refer" files in the database to the HTML indices which users actually see when they access GPOL via a Web browser. In addition, papers can easily be found if corrections or changes are necessary.

At present, the most pressing concern with GPOL is the limited the number of papers available on-line. Due to copyright issues, there is still some debate whether papers written by a significant fraction of GISS personnel may be included, as these persons are actually research faculty or post-docs employed by nearby Columbia University. An effort is currently being made to aggressively solicit additional on-line papers in conjunction with assembly of an annual hardcopy volume of GISS research publications abstracts.

A lesser issue to be addressed is "in-house" handling of users who wish to search the GPOL database. At present, this has been enabled by using an "off-site" WAIS engine at Langley Research Center, which requires that LTRS periodically copy the entire GPOL database. However, it is expected that an in-house WAIS engine will be installed in early 1995.

The Institute for Computer Applications in Science and Engineering Technical Reports Server

The Institute for Computer Applications in Science and Engineering (ICASE) is a center of research in applied mathematics, numerical analysis, fluid dynamics, and computer science, providing a natural mechanism for interactions among NASA scientists and engineers, the ICASE staff, and the wider community in universities and related industries. Therefore, its main product is research and subsequently the papers that are produced as the result of this research. For this reason, ICASE chose to participate in the NTRS project by supplying the ICASE Technical Report Server (ITRS).

The ITRS follows the paradigm of the NTRS and uses the same basic tools. Abstract lists are available for approximately half of the technical reports generated in 1993 and all of the reports generated in 1994. Ninety percent of the 1994 documents are available as compressed PostScript documents and can, therefore, be instantly retrieved by the user. If an on-line version of the document is not available, there is an electronic form provided for the user to request a hardcopy to be sent to him/her. Only about half of the 1993 technical reports and abstracts are available on-line due to the difficulties of acquiring an electronic version of the paper that can be converted to PostScript. Ideally, all future ICASE technical reports will be available on-line via the ITRS.

Langley Technical Report Server

The LTRS contains all "unclassified, unlimited" formal (4-digit) technical papers and technical memorandums since 1992. Some 1991 formal reports are available. Authors have contributed conference papers, high-number (6-digit) technical memorandums, LaRC-sponsored contractor reports, journal articles, and LaRC-sponsored theses to LTRS, dating as far back as 1986. The author-contributed reports currently account for roughly 30 percent of LTRS contents. Since there is currently no mandatory requirement for LaRC authors to participate, only a subset of LaRC's research activities are present. Increased participation is anticipated in the near future.

Some of the reports do not contain all the original graphics, though most have at least some. There are no statistics on the level of completeness for the reports. Generally, older reports are less likely to be totally electronic. Most reports are PostScript, and are compressed with the UNIX compress utility. Some reports are in HTML, if the authors have made the conversion.

Numerical Aerodynamic Simulation Facility Database

In the first quarter of 1994, several people at the Numerical Aerodynamic Simulation (NAS) facility at NASA Ames Research Center began publishing branch technical reports via the WWW. After the publication of the first few reports went well, NAS management decided to use the World Wide Web as the primary publication vector for NAS technical reports.

The initial attempt far exceeded expectations. This success lead NAS to stop printing paper reports, which resulted in a several orders of magnitude increase in total reports disseminated. There was also an enormous subjective improvement in the ease of using the technical reports as references, largely due to the ease of searching the reports with NTRS.

Based on initial success with the technical reports server, NAS is currently undertaking projects to develop new procedures to index and eventually publish other forms of data, including scientist's research notes, datasets, software, and training videos. It is difficult to estimate a dollar amount saved by switching to electronic publication, but the consensus at NAS is that WWW publication of technical reports has been highly successful.

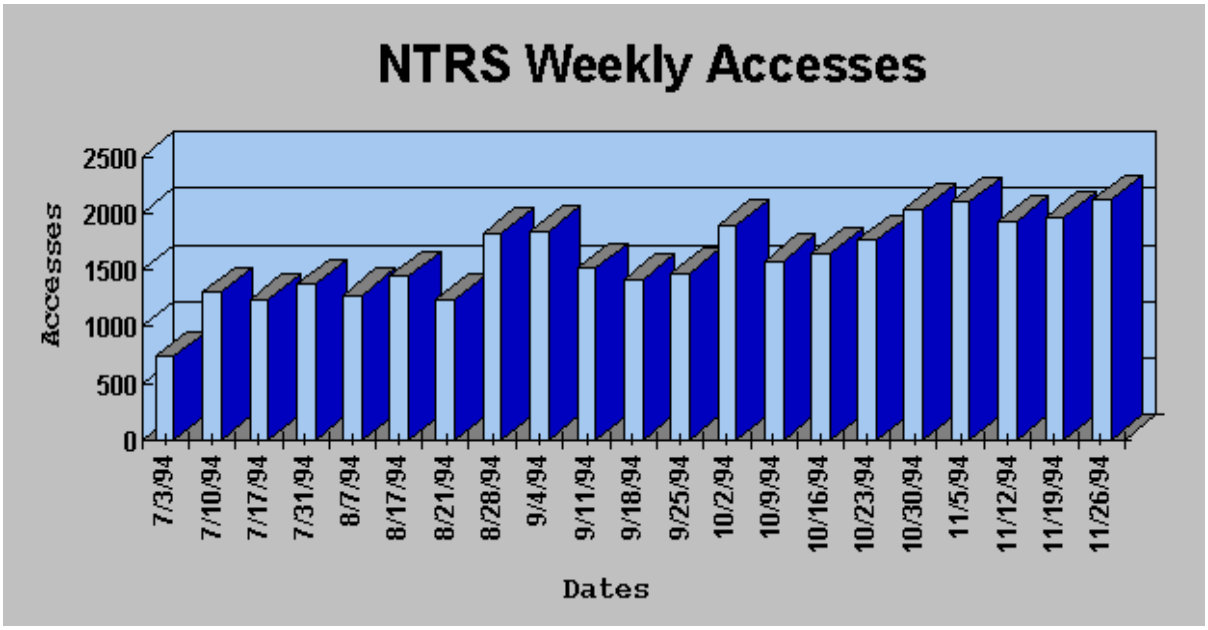


Fig. 4 NTRS weekly accesses.

RECON, SCAN, & STELAR

NTRS employs “database reuse” as well. The experimental WAIS subset of RECON and the SCAN current awareness notes, both offered by the STI program and the publicly accessible version of STELAR, are also accessible via NTRS. RECON is a large, all-encompassing collection of aerospace related references, and the years 1990-1994 are offered experimentally in WAIS. SCAN is a current awareness service, whose contents are constantly updated. STELAR is an early experiment in providing access to astronomical abstracts with WAIS. These databases were on-line before the popularity of WWW and Mosaic. They offer WAIS access, but the data are not in HTML and do not offer the option of downloading a publication. Despite this, they make a useful addition to NTRS. This is especially true of RECON, which has the breadth to provide search results for most topics.

NTRS Access Statistics

The NTRS has been well received by the WWW user community. During its initial 6-month operating period, (June 6, 1994 through December 6, 1994) the NTRS had logged nearly 37,000 server connections and 18,500 database searches were performed. Assuming an average of 1 user per computer, the NTRS has had nearly 10,000 customers.

NTRS accesses have been steadily increasing since June 6, 1994. Figure 4 is a chart of the NTRS weekly accesses. The average weekly accesses have increased from 1,100 in July, 1994 to over 2,300 in December, 1994.

The following table shows a distribution of where the NTRS accesses originate by domain. This table is not complete. It shows the broad base of world-wide interest in the information that NASA has made available.

Table 3. Example of NTRS accesses by domain.

Domain	Total accesses	Percent of total accesses
Argentina	40	
Australia	87	2.4
Austria	110	0.7
Belgium	218	
Brazil	39	
Canada	1025	2.8
Chile	20	
China	4	
Colombia	12	
Costa Rica	2	
Croatia	11	
Czech Republic	7	
Denmark	170	0.5
Egypt	7	
Estonia	3	
Finland	303	0.8
France	939	2.5
Germany	1992	5.4
Great Britain	49	

Table 3. Concluded.

Domain	Total accesses	Percent
Greece	30	
Hong Kong	25	
Hungary	17	
Iceland	12	
Indonesia	9	
International	10	
Ireland	67	
Israel	41	
Italy	460	1.2
Japan	577	1.6
Korea (South)	234	0.6
Luxembourg	3	
Malaysia	1	
Mexico	43	
Netherlands	656	1.8
Network	604	1.6
New Zealand	28	
Non-Profit	205	0.6
Norway	453	1.2
Peru	2	
Philippines	3	
Poland	36	
Portugal	53	
Singapore	37	
Slovak Republic	15	
Slovenia	23	
South Africa	77	
Soviet Union	28	
Spain	144	0.4
Sweden	586	1.5
Switzerland	319	0.9
Taiwan	22	
Thailand	17	
Turkey	17	
US Commercial	4817	13.0
US Educational	7242	20.0
US Government	1370	3.7*
US Military	1001	2.7
US NASA	3607	9.7
United Kingdom	2636	7.1
unresolved	5025	14.0

*Note: Not including NASA

User Feedback

A simple on-line form is provided for users to give feedback about the NTRS. A sampling of user feedback can be found at:

http://techreports.larc.nasa.gov/ntrs/feedback/

To date, feedback generally falls in the categories of: (1) compliments and suggestions, (2) syntax questions (i.e., how to correctly invoke Boolean searching, etc.), and (3) how to obtain an electronic or hardcopy of a paper where only the abstract is provided in NTRS.

Cost Savings

Use of NTRS results in direct savings over the traditional method of report distribution. Printing and binding a report is estimated to cost $1.75. The estimated cost is based upon an average report length of 25 pages, 100 copy print runs, $0.04 per page duplication, labor at $15.00/hour, and paper at $2.00/ream. There is an additional cost for postage and handling estimated at $1.00 per report.

Using the raw numbers of reports distributed through the NTRS, the cost of distribution of 29,500 report copies at $2.75 per report is over $81,000.00. All NASA facilities are already networked, so the only direct cost to NASA is the increased bandwidth usage of the campus network. It is interesting to note that NAS has tentatively ceased paper production of the NAS report series and the on-line versions are the only method of distribution.

A more useful metric of the NTRS is in time savings to the customer. Using NTRS, downloading a report takes no longer than a few minutes, and depending on the printer speed, it should take no longer than 1-2 hours to print. In contrast, it can currently take 2-4 weeks to request and receive a paper copy of a report, based upon researchers' experience at NASA Dryden FRC. When this time delay is multiplied over 29,500 reports, the time savings to the customers exceeds 1000 years.

These numbers assume the 100-percent usage of all reports distributed, which is definitely not correct. But if even a significant fraction of the service is used to aid research, the payback in time and cost savings is tremendous, with the largest potential savings by NASA customers. Use of the NTRS will continue to increase as the databases are populated with increasing amounts of information.

Lessons Learned

Implementing NTRS taught or reinforced many lessons about Internet-based information systems. Some of the observations are shared in the following:

- The most frequent problems encountered were incorrectly configured machines, and the somewhat unstable nature of early WWW software. These problems were especially pronounced for PC and Macintosh platforms. This will be resolved in the near future with availability of more robust, commercial-grade WWW browsers.

- The search syntax (or lack of) was not immediately intuitive to some users. The free text searching of WAIS would sometimes puzzle users that were expecting something more complicated.

- If the document that a user wanted was not on-line, it did not make them feel better knowing that hundreds of other documents were available. More and more, users expect all information to be on-line. It is anticipated that services which can meet these expectations will be the ones users come to depend on. It is possible that those collections that do not have an upgrade path from paper or microfiche to electronic stand a good chance of being "forgotten" by on-line researchers.

- The feedback from NTRS (and other WWW services) indicate that the public (those on-line, at least) have a real interest in obtaining NASA information. Some NTRS users were looking for specific reports, and some were searching for NASA activity in general subject areas.

- The previous revelation also highlights another lesson: the public considers NASA information difficult to obtain. Users report that many libraries do not receive the information in a timely fashion if at all. Other users expect NASA to deliver information in a media comfortable to them (on-line) and do not consider traditional distribution methods (NTIS, CASI, libraries) acceptable.

- Interestingly, some user feedback ranges from surprise that NASA is providing this information to impatience as to why all of NASA's information is not available via WWW.

- ADS has implemented the optimal WAIS solution. To better serve the astronomical community, they provided a WAIS interface to their existing abstracts database. After testing several packages currently available in the public domain, they selected a variation of the Clearinghouse for Networked Information, Discovery, and Retrieval (CNIDR's) freeWAIS called freeWAIS-sf[16] since they found it to be the most advanced at this time; in particular, it is superior to the CNIDR freeWAIS version in that it introduces the concepts of structured fields in the document that can be searched separately.

Some time was spent in enhancing the freeWAIS code to run faster by storing some frequently accessed data in shared memory, to allow better control of what words would be ignored when indexing, and to support extended headlines (the document identifier strings returned by the server upon completion of a query). These changes have since been incorporated in the freeWAIS code.

Thanks to the efforts of the author of freeWAIS-sf, typical bottlenecks in the indexing of source documents (hard-coded limits about the size of headlines, inverted indexes, etc.) have been either eliminated or isolated. These enhancements over freeWAIS make freeWAIS-sf the best public-domain, general purpose, full-text indexing and search engine available today.

Future Implementation Plans

The most pressing issue for NTRS is the inclusion of other NASA centers and institutions, especially the remaining NASA research centers. Of these, Lewis Research Center has a working prototype for LeTRS. Ames Research Center, Goddard Space Flight Center, and Kennedy Space Center have preliminary prototypes of ATRS, GTRS, and KTRS (respectively). Marshall Space Flight Center and the Jet Propulsion Laboratory have also expressed interest in participating. There may be a need for more sophisticated searching capabilities, possibly even a new search engine. Perhaps two different interfaces to NTRS, "beginner" and "expert", are needed to service both groups of users.

Other issues to be addressed include (1) increasing the number of reports available, (2) ensuring that all present and future reports are stored electronically, (3) converting older archives and completing partially electronic reports (e.g., making all figures electronic), and (4) presenting a choice of formats for the user (PostScript, HTML, native word processing format if applicable, etc.).

The ADS plans to expand its abstract service to include scanned images of full-text articles from selected journals. As a test case, they are scanning all issues of the Astrophysical Journal (Letters) from 1975 to date. If an article is available for a selected abstract, links to the bitmaps (stored in PostScript format) will be present in the returned report.

Conclusion

The NASA Technical Report Server is an experimental, grass-roots project with the goal of providing maximum connectivity to existing electronic publications and publications servers. Using World Wide Web and other Internet information systems, NTRS has achieved great initial success in providing access to NASA research publications to

the world-wide scientific and research community. The WWW also allows the implementation of NTRS to be logically central and physically distributed. All participating institutions maintain their own servers, and the NTRS WWW page provides a single access point for the entire collection. NTRS demonstrates that significant results in technology awareness and distribution can be achieved with minimal resource investment using the latest in information technology.

Acknowledgments

The authors would like to thank James Gass (GSFC) for his work w/ STELAR and the NASA Web in general; Rick Dunbar and Gregory Decker (NASA STI) for their work with RECON and SCAN; Michael Grabenstein (GSFC), Steve Eubanks and Omar Syed (LeRC) for working on servers for their respective centers; and Richard Tuey (NASA STI) for organizational support.

References

[1]Berners-Lee, Tim, Cailliau, Robert, Groff, Jean-Francois, and Pollermann Bernd, "World-Wide Web: The Information Universe," Electronic Networking: Research, Applications and Policy, vol. 2, no. 1, 1992, pp. 52–58.

[2]Andreessen, Marc and Bina, Eric, "NCSA Mosaic: A Global Hypermedia System," Internet Research: Electronic Networking Applications and Policy, vol. 4, no. 1, 1994, pp. 7–17.

[3]Smith, Ruth S., "The NACA Collection: When Old is New Again - A Status Report on the NACA Documents Project," Sci-Tech News, vol. 46, no. 2, Apr. 1992, pp. 15–20.

[4]Nelson, Michael L. and Gottlich, Gretchen L., "Electronic Document Distribution: Design of the Anonymous FTP Langley Technical Report Server," NASA TM-4567, Mar. 1994.

[5]Nelson, Michael L., Gottlich, Gretchen L., and Bianco, David J., "World Wide Web Implementation of the Langley Technical Report Server," NASA TM-109162, Sept. 1994.

[6]Ginsparg, P., "First Steps Towards Electronic Research Communication," Computers in Physics, vol. 8, no. 4, Jul/Aug. 1994, pp. 390–396.

[7]Maly, K., French, J., Selman, A., and Fox, E., "Wide Area Technical Report Service," TR_94_13, June 6, 1994. Old Dominion University, Dept. of Computer Science.

[8]Davis, James R. and Lagoze, Carl, "A Protocol and Server for a Distributed Digital Technical Report Library," Cornell University Technical Report TR94-1418, June 1994.

[9]Accomazzi, Alberto, Grant, Carolyn Stern, Eichhorn, Guenther, Kurtz, Michael, and Murray, Stephen, "ADS Abstract Service Enhancements," Proceedings of the Fourth Annual Conference on Astronomical Data Analysis Software and Systems, H. Payne, D. Shaw, J. Hayes (ed), 1995. Publications of the Astronomical Society of the Pacific series.

[10]Van Steenberg, M. E., Gass, J., Brotzman, L., Warnock, A., Kovalsky, D., and Giovane, F., "STELAR: An Experiment in the Electronic Distribution of Astronomical Literature," American Astronomical Society Newsletter, No. 67, Oct. 1992.

[11]Smith, Stephanie L., "NELS 2.0: A General System for Enterprise Wide Information Management," Presented at AIAA Computing in Aerospace 9, San Diego, CA, 1993, pp. 1358–1365. (Available as AIAA 93-4706).

[12]Hunter, Judy, Generous, Curtis, and Duncan, Denise, "NASA Access Mechanism - Graphical User Interface Information Retrieval System," Jan. 1993, AIAA 93-0583.

[13]Kahle, B., Morris, H., Davis, F., Tiene, K., Hart, C., and Palmer, R., "Wide Area Information Servers: An Executive Information System for Unstructured Files," Electronic Networking: Research, Applications and Policy, vol. 2, no. 1, 1992, pp. 59-68.

[14]Lamport, Leslie, "LaTeX: A Document Preparation System," Addison-Wesley, 1986.

[15]Cohen, Dan, "A Format for E-mailing Bibliographic Records," USC - Information Sciences Institute, Internet RFC-1357, July 1992.

[16]Pfeifer, Ulrich and Huynh, Thung, "FreeWAIS-sf," Internal report, University of Dortmund, 1994.

AIAA 95-0964

The Widest Practicable Dissemination: The NASA Technical Report Server

Michael L. Nelson, Gretchen L. Gottlich
NASA Langley Research Center, Hampton, Virginia

David J. Bianco
Computer Sciences Corporation/NASA Langley Research Center, Hampton, Virginia

Robert L. Binkley, Yvonne D. Kellogg
NASA Dryden Flight Research Center, Edwards, California

Sharon S. Paulson
Institute for Computer Applications in Science and Engineering, Hampton, Virginia

Chris J. Beaumont
Computer Sciences Corporation/NASA Ames Research Center, Moffett Field, California

Robert B. Schmunk
NASA Goddard Institute of Space Studies/
National Research Council, New York, New York

Michael J. Kurtz, Alberto Accomazzi
Smithsonian Astrophysical Observatory, Cambridge, Massachusetts

Computers in Aerospace 10
March 28-30, 1995/San Antonio, TX