CR–206873

SPACE TELESCOPE SCIENCE INSTITUTE

3700 San Martin Drive
Baltimore, MD 21218
(410) 338-4700
FAX (410) 338-4211

Letter Number: M98-0017
February 17, 1998

National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland 20771

Attention: Mr. Michael E. McGrath
Contracting Officer
Mail Code 216.0

Subject: NASA Contract NAS5-32496 - "Schmidt Plate Astrometry and the Subplate Overlap Code" and "Astronomical Software Directory Service" Studies

Reference: (a) ST ScI Internal Accounting Contract Number C001.34300

Pursuant to the contractual requirements contained in the referenced contract, the Association of Universities for Research in Astronomy, Inc. operating the Space Telescope Science Institute (ST ScI) hereby provides the enclosed final technical report and the final government property report for equipment procured in support of this effort.

Should you have questions or require any additional information, please contact me at (410) 338-4208 or by E-mail at "mwilson@stsci.edu".

Sincerely,

Monica S. Wilson

Monica S. Wilson
Senior Contracts Administrator

Enclosures

cc: Dr. Donald West, COTR - Mail Code 684 (1 Copy)
Publications and Graphics Services Section - Mail Code 253.1 (1 Copy)
Center for Aerospace Information - Linthicum Heights, Maryland (2 Copies)

Final Technical Report:
Astronomical Software Directory Service

Contract No. NAS5–32496
Funded by NRA–OSSA–92–15
Astrophysics Data Program

Submitted by Robert J. Hanisch, Principal Investigator
Space Telescope Science Institute
3700 San Martin Drive
Baltimore, MD 21218
hanisch@stsci.edu

Co-Investigators:
Harry Payne
Jeffrey Hayes
Space Telescope Science Institute

Consultant:
Archibald Warnock, A/WWW Enterprises

With the support of NASA's Astrophysics Data Program (NRA 92-OSSA-15), we have developed the Astronomical Software Directory Service (ASDS): a distributed, searchable, WWW-based database of software packages and their related documentation. ASDS provides integrated access to 56 astronomical software packages, with more than 16,000 URLs indexed for full-text searching. Users are performing about 400 searches per month. A new aspect of our service is the inclusion of telescope and instrumentation manuals, which prompted us to change the name to the Astronomical Software and Documentation Service. We have already incorporated 68 on-line manuals from 23 sites, with about 6500 URLs indexed for searching. Our home page, shown in Figure 1, is available on the WWW at **`http://asds.stsci.edu/asds/`**.

ASDS was originally conceived to serve two purposes: to provide a useful Internet service in an area of expertise of the investigators (astronomical software), and as a research project to investigate various architectures for searching through a set of documents distributed across the Internet. Two of the co-investigators were then installing and maintaining astronomical software as their primary job responsibility. We felt that a service which incorporated our experience in this area would be more useful than a straightforward listing of software packages. The original concept was for a service based on the client/server model, which would function as a directory/referral service rather than as an archive. This concept predated the emergence of the World Wide Web (WWW). Once we the WWW came into existence, we quickly settled on the idea of a Web site devoted to searching through documents associated with astronomical software.

Our original concept of how our site would be used was that astronomers, or astronomical programmers, having some idea of what they needed in the way of astronomical software, would perform full text searches of our document collection, with the idea that relevant documents would come from packages deserving the user's attention. Once interesting packages had been identified, users could then go to our package descriptions to find out how to get the packages, what resources were required to run them, and perhaps what our experience was when we installed them. We created a standardized template for describing the software packages, with fields describing many of the general features of each package, ensuring that each package in our collection had a "package description" Web page. The other documents in our collection include user guides, programmer guides, on-line help files, UNIX man pages, and custom Web interfaces to software documentation—basically, anything available on-line in HTML format, or some other format suitable for full-text searches. But rather than collecting all of these documents into a centralized archive that we maintain, we chose to index the documents maintained by their authors. The results of a search include links to the original documents, so that users retrieve documents from the host sites. This insures that users always get the most up-to-date documentation.

For performing the searches, we began our investigation with a decision to evaluate the Isite software from the Center for Networked Information Discovery and Retrieval (CNIDR). This software was intended as a replacement for WAIS (Wide-Area Information Service), a client/server technology for performing full-text searches through a set of documents. Isite had some additional features that we considered attractive, and we enjoyed the cooperation of the Isite developers, who were happy to have ASDS as a demonstration project. We ended up staying with the software throughout the project, making modifications to take advantage of new features as they came along, as well as influencing the software development. The Web interface to the search engine is provided by a gateway program written in C++ by a consultant to the project (A. Warnock).

The key modification to the Isite software that made it applicable to our project was allowing it to open a WWW URL for indexing and searching—the original software, like WAIS, could only handle files in the host's own file system. Most of what made ASDS a research project had to do with the way in which this modification was implemented. The original version of Isite did not store information in the index if it

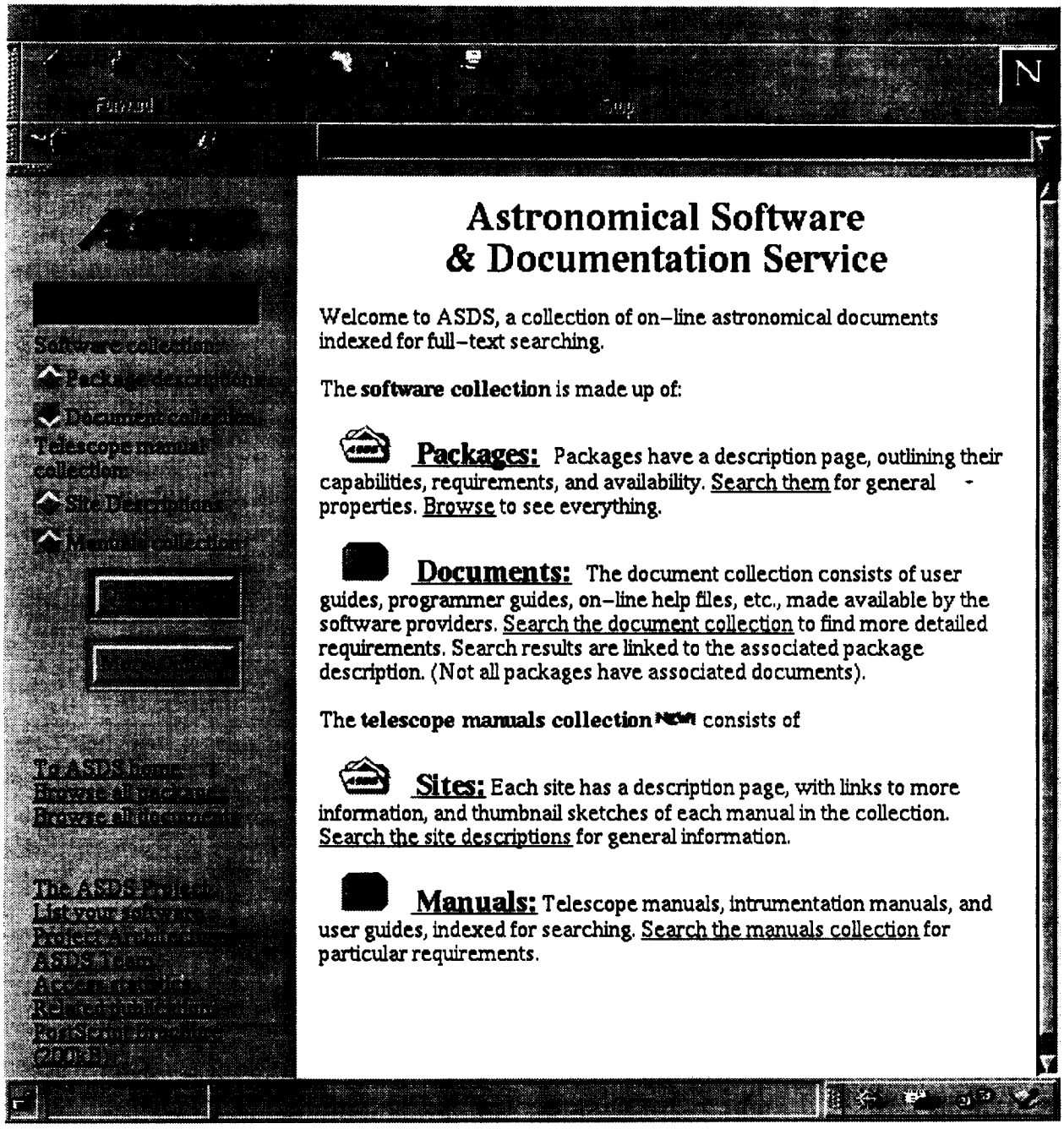

Figure 1: The ASDS home page.

could be retrieved at search time from the documents themselves. Similarly, the original implementation of indexing URLs required the search engine to go out on the Internet at search time to retrieve the information not stored in the index. This made our system similar to those systems that query distributed servers by sending a query to each one, and collecting the results as they come back, i.e., slow. If remote systems were down, or if a software author changed a document, then searches would fail. In subsequent iterations, more information has been in the index, eliminating network traffic at search time. The current system is robust against changes in the documents, although the index is updated regularly to track authors' changes.

At the end of the project, we have package descriptions for 56 packages, grouped into seven functional categories:

- Data reduction and analysis,
- Document preparation,
- Graphics, plotting, image display, and visualization,
- Modelling and simulation,
- Math and statistics,
- Subroutine libraries, and
- Utility programs.

Of these 56 packages, 41 have an associated searchable index of on-line documents. A total of 16,000 URLs have been indexed. It takes the search engine less than a second to perform a simple search of all 16,000 documents. Through the Web interface, the search time is completely dominated by the time it takes for the user to transmit the query to us, and for us to transmit the results back to the user.

Most of the effort invested in ASDS has been devoted to three areas. First, as with most projects dating back to the start of the Web, we did a lot of experimenting with user interface issues: arranging the content, navigating the site, establishing convenient links between document searches and the corresponding package descriptions, and establishing a balance between searching and browsing. One of the last tasks on this project was giving the site a finished look and feel. The second major effort was in automating our Web site maintenance. Almost all of the Web pages from our site are generated by Perl scripts, either on-the-fly for each user, or saved periodically into static pages as part of routine maintenance. After a new package or document has been added to our collection, the entire Web site can be rebuilt with a simple UNIX "make" command. These two aspects of the project have been quite successful, and will have continued benefits in allowing ASDS to be maintained at relatively low cost after the end of the project, and in future Web services being developed by the investigators. The third major task was adding content to the system: searching the Web for software, searching through Web sites for the information required to complete a package description page, and collecting URLs of pages to be indexed.

We recently extended the ASDS architecture to the subject of telescope manuals, extending the package description/software manuals paradigm to a site description/telescope manuals paradigm, as part of a project to index more of the astronomical grey literature. This project will produce Web forms to facilitate maintenance of the ASDS Web site, reducing the cost of adding new content.

The ASDS remains available, and will continue to be available as resources allow. We plan to continue to add packages and manuals, and to respond to updates in the Isite software. We are considering options for further enhancing the service, and broadening its scope, via other funding sources.

Space Telescope Science Institute
Property 1018 Receive Report by Contract-Grant number
as of 02.13.98
for dollar values greater than 0.00 and less than 1000000.00

Contract/Grant	Tag No	Description	Model Number	Location	Date	PO	Serial Number	Unit Cost	Trans Code	Current Custodian
NAS5-32496	GV004333	DUAL ENCLOSURE W/TWO 4.29 GB DRIVES	AD5230/E-STL	115	12/21/95	17924	36607-5511	2,165.00	R	CLARKE, ANGIE
NAS5-32496	GV004334	DUAL ENCLOSURE W/TWO 4.29 GB DRIVES	AD5230/E-STL	115	12/21/95	17924	36607-5510	2,165.00	R	CLARKE, ANGIE
**								---		
NAS5-32496								4,330.00		
								==========		
								4,330.00		

2 records listed. 2 values listed.