DEVELOPMENT AND OPERATION OF THE ASTROPHYSICS DATA SYSTEM

NASA Cooperative Agreement NCCW-0024

Final Report

For the Period 1 October 1992 through 31 March 1997

Principal Investigator
Dr. Stephen S. Murray

June 1997

Prepared for:

National Aeronautics and Space Administration
Headquarters
Washington, DC 20546

Smithsonian Institution
Astrophysical Observatory
Cambridge, Massachusetts 02138

The Smithsonian Astrophysical Observatory
is a member of the
Harvard-Smithsonian Center for Astrophysics

Contents

1 Introduction

The Astrophysics Data System (ADS) provides access to astronomical bibliographic information. including references, abstracts, and full journal articles, as well as links to other on-line information sources like on-line electronic journals and on-line data.

This section will first provide a brief history of the ADS. a general introduction of the current system. and a more detailed description of some of the parts of the ADS.

2 History

The Astrophysics Data System (ADS) provides access to astronomical bibliographic information, including references, abstracts, and full journal articles, as well as links to other on-line information sources like on-line electronic journals and on-line data.

The ADS started as a system to access the various data repositories containing data from NASA's space missions. In the late 1980s NASA asked the astronomical community for an evaluation of the status of data access (Squibb, 1987). The recommendation was that a system be established that allows access to the data at the various data centers through one system (Squibb, 1988). In 1991 an operational version of such a data access system was released (Good, 1991, Murray, 1993, Eichhorn, 1994). It was based on proprietary technology since at that time there was no public domain distributed data system available. This system allowed access to data at various collaborating data centers at the record level. It allowed users to query different databases and to retrieve data for display and further analysis.

In 1993 the abstract service was added to the ADS. It contained at the beginning about 160,000 abstracts. It allowed fielded queries with a sophisticated relevance scoring system (Kurtz, 1992, Eichhorn, 1995a)

In 1994 the emergence of the World Wide Web (WWW) made this custom built distributed data system obsolete. The open architecture of the WWW allowed the data centers to bring their data on-line easily and to provide search capabilities tailored to their particular needs.

The ADS abstract service was the first part of the ADS that was interfaced to the WWW in February 1994. A WWW catalog interface tool followed quickly (Eichhorn, 199b, Accomazzi, 1995). With the introduction of the WWW access tools to the abstract service, the usage tripled within one month. Since then the usage has been increasing almost continuously. It showed a fairly slow increase during the first year with the classic ADS interface and the rapid growth after the introduction of the WWW interface.

With the WWW technology other features became feasible, namely linking between different data systems. Very quickly the ADS made use of this technology and started linking data at different data centers to the references in which these data are described.

The first such links were from the abstract service to the SIMBAD database at the Centre de Donnees Astronomiques de Strasbourg (CDS) (Egret, 1991). Queries for astronomical objects would be sent to the SIMBAD server at the CDS. The list of references that was retrieved would then be matched against the ADS database and the matching references returned to the user. Subsequently links to the data tables at the CDS, the General Catalogue of Photometric Data, maintained at the University of Lausanne, Switzerland (Mermilliod,

1996), and the Astronomy Digital Image Library (ADIL) project at the National Center for Supercomputing Applications (NCSA) (Plante, 1996), were included where appropriate in the returned ADS references. Since then the SIMBAD project has provided the necessary information to us to include links to lists of objects that are mentioned in journal articles. This now allows our users to directly get information about objects mentioned in retrieved journal references.

With a large part of the astronomical literature references and abstracts available through the ADS, we proceeded to make full journal articles available on-line. After obtaining permission from the copyright holders of the various journals, we scanned these journals. Scanning was done at 600 dpi to ensure highest quality display and printing capabilities. This service came on-line in January of 1995 (Accomazzi, 1996). It currently provides access to eleven journals for various time spans. The near term goal is to have complete coverage for these journals from 1975 to present. The longer term goal is to provide complete coverage for all these journals (and hopefully more) from Volume 1 to present.

In order to protect the revenue of the journal, we make the latest issues available only after a grace period whose length depends on the preference of the journal. This can be as short as six months to as long as two years.

In 1996 the Astrophysical Journal Letters (ApJL) came on-line in electronic form (Boyce, 1996). More journals will be on-line in 1997. The ADS provides links to these on-line articles. The on-line journals in turn utilize the ADS abstract service to make links to abstracts available directly from their reference lists in the on-line articles. This close cooperation between the different data and information systems in Astronomy is the first example of a complete digital library that allows the user to easily move from one information source to another and to quickly find related data. The collaboration of different groups is now called Urania (Universal Research Archive of Networked Information in Astronomy).

3 Current ADS System

The ADS has two major parts, the abstract service and the on-line journal articles. This section will give an overview over the two parts and will describe some parts of the system in more detail.

3.1 ADS Abstract Service

3.1.1 General Description

The abstract service provides access to abstracts and references of most of the astronomical literature since 1975. The references are accessible in three databases: Astronomy, Space Instrumentation/Optics, and Physics/Geophysics. As of January, 1997, the astronomy database contains about 250,000 references, the Instrumentation database has almost 500,000 references, and the physics database about 250,000 references. Keeping separate databases for different topics improves the quality of the information returned from the search engine.

The databases are currently updated approximately every two weeks. In the near future a quick update capability will be implemented that will allow much more frequent partial

updates (one or more per day). Full re-indexing of the complete database will still be required every few weeks. This new capability is expected to be available in March, 1997.

Most of our data came from the NASA Scientific and Technical Information (STI) service. This project was abstracting all the astronomical literature, including journals, conference proceedings, Ph.D. Theses, NASA internal reports, etc. In 1995 NASA/STI discontinued abstracting most of the literature. Since then we are receiving abstracts or at least references (titles and author lists) from most journals directly. For the Astronomy database we receive abstracts from all major journals directly. For the Instrumentation database we receive all conference proceedings abstracts that the Society for Optical Engineering (SPIE) publishes.

The databases can be searched with a sophisticated search engine, allowing fielded searches. the retrieved references are ranked according to their relevance to the query, taking into account word frequencies for weighting.

The database is indexed such that retrieval of abstracts that contain synonyms for the search words (e.g., M31 and Andromeda) is possible. The list of synonyms was compiled manually taking into consideration expert knowledge of the technical language used in Astronomy. This synonym list has proven to be one of the most valuable assets of the ADS because it greatly improves the quality of the searches. For added flexibility, the user has the option of turning off the automatic synonym replacement.

Parsing of the abstracts before indexing and of the user input before searching allows further refinements of the search capability. For instance the words "Be Star" are indexed together when they appear in this sequence, making it possible to find references about Be Stars. Searches for the two words Be and Star separately would not give useful results.

3.1.2 Search Engine

- General

  The search engine for the abstract service was written specifically for this application. This allowed us to implement optimizations specifically tailored to this type of data. It allows us to quickly implement suggestions from our users to improve the quality, ease of use, and speed of the searches. Separate indexes for authors, title words and text words are created during indexing. Additionally, indexes for word pairs are built to enable searching for multi-word phrases. All searchable tables (index tables for each field, list of bibliographic codes, etc) are maintained in memory. When a new search is initiated, the software attaches to the tables in memory. No file accesses or loading of tables has to be done. This provides significant execution speed improvements compared with conventional databases. All tables in memory are sorted. For correlations that need to be searched in both directions, inverted index tables are held in memory also. This means that all searches can be done as binary searches.

  During a search, the database is searched for the specified words in each search field. The results are then combined according to the user specified logic. For each field, the following logic options are available:

  Combine with 'OR': A reference is returned when at least one of the words specified in the field was found.

Combine with 'AND': A reference is returned only when all words specified in the field were found.

Simple Logic: This allows the user to specify for each word individually whether it is optional for return, required, or must not be in the returned reference. Required words have to be prefixed with a '+', words that must not be present are prefixed with '-'. Words without a prefix are optional.

Complex Logic: This mode allows the user to specify full boolean expressions containing 'and', 'or', 'not', and parentheses for grouping.

For all logic options, the following constructs are recognized:

Prefix '=': This specifies that no synonym replacement should be done for this word. Normally a search is done for a word and all its synonyms.

Prefix '#': If synonym replacement has been turned of, this prefix specifies that synonym replacement should be done for this particular word.

Phrase Searching: The search engine can search for sequences of words. This allows for instance to search for "black hole" as a phrase, rather than searching for 'black' and 'hole' separately. Phrases can be specified by enclosing them in single or double quotes, or by concatenating the words with '.' or '-' (without any blanks in between).

Once the searches for each field are completed, and the results within each field are combined as per the logic selected, the results from the different fields are combined. Again, the user can determine how this is done. Normally, references are returned if at least one field shows a match. The user can select one or more of the fields to be required for the result. If this option is selected for a particular field, only references that have at least some match for this field will be returned. This allows for instance to find references by a particular author by requiring the author field to be present. The other fields would then be used to further select and sort references from this set.

- Scoring of the retrieved references

  Each selected reference is assigned a score that shows how closely this reference matches the query. By default, the score is calculated for each field from the number of items matched. For the author field and the object field, the score is the number of matched items divided by the total number of items in the search field. For the title and text fields the score is determined as the sum of the weights for each matched word divided by the sum of the weights for all words specified in the field. The weight of a word is $1/\log 10(freq)$ where freq is the frequency of the word in the database. This scoring algorithm weighs less frequent words more, since they are presumably better indicators of the relevance of an abstract to a particular subject. This frequency weighted scoring for title and text fields can be turned off if desired.

  The total score is finally determined by multiplying the score for each field with the weight for this field and adding up these products. The weights for each field can be adjusted by the user. This allows the user to for instance put more weight on the object field if this seems important. By specifying a weight of -1.0 in a particular field, this weighting scheme can be used to select against a particular field. If for instance an

4

author is specified, together with words in the text field, and the weight of the author field is set to -1.0, only references that have matches in the text field, but do NOT contain this author in their author list are returned.

The final score is then normalized by the total possible score for the query. A final score of 1 means that all items in the query were matched by this reference.

- Filters

  The final output can be further modified with several filters. These filters let the user select further attributes of the returned references. The following filters are available:

  Publication date: Selects references with a publication date in the specified range.

  Entry date: Selects references that were entered into the database after a specified date.

  Minimum score: Selects references whose score is larger than the minimum score specified.

  Bibliographic sources: Selects references from specified sources (refereed journals only, non-refereed journals only, specified journals only).

  Available data: Selects references with specified data links (e.g., full articles, or electronic on-line version, etc).

  Groups: Selects references from the bibliography of the specified group.

3.1.3 Returned Information

The result of a search is a list of references. The list contains the bibliographic code, the authors, the title, the publication date, the score and a list of links.

- Bibliographic Codes

  The bibliographic codes are a very important part of the whole digital library in Astronomy. They allow the different systems to work together by uniquely identifying every reference. They are described in detail by Schmitz et al. (1995).

  These codes are nineteen characters long. They have the following form:

  YYYYJJJJJVVVVMPPPPA

  where YYYY is the year, JJJJJ is the abbreviation for the journal (e.g., ApJ, AJ, MNRAS, Sci, PASP, etc.), VVVV is the volume number, M is used when needed to indicate "special" issues (such as "L" for Letters, "P" for pink pages), PPPP is the page number, and "A" is the first letter of the first author's surname. The fields are padded with periods (.) so that the code is always nineteen characters long. The journal is left-justified within its five characters, and the volume and page are right-justified. A list of journal abbreviations already in use is available at:

  http://adsabs.harvard.edu/abs_doc/journal_abbr.html

  These bibliographic codes can be generated automatically from a regular reference in the reference section of a journal. This allows the electronic journals for instance to generate links to abstracts in the ADS in their reference section. We provide a bibcode

verification tool that allows the journal editors to check whether the bibcodes that they generated are indeed valid bibcodes.

These bibcodes are the glue that holds the digital library together and allows different parts of the system to easily interact with other parts.

- Links to other information

  One of the most important aspects of a distributed digital library on the Web is the cross-linking between different parts of the library. Therefore the generation and maintenance of links from the ADS references to other information connected with each reference has very high priority. These links can be accessed directly from the retrieved references. They are anchored to a string of letters. The following links, anchored to the specified letters, are currently available:

TABLE 1

Link Letter	Information available at this link
A:	Abstract provided by NASA/STI.
O:	Original author abstract provided by the journal or author.
F:	Full article available through the ADS article service as scanned images.
E:	On-line electronic version of the journal article, usually at the site of the publisher. Access to the electronic version may be restricted to subscribers of the journal.
P:	PDF version of the article. This too may be restricted to subscribers of the journal.
D:	On-line data from the reference in electronic format.
S:	List of astronomical objects referred to in the article (provided by the SIMBAD database).
M:	The article can be mail-ordered from a document delivery service. This service is fee based and is handled directly through the document delivery service.
I:	Additional information about the article provided by the author.
R:	List of references for this article. It generally includes only the references to other journal articles.
C:	List of references that cite this article.
T:	Table of contents for this reference (for books and proceedings volumes).

Following is a more detailed description for some of these links.

  - Abstracts ('A' and 'O' links)

    These links provide access to the full information available in the ADS. This includes the title, author list, journal information, publication date, author affiliation, keywords, source of this information, and the full abstract (where available). the returned page also contains links to all the information mentioned in this section, as well as other relevant links (links on author names to a list of email addresses and current affiliations and the StarPages (see other section in this book), and a mechanism to use the current abstract as input to a new query to find related articles.

– References and Citations ('R' and 'C' links)

In late 1996, the America Astronomical Society (AAS) purchased a list of references from the Institute for Scientific Information (ISI). These new data allowed us to build the list of references and citations for most journal articles included in the Astronomy database. It includes only references to journal articles, not to conference proceedings. Each reference for which we have this information has a link to the set of references in that article and to the set of references that cite this article.

– On-line Electronic HTML and PDF articles ('E' and 'P' links)

These links allow our users to view on-line journals in html format and to download PDF files for printing. We are currently working closely with the AAS and University of Chicago Press, which publish the Astrophysical Journal, to provide direct links to their on-line articles. The AAS has agreed to use the same bibliographic codes (see above, and Schmitz, 1995) that we use to identify articles. This makes it very easy to provide these links automatically right after publication. Other on-line journals are using the same codes or have indicated that they will use them when they come on-line.

– On-line Data ('D' links)

One important aspect of the on-line digital library is to provide access not only to the literature, but also to on-line data. Traditionally it has been very difficult to get data from an article for further analysis. Hand typing the data was the only choice in most cases, which is error prone. The ADS is working with several data centers to provide links from the references to the data associated with these articles. This allows our users to directly access these data tables and download them in electronic format for further analysis.

– Document Delivery Services ('M' links)

These links allow our users to order specific articles from a document delivery service. All such transactions are handled directly between the user and the delivery service. Our M-link provides the information about the article to the delivery service. We currently have such an arrangement with SPIE. All their articles can be ordered through their delivery service. Our link to the SPIE delivery service contains the information for the article that is being ordered. They then return the order form for that particular article and ask the user to supply shipping and payment information. We are discussing similar arrangements with other document delivery services.

- Table of Contents

For a considerable number of journals we receive the table of contents (ToC) on a regular basis. For these journals we provide direct access to the ToCs. One of our pages provides access to either the latest published volume of each journal or the last volume that the reader has not yet read, depending on the settings of the user preferences. The last read volume for each of the major journals is remembered through the WWW cookie preference system (see next section about "Customizing the ADS").

7

This works only for users whose browsers handle WWW cookies. For the other users, this page always returns the latest published volume. There are two versions of this page available, one with images of the journals for easy journal selection, and one with plain text links for faster downloading. The ToCs for older volumes can be retrieved through a second ToC query page. That page allows the user to retrieve ToCs for selected journals by either volume or publication year and month. For older volumes there may not be all articles for a given volume in our system.

3.1.4 Customizing the ADS

Most frequently used WWW browsers now support a scheme called "Cookies." This scheme allows a server to uniquely, but anonymously, identify a user. When our server communicates with a cookie-aware browser, we set a so-called cookie. This is just a unique string that will later identify his user. The browser stores this cookie on the users local machine together with the information about which host has sent that cookie. In all subsequent transactions, the browser will include this cookie with any information that is sent to this host. This allows the ADS server to determine from which user such a request originated. We provide a form that allows our users that have cookie-handling browsers to set certain preferences in searching and returning information. These preferences are stored in a database on the ADS server. They are used to customize what information is returned and how it is returned. The preference settings form allows the user for instance to select the default resolution for printing articles and how to return them (to printer directly, to file, via fax or email). It allows the user to customize font sizes and colors for the returned pages. The user can also store a default query form that can be tailored to the needs of the particular science subjects. We intend to continue to exploit this capability to make the use of the ADS more convenient for our users.

3.1.5 Alternate Access Methods

Usually, the ADS databases are accessed through WWW forms. However, there are other possibilities for using these databases:

1. Direct Hyperlinks: Direct links to individual abstracts and articles can be embedded in html documents. These links have the form:
   `http://adsabs.harvard.edu/cgi-bin/bib_query?bibcode`
   where bibcode is the bibliographic code for the requested reference. Such a link will return the formatted abstract with links to all available information.

   Direct links to journal articles have the form:
   `http://adsabs.harvard.edu/cgi-bin/article_query?bibcode`

2. Embedded Queries: Such embedded query hyperlinks in documents perform ADS abstract service queries. This allows users for instance to include links in documents that return their bibliography. Such links have the form:
   `http://adsabs.harvard.edu/cgi-bin/abs_connect?parameters=values&return_req=no_params`

The parameters describe what query is to be executed. A description of these parameters is available at:

http://adsdoc.harvard.edu/abs_doc/embed_form.html

An example of such a link is:

http://adsabs.harvard.edu/cgi-bin/abs_connect?author=eichhorn,+g&return_req=no_params

This link returns the bibliography for the author of this section.

3. Perl Scripts: We provide Perl scripts that allow developers to directly access our database and retrieve information in Perl arrays. This allows other systems to build their own query forms and custom format the returned abstracts. The Perl scripts and a description on how to use them are available at:

http://adsdoc.harvard.edu/abs_doc/direct_access.html

4. Email Queries: The ADS can be queried through email without needing WWW access. The results are returned via email. This allows our users for instance to automatically query the ADS every month to get the latest entries in our database. The user can build a query form that produces the desired results and the save the form locally. Sending this saved query form via email to adsquery@cfa.harvard.edu will execute this query and return the results to the sender. By specifying -31 in the Entry Date Day field, the query will return matching references that were entered in the database within the last 31 days. A more detailed description of this capability is available at:

http://adsdoc.harvard.edu/abs_doc/whatsnew_help.html

3.1.6 Usage Statistics

The ADS abstract service has seen rapid growth since the introduction of the WWW interface. In May, 1997 the ADS was accessed from more than 10,000 hosts. This translates into many more users since many hosts (especially hosts of Internet access providers) are use by more than one user. These users made over 235,000 queries and retrieved 4.5 million references.

The number of people using the abstracts service (more than 10,000) is comparable to the number of astronomers world wide. This indicates that almost all astronomers by now use the ADS at one time or another. Out of the 10,000 users in May, about 2,000 users issued as many queries as there were working days in May (on average one query per working day). This means that about 20% of astronomers use the ADS on average on a daily basis.

3.2 ADS Article Service

3.2.1 General Description

The ADS article service provides access to scanned images of full journal articles. We have agreements with the following publications to scan their journals and make them available on-line a certain time after publication:

9

TABLE 2

Astronomical Journal (USA)
Astrophysical Journal (USA)
Astronomy & Astrophysics (Germany)
Baltic Astronomy (Lithuania)
Bulletin of the Astronomical Society of India (India)
Contributions of the Astronomical Observatory Skalnate Pleso (Slovak Republic)
Monthly Notices of the Royal Astronomical Society (Great Britain)
Publications of the Astronomical Society of Australia (Australia)
Publications of the Astronomical Society of Japan (Japan)
Publications of the Astronomical Society of the Pacific (USA)
Revista Mexicana de Astronomia y Astrofisica (Mexico)

Our scanning is not yet complete as of February 1997. We have scanned 700,000 pages so far and have about 400,000 pages on-line. About 200,000 pages are waiting to be scanned. We are also discussing similar arrangements with other journals. Currently the article database uses about 140 GBytes of disk storage. Eventually we will need up to 500 GBytes. We have decided to use regular magnetic disks for the storage of the article images. This is quite feasible with todays disk technology and will be even easier with the next generation of magnetic disks which should hold more than 20 GBytes per disk.

3.2.2 Image Storage Technology

The journals were scanned with a resolution of 600 dots per inch (dpi). This provides very high quality images. Printed with a 600 dpi printer, they are almost indistinguishable from the original. The images are stored on our server in two resolutions (200 dpi and 600 dpi). This allows the fast retrieval of a smaller, lower resolution version. The images are stored in TIFF format with G4 compression. This format provides a very good compression ratio. An added advantage is the fact that Postscript Level 2 uses the same compression algorithm. The tiff files can therefore be directly converted into Postscript Level 2 files for printing.

The user can retrieve images in several formats. All these formats are generated in real-time when the user requests them:

1. GIF format, 100 dpi, 4 level gray-scale, anti-aliased for screen viewing. These images are about 100–200 KBytes per page in size. They download reasonably quickly. The quality and resolution is good enough so they can be read on the screen. To save disk space, but also conserve compute resources, the gif images are created on demand and then cached. This allows for fast downloading of frequently requested articles, while saving space by not storing screen-view images of older articles that are read only infrequently. The time to create the gif images in real-time is small compared to the transfer time and does therefore not significantly increase the time to download a page. The caching is done more to decrease the load on our server than to speed up the image transfer.

2. Postscript files for printing. We provide three Postscript versions: 600 dpi Postscript Level 2, 200 dpi Postscript Level 2, and 200 dpi Postscript Level 1. All three versions

are produced from the stored tiff files in real-time. The 600 dpi version takes longer to download and to print, but it provides the best print quality. We store the 200 dpi tiff files in order to make the retrieval of the low resolution version as fast as possible. This version prints very quickly. The quality is still better than a fax and generally as good as a good quality photocopy. The 200 dpi version can also be retrieved in Postscript Level 1 format. This is for back compatibility with older Postscript printers. This version is rather large, and takes long to download and print.

3. PCL files for printing on PC printers. These versions too are created in real-time. They print on most PC printers. They are available at 300 dpi and 150 dpi (PCL does not allow for 600 dpi printing).

4. TIFF G4 compressed files. The original tiff files can be downloaded for viewing, printing or other processing.

One important note to our users: All these journal article images are copyrighted by the owner of the original copyright of the journal. They are to be downloaded only for personal use. Any commercial use is prohibited, unless express written permission of the owner of the copyright is given. Images should in general not be stored on the user's system and should not be distributed further. All access should be done only through the ADS article service.

3.2.3 Accessing the Article Service

The article database is closely linked with the ADS abstract service. Any retrieved abstract for which a scanned article is available is directly linked to the on-line article. The articles can also be retrieved directly by volume and page number. The user can retrieve articles either for viewing on-screen, for printing, or for storing on local disk. The ADS server will instruct the browser to handle the file according to the requested retrieval method. Most browsers can handle direct printing requests, so the file does not have to be stored and then printed. The user can select whether to print the complete article or only selected pages.

For users whose browser can handle WWW cookies, the system returns a form with just one default print button and a link to another page that has all the retrieval options. The default behavior can be set in the user preferences. This includes for instance the resolution, whether to return Postscript or PCL, and whether to send the file directly to the printer, to local storage, through fax or through email. For users whose browser does not handle cookies, a page will be returned that allows the selection of the most important of these options directly and all the others through another page.

3.2.4 Statistics

In 1995 we had only scanned and made available on-line the Astrophysical Journal Letters. By the end of 1995, more journals were added on a regular basis. In May 1997, over 235,000 pages were retrieved. The number of pages retrieved is more than what any astronomical library handles.

11

4 Mirrors

In order to provide better service to non-US users, we are trying to duplicate our database in different places. We currently have mirrors at the CDS in France and at the Astronomical Data Analysis Center (NAOJ) in Tokyo, Japan. This greatly improves access times from Europe and East Asia.

The data and software are mirrored over the network every time they are updated. Semi-automatic procedures make the mirroring operation fairly easy and quick.

The ADS software was developed with the possibility of mirrors in mind. All parts of the software that depend on the location of the server and on the directory structure are guided by two configuration files. These configuration files are read during startup of the server. No re-compilation of the software is necessary to move the ADS data and software to a mirror server.

5 Future Plans

In the near future we plan to complete the scanning of the astronomical literature from 1975 to present. There are still several important journals in the planetary and solar system sciences that have not given us permission to scan their journal. Hopefully we will receive permission from them to complete the digital library in Astronomy.

In the longer term we hope to scan the historical literature as well. This is a much larger and more complicated effort and needs to be done in cooperation with the traditional astronomy libraries.

For the abstract service we plan to improve the coverage of the astronomical journals. For several journals we have only references, not abstracts. It should be in the interest of the journals to be represented in the main search system for the astronomical literature. Any journal that is not represented with abstracts will per force not be represented well in the references returned to our users. Since our users comprise the vast majority of the astronomical community, this will be a significant disadvantage for these journals.

We are currently working with the publisher of the IAU Circulars and the Minor Planet Center to include their publications in the ADS. These would be included as soon as they are published (within less than one hour). We are also working with the Information Bulletin on Variable Starts to index their publication in the ADS.

We hope to be able to mirror the ADS at a few more sites. Main targets for this would be Australia, India, Germany, and Great Britain.

The collaboration with the journals of the American Astronomical Society to cross-link between the journal references and the ADS abstracts is working very well. Hopefully we will be able to start similar cooperations with other journals as they prepare to come on-line.

This overall collaboration between all the parts of the astronomical community provides the basis for an example of how a complete digital library can work. Hopefully the continuation of this collaboration will provide the astronomical researchers with a constantly improving tool for their research.

6 Addresses

The ADS Homepage is at:

`http://adswww.harvard.edu`

The abstract service can be accessed from:

`http://adsabs.harvard.edu/ads_abstracts`

This form has links to both the primary site at SAO and to the mirrors in France and Japan.

The ADS can be contacted by sending email to:

`ads@cfa.harvard.edu`

7 References

Accomazzi, A., Grant, C.S., Eichhorn, G., Kurtz, M.J., Murray, S.S, ADS Abstract Service Enhancements, Astronomical Data Analysis and Software and Systems IV, R.A. Shaw, H.E. Payne, and J.J.E. Hayes, eds., Astron. Soc. Pac. Conf. Ser., Vol. 77, pp. 36-39, 1995.

Accomazzi, A., Grant, C.S., Eichhorn, G., Kurtz, M.J., Murray, S.S., The Article Service Data Holdings and Access Methods, Astronomical Data Analysis and Software and Systems V, Astron. Soc. Pac. Conf. Ser., Vol. 101, pp. 558-561, 1996.

Boyce, P.B., Dalterio, H., Electronic Publishing of Scientific Journals, Physics Today, Vol. 49, No. 1, pp. 42-47, 1996

Egret, D, Wenger, M., and Dubois, P., The SIMBAD Astronomical Database, Databases and On-line Data in Astronomy, D. Egret and M. Albrecht (eds), Kluwer, pp. 79-88, 1991

Eichhorn, G., An Overview of the Astrophysics Data System, Experimental Astronomy, Vol. 5, pp. 205-220, 1994.

Eichhorn, G.,Accomazzi, A., Grant, C.S., Kurtz, M.J., and Murray, S.S., Access to the Astrophysics Science Information and Abstract System, Vistas in Astronomy, Vol. 39, pp. 217-225, 1995a.

Eichhorn, G., Murray, S.S., Kurtz, M.J., Accomazzi, A., Grant, C.S., The New Astrophysics Data System, Astronomical Data Analysis and Software and Systems IV, Astron. Soc. Pac. Conf. Ser., Vol. 77, pp. 28-31, 1995b.

Good, J.C., Overview of the Astrophysics Data System (ADS), Astronomical Data Analysis Software and Systems I, D.M. Worrall, C. Biemesderfer, and J. Barnes (eds), Astron. Soc. Pac. Conf. Ser., pp. 35-43, 1992

Kurtz, M.J., Karakashian, T., Stern, C.P., Eichhorn, G., Murray, S.S., Watson, J.M., Ossorio, P.G. and Stoner, J.L., Intelligent Text Retrieval in the NASA Astrophysics Data

System, Astronomical Data Analysis and Software and Systems II, eds. R. Hanish, R. Brissenden, and J. Barnes, Astron. Soc. Pac. Conf. Ser., Vol. 52, pp. 132-136, 1993.

Mermilliod, J.-C., Weidmann, N., Hauck, B., The Lausanne Photometry Server on the World Wide Web, Baltic Astronomy, Vol. 5, pp. 413-416, 1996

Murray, S.S, Brugel, E.W., Eichhorn, G., Farris, A., Good, J.C, Kurtz, M.J., Nousek, J.A., and Stoner, J.L., The NASA Astrophysics Data System: A Heterogeneous Distributed Processing System Application, Astronomy from Large Databases II, eds. A. Heck and F. Murtagh, European Southern Observatory Proceeding, p. 387, 1992.

Plante, R.L., Crutcher, R.M., Sharpe, R.K., The NCSA Astronomy Digital Image Library, Astronomical Data Analysis and Software and Systems V, Astron. Soc. Pac. Conf. Ser., Vol. 101, pp. 569-572, 1996.

Schmitz, M., Helou, G., Dubois, P., LaGue, C., Madore, B., Corwin Jr., H.G., and Lesteven, S., NED and SIMBAD Conventions for Bibliographic Reference Coding, Information & On-line Data in Astronomy, eds. D. Egret and M.A. Albrecht, Kluwer Academic Publishers, Sect. 24, 1995.

Squibb, G, NASA Report on Astrophysics Data System Workshop, 1987

Squibb, G, NASA Astrophysics Data System Study, Final Report, 1988