Access to the Astrophysics Science Information and Abstract System ⋆

G. Eichhorn, A. Accomazzi, C.S. Grant, M.J. Kurtz, S.S. Murray

Smithsonian Astrophysical Observatory, Cambridge, MA, USA

Abstract

The Astrophysics Science Information and Abstract System (ASIAS) is sponsored by NASA to make astronomical data and literature references available to the astronomical community. It provides access to abstracts, full article images in bitmapped form, and catalog and archive data through the World Wide Web. These services are accessible through the ADS Data Services page at http://adswww.harvard.edu/ads_services.html.

1 Introduction

The Astrophysics Data System (ADS) has provided access to astronomical and astrophysical data and references since 1991. Originally it provided access to the data through a distributed client/server system, the first of its kind at that time [1–3]. This access software was based on a proprietary protocol and was therefore somewhat restricted.

When in 1994 the World Wide Web (WWW) became popular, the direction of the Astrophysics Data System (ADS) changed. We began to make our services available on the WWW [4,5] to avoid the complications of a system that was based on proprietary software. At the same time, because of funding restrictions, we needed to limit our activities to the most essential part, the abstract service [6,7] which was the most used part of the ADS.

We now concentrate mainly on making bibliographic information available through the World Wide Web, but we continue some support for our catalog

⋆ This work is funded by the NASA Astrophysics Program under grant NCCW-0024.

and archive services on the WWW. All the ASIAS services have been available through the World Wide Web since early 1994. We are now concentrating on expanding the bibliographic services to provide more information and functionality. This includes making more abstracts available, adding access to full journal articles, and providing access to reference books to build a complete digital library.

2 ASIAS Abstract Service

2.1 Abstract Service Data Base

Most of the abstracts are provided by the NASA Scientific and Technical Information (STI) group [8]. These abstracts are of a uniform format in straight ASCII text. They have keywords in a common keyword system and provide information such as the author affiliation, etc. Other references come from different sources, such as the SIMBAD project [9] at the Centre de Données astronomiques de Strasbourg (CDS), the Astrophysical Journal (ApJ), the Publications of the Astronomical Society of the Pacific (PASP), individual authors, and others. All these references are integrated in two databases, one for astronomy and astrophysics references, and one for space instrumentation references. These two databases can be searched separately with the abstract service search engine.

2.2 Abstract Service Search Engine

The abstract service provides access to currently almost half a million references for the Astronomy and Astrophysics and Space Instrumentation literature. The user can query for abstracts by specifying certain selection criteria on a WWW query form.

The abstract service allows the user to query the database by author names, astronomical object names, keywords, and title and text words [6]. The results of each search in the different fields can be combined in user selectable ways.

The user can specify whether or not a search in a particular field must be successful in order for the abstract to be selected. Requiring the author field for instance allows the user to retrieve only references that have a specified author.

Within a field, the user can select whether to require all entries to be found or only at least one of them. This allows for instance to search for articles that refer to two or more specified astronomical objects by selecting the "AND" combination for the object field.

For author names and regular words in the title or abstract, we have built a list of equivalent words (synonyms). When synonym replacement is requested (the default), the search will return not only the abstracts with the specified words, but all abstracts that have equivalent words in them. This allows for instance to correct for misspellings in author names and text words.

In order to allow queries by object name, the abstract service has a direct connection to the SIMBAD database at the CDS. When a user issues an object query, the object name(s) are sent to Strasbourg to the SIMBAD database. The returned list of references is then parsed and cross-referenced with our database and the resulting references with abstracts are returned to the user.

The results of a query are returned sorted by a score. This score is a measure for the relevance of each article to the query. The more words from the query that were found in the abstract, the higher the score will be. The user can select whether to use direct scoring, where the score is twice as high if twice as many words are found, or weighted scoring, where each word is given a different weight. The weight is calculated from the inverse log of the frequency of the word in the complete database. This means that rare words will have higher scores (since they are presumably more important to the search).

The total weight is calculated by combining the weights from the different query fields. The relative importance of the different fields can be adjusted by modifying the relative weight of each field. The total score is then normalized to 1. The returned list is sorted by decreasing score. For references with equal scores, the most recent reference is listed first.

2.3 Abstract Queries

The ASIAS abstract service is accessed through a WWW forms interface. The Abstract Query Form at http://adsabs.harvard.edu/abstract_service.html provides the user with easy access to the capabilities of the abstract service. The top part of the form provides input fields to specify the query parameters for the different query fields, like author names, publication date limits, titles, etc. The bottom part of the form allows the user to configure the abstract service search engine. A complete description of these capabilities is available on-line at http://adsdoc.harvard.edu/abs_doc/abs_help.html.

The abstract service returns a list of abstracts. This list includes the bibliographic codes [10], the score of the abstract in the search, the publication date, the list of authors, the title, and a list of items that are available for this reference. For instance a query with "Eichhorn, G" in the author field and "gamma ray detectors in antarctica, SN 1987A" in the abstract text field with the default settings returns a list of abstracts with different scores.

Two abstracts have a score of 1.0 since all the words were found in an abstract with the specified author. Subsequent references have progressively smaller scores, depending on how many words were matched. The bibliographic codes in this list are links to the full abstracts. Selecting such a link returns all the information that is available for this reference. This usually, but not always includes the abstract. It is also possible to retrieve several abstracts at once by checking the checkbox next to the selected references and clicking on the "Retrieve selected abstracts" button at the bottom of the page. This retrieval allows the user to select either screen viewing, sending to a printer or saving to a local file.

The list of available items is a series of letters with hyperlinks. The letters stand for the following:

- **A**. A NASA/STI abstract is available (some of the references have only authors and titles).
- **O**. An original author abstract is available. We get these abstracts from some journals, conference proceedings, and from contributions by individual authors. We are trying to get them from all major journals.
- **F**. The full article is available in image form (see below).
- **N**. The author has submitted an author note (see below).
- **D**. Data tables are available for this reference. These data tables are made available by the CDS.
- **E**. The full article is available in electronic form. The AAS has started to work on electronic publishing. They will soon have the Astrophysical Letters on-line. We will link to any electronic publications that are made publicly available.

These links are also at the top of each abstract that is retrieved from the abstract service.

2.3.1 NASA/STI Abstracts

This link indicates that we have an abstract from NASA/STI in our database. For references that we get from other sources, this may not be the case (for instance references from SIMBAD).

2.3.2 Original Author Abstracts

We include original author abstracts into our system whenever they are available to us. We currently have abstracts from Astronomy and Astrophysics, the Publications of the Astronomical Society of the Pacific, and Contributions of the Astronomical Observatory Skalnate Pleso in our system. We are trying to get abstracts from other sources such as the American Astronomical Society. We would welcome any contributions of abstracts from journals or conference proceedings, as long as they are within our subject area and conform to our data format (see http://adsabs.harvard.edu/abs_doc/abstract_format.html). In certain cases we may even be able to include abstracts that are in a different format. If you have abstracts that you want to contribute, please contact the ADS project at ads@cfa.harvard.edu.

2.3.3 Full Article Images

For some of the papers we have the images of the full articles on-line. If an article is available for a selected abstract, links to the images are available in the returned abstract. An example of this can be found when retrieving the second abstract in the previous list. These links allow the user to either view the pages on the screen or to send the images with different resolution directly to the printer.

2.3.4 Author Notes

We give authors the capability to submit additional information to any reference in our system on which they are an author. This can be used for instance to give more detailed information about the topic of the article, or to refer to other publications. These author notes can be either plain text or fully formatted html files, including hyperlinks. The author is responsible for the content of the note, but we reserve the right to reject author notes if we deem it necessary.

2.3.5 Data Access

The CDS provides access to data tables published in some astronomical articles. We have interfaced with the CDS system to provide links from our abstracts to their data tables whenever they are available for a selected abstract. An example can be found with a bibcode query for 1982ApJS...49...27U, an article in the 1982 ApJ Supplements. The full abstract has a link to the data tables, which can be retrieved directly from the CDS.

2.4 Find Similar Abstracts

This mechanism to find similar abstracts allows the user to automatically formulate a new query from the retrieved abstract. It is available at the bottom of each abstracts that is retrieved from the abstract service. This is an easy way to do exhaustive literature searches for a given subject. The user specifies which parts of the current abstract should be used for a new query and then executes this query without having to transfer data from one window to another.

As an example you can look at the abstract from the previous example about Gamma Ray Observations from Antarctica. This abstract has links to the full article as well as the 'Find Similar Abstracts' form at the bottom. This form allows to either execute the query immediately or to return a new form that has the specified fields filled with the information from the current abstract. This allows further modifications to the new query before sending it off.

2.5 Article Service

In January of 1995 the ASIAS released a new service that provides access to the full journal articles. We have so far scanned 20 years of Astrophysical Journal Letters (1975-1994). We are in the process of scanning the Astrophysical Journal and the Astronomical Journal for the same time period. We also have permission to scan the Contributions of the Astronomical Observatory Skalnate Pleso and the Revista Mexicana de Astronomia y Astrofisica. All these journals will come on-line during the summer of 1995.

The images of these articles are stored in three different versions. One version is for viewing on the screen, the other versions are for printing.

The screen view images are in GIF format in 4 level greyscale. The images for the ApJ Letters are stored with resolution of 75 dots per inch (dpi), for all the other journals, the resolution is 100 dpi. This makes them readable on the screen without too much effort. However, this resolution is not good enough for printing. For printing there are two versions available: A version with a resolution of 200 dpi (150 dpi for ApJL) and one with 600 dpi (300 dpi for ApJL). The 200 dpi version prints very fast, but is not quite as readable as the images with the highest resolution, especially for plots and figures. The higher resolution version prints slower, but considerably improves the image quality. The lower resolution printing versions are available as Postscript Level 1 or Level 2 files and as TIFF files with G4 compression. The higher resolution versions are available as Postscript Level 2 and as TIFF files with G4 compression. The link to "More Article Retrieval Options" allows to retrieve different versions and either print them, view them, or save them on the local computer.

If the print files are not printed automatically, the browser needs to be configured properly. To enable the browser to do that on a Unix system, add the following line to your .mailcap file in your home directory (or create the file .mailcap with this one line in it if it doesn't exist):

application/remote-printing; lpr %s

where "lpr" should be the print command on your system. After restarting the browser, the transferred print files should automatically be sent to the printer.

The article images are accessible through the abstract service as mentioned above, or directly through the Reference Query Page at http://adsabs.harvard.edu/article_service.html if the user knows the journal reference of a paper. Here, the user specifies the journal, volume and page of an article to retrieve the bitmapped images.

2.6 Bibcode Query

We also provide the capability to Query by Bibcode at http://adsabs.harvard.edu/bib_abs.html. This query form allows access to our database via complete or partial specification of bibcodes. The partial bibcodes can include a '*' as a wildcard at the beginning and/or end of the string. The returned list is in the same format as the list retrieved from an abstract query. This form allows the user for instance to find all articles in a particular volume of the A&A by specifying a partial bibcode like *A&A...281*.

2.7 Table of Contents Query

For the major journals we provide the capability to get a Table of Contents at http://adsabs.harvard.edu/toc_service.html. This form allows access to a list of publications by year, month, or volume number. The returned list is in the same format as from an abstract query. The capability to get lists by publication month is available only for the journals that we have scanned. As we get more complete lists of references from other journals, we will provide this capability for these journals as well.

2.8 List Queries

The List Query form at http://adsabs.harvard.edu/list_abs.html allows the user to retrieve lists of data items in the abstract service system. For authors and keywords, the user specifies search templates. The templates can include a '*' as a wildcard at the beginning and/or end of a string. The abstract service then returns a list of authors or keywords that match the template. For synonym lists, the user specifies a word and the abstract service returns the list of synonyms for this word. When the user enters an object name in the object query field, the abstract service queries the SIMBAD database at the CDS for all names of the specified object.

Full lists of authors and keywords are also available. These lists are accessible through hyperlinks over the text entry boxes on the abstract query form.

2.9 Direct Access via Bibliographic Codes

This access method allows access to the ASIAS bibliographic data directly through bibliographic codes [10] or NASA/STI accession number. The bibliographic code can be constructed if the reference of an article is known and allows direct access to this abstract or article as long as it is in the database.

Both the abstracts and the full article images are available directly for instance for building links in other documents. The URL to use for the direct access to abstracts is:

http://adsabs.harvard.edu/cgi-bin/bib_query?"bibcode"

and for full article images it is:

http://adsabs.harvard.edu/cgi-bin/article_query?"bibcode"

where "bibcode" is the bibliographic code for the requested abstract or article, or its NASA/STI accession number. This allows any data provider to access our database and link to our data from other documents. This can be used for instance to build reference lists in electronically published articles that contain links to the referenced papers. If there is a '&' in the bibcode, it has to be escaped as %26 (its hex code). A bibcode verification utility is available at http://adsabs.harvard.edu/verify.html to check bibcodes for correctness and availability of the different data items. This utility allows the user to input a bibcode and get back information on whether the bibcode is a valid bibcode and what data items are in our database. It can also be accessed directly via:

http://adsabs.harvard.edu/cgi-bin/verify?"bibcode"

This allows batch validation of bibcodes for systems that automatically build and include reference links.

3 ASIAS Astronomical Catalogs

The ASIAS WWW Catalog Service at http://adscat.harvard.edu/catalog_service.html provides access to more than 170 astronomical catalogs. These are some, but not all of the catalogs that were in the Classic ADS. We try to maintain access to as many of the catalogs as possible within the current budget limitations. If you need a particular catalog on-line, please contact us. Maybe we can include it in our database.

The main catalog service query page allows the user to select how the catalog lists are to be sorted (by node name, subject, or catalog name) and how the catalogs are to be queried (QBT: Query by Table, SQL: Standard Query Language (SQL) query). it also allows to issue a WAIS query to find catalogs by searching for words in the catalog description.

Selecting "List by nodes" presents the user with a list of data centers that provide access to catalogs. After selecting one node and clicking on the "Get Catalog List" button, the list of catalogs available from this node is available for selection of a specific catalog. Similarly, the "List by subject" form presents a list of astronomical subjects. The "List by name" gives an alphabetical list of all catalogs in the system.

Once a list of catalogs is displayed, the user can retrieve the query form for the desired catalog. The SQL query form allows the user to specify any SQL query that the database of the selected catalog understands. This type of catalog query requires sufficient knowledge of the SQL language. It allows the user to use all the capabilities of the database where the catalog is stored when selecting data from that catalog.

Easier to use is the QBT form. This form presents the user with a lists of all the fields in a catalog. It allows the user to specify how the catalog is to be queried and in what format the data are to be returned. The user selects which fields should be retrieved from the catalog, specifies the query criteria and the desired output format (if different from the default format). It is slightly more limited in its capabilities, but in most cases provides more than enough options to select the desired data.

The data can be returned in three formats:

(i) An ASCII table in a format that most relational databases can use,
(ii) An EOS table, the format that the Classic ADS used, and
(iii) A FITS table.

Each of the three formats can be requested either as TABLE or as FILE. Selecting the TABLE format will display the data in the window of the WWW browser. If the FILE format is selected, the WWW browser will ask for a filename and then automatically store the data in the requested format on a local disk.

4 ASIAS Astronomical Data Archives

The ADS currently provides access to one data archive at the Harvard/Smithsonian Center for Astrophysics, the Einstein Archive at http://adsarc.harvard.edu/einstein_service.html, one archive at the Center for Astrophysics and Space Astronomy, an Atlas of Ultraviolet P Cygni Profiles at http://adswww.colorado.edu/pcyg/pcyg.html, and one archive at the National Institute for Standards and Technology with an Atomic Spectroscopic Database at http://aeldata.phy.nist.gov/nist_atomic_spectra.html.

4.1 Einstein Archives

The Einstein Archive service provides access to images and photon event lists collected by the Einstein satellite. The main window provides a selection of search types:

- By Position (Box)
- By Position (Circle)
- By Field Title
- By Sequence Number
- By 2E Catalog Number

Selecting one of the search types returns a WWW form with query fields appropriate for the query type. For example the Query by Field Title Form contains a text input field for title words. All the query forms contain the selection fields for the data types and the instruments for which the user requests the data. The user selects what type of data to return from which instrument, fills in the selection criteria and sends the query. The Einstein server returns a list of files that satisfy the query condition.

As an example, a query with “COMA” in the title field and default values for the rest of the form returns a list of eight files with general information about these files. The filename of each file is a hyperlink to the actual file on the Einstein server. This allows the user to retrieve files of interest simply by clicking on the name of the selected file. The file is then transferred and stored on the local disk for further analysis or display.

The files are returned as FITS files. If the user configures the browser properly, it can automatically start up a viewer with the transferred FITS file. To enable the browser to do that on a Unix system, add the following line to your .mailcap file in your home directory (or create the file .mailcap with this one line in it if it doesn't exist):

image/x-fits; saoimage -fits %s

and then restart the browser. This will direct your browser to automatically start SAOimage and display the transferred image. You can replace SAOimage with any other viewer that can display FITS images by replacing "saoimage" in the .mailcap file with the name of the viewer that you would like to use.

4.2 Ultraviolet P Cygni Profiles Archives

The Atlas of Ultraviolet P Cygni Profiles allows the user to browse the text, tables, and figures of "An Atlas of Ultraviolet P Cygni Profiles" [11] as well as to provide access to the spectral data used in the Atlas. GIF images of the spectra as they appear in the Atlas can be retrieved. The spectral data used in the Atlas can be transferred in the form of binary FITS tables.

4.3 Atomic Spectroscopic Data Archives

The NIST Atomic Spectroscopic Database provides access and search capability to critically evaluated atomic spectroscopic data on energy levels, wavelengths, and transition probabilities that are reasonably up-to-date. The Atomic Energy Levels Data Center and Data Center on Atomic Transition Probabilities and Line Shapes have carried out these critical compilations. Both Data Centers are located in the Physics Laboratory at the National Institute of Standards and Technology (NIST). It allows to query for information about what data holdings are in the database. Once desired data sets are identified, they can be retrieved through a forms interface. These interfaces were developed jointly by the ADS project and NIST.

5 Future Plans

The major effort in the near future will be spent on making more journals available. We plan to increase the number of articles from our present journals as well as to start scanning issues from other journals. In the next few months we plan to have about 20 years each of Astrophysical Journal and the Astronomical Journal on-line. We will also bring on-line three other journals (Publications of the Astronomical Society of the Pacific, Revista Mexicana, Observatory Reports from Skalnate Pleso) for which we have already permission. We are negotiating with other journals to obtain permission to make their journal articles available. We are also talking with editors of several conference proceedings to bring their abstracts and possibly articles on-line.

We welcome any suggestions on how we can improve the ASIAS services. Please contact us if you have suggestions for new services or know about sources of abstracts that we may use. If you are involved in publishing a journal or proceedings and are interested to bring your abstracts or articles on-line, please let us know.

6 More Information about ADS

More information about the ADS project is available through the ADS Home Page at http://adswww.harvard.edu/, or you can send e-mail to us at ads@cfa.harvard.edu with suggestions, improvements, comments, abstracts, etc.

References

[1] J.C. Good, Overview of the Astrophysics Data System (ADS), in: D.W. Worral, C. Biemesderfer, and J. Barnes, eds., *Astronomical Data Analysis and Software and Systems* **I** (Astron. Soc. Pac. Conference Series, 1992) 205.

[2] S.S. Murray, E.W. Bruegel, G. Eichhorn, A. Farris, J.C. Good, M.J. Kurtz. J.A. Nousek, J.L. Stoner, The NASA Astrophysics Data System: A Heterogeneous Distributed Processing System Application, in: A. Heck and F. Murtagh, eds., *Astronomy from Large Databases* **II** (European Southern Observatory, 1992), 387.

[3] G. Eichhorn, An Overview of the Astrophysics Data System, *Experimental Astronomy* **5** (1994) 205.

[4] G. Eichhorn, The New Astrophysics Data System, *Information Systems Newsletter* **34** (1994) 31.

[5] G. Eichhorn, S. S. Murray, M. J. Kurtz, A. Accomazzi, and C. S. Grant, The New Astrophysics Data System, in: R. Hanisch, R. Brissenden, and J. Barnes, eds., *Astronomical Data Analysis and Software and Systems* **IV** (Astron. Soc. Pac. Conference Series, in press).

[6] M.J. Kurtz, T. Karakashian, C.P. Stern, G. Eichhorn, S.S. Murray, J.M. Watson, P.G. Ossorio, and J.L. Stoner, Intelligent Text Retrieval in the NASA Astrophysics Data System (ADS), in: R. Hanisch, R. Brissenden, and J. Barnes, eds., *Astronomical Data Analysis and Software and Systems* **II** (Astron. Soc. Pac. Conference Series, 1992), 205.

[7] A. Accomazzi, C. S. Grant, G. Eichhorn, M. Kurtz, and S. Murray, ADS Abstract Service Enhancements, in: R. Hanisch, R. Brissenden, and J. Barnes, eds., *Astronomical Data Analysis and Software and Systems* **IV** (Astron. Soc. Pac. Conference Series, in press).

[8] T.E. Pinelli, NASA Scientific and Technical Program, *Special Edition of Government Information Quarterly* **5** (1990) 119.

[9] D. Egret, M. Wenger, P. Dubois, 1991, The SIMBAD Astronomical Database, in: D. Egret and M. Albrecht, eds., *Databases and On-line Data in Astronomy* (Kluwer, 1991) 89.

[10] M. Schmitz, G. Helou, P. Dubois, C. LaGue, B. Madore, H.G. Corwin Jr., and S. Lesteven, NED and SIMBAD Conventions for Bibliographic Reference Coding, in: D. Egret and M. Albrecht, eds., *Information and On-line Data in Astronomy* (Kluwer, in press).

[11] T.P. Snow, H.J.G.L.M. Lamers, D.M. Lindholm, and A.P Odell, An Atlas of Ultraviolet P Cygni Profiles, ApJ Supplements, *Astrophysical Journal Suppl.* **95** (1994) 163.