

## Linking in the Astrophysics Data System

Guenther Eichhorn

*Smithsonian Astrophysical Observatory, Cambridge, MA 02138, USA,*  
*e-mail: gei@cfa.harvard.edu*

### Abstract.

The Astrophysics Data System (ADS) provides access to references, abstracts and full journal articles of the Astronomy/Astrophysics, Space Instrumentation, and Physics/Geophysics literature. The Astrophysics Data System is part of an Astronomy-wide collaboration called Urania. The members of this collaboration make a variety of information available on-line and provide cross-links between their data holdings and other information providers. This allows the users of this system to easily find related information, regardless of where they start.

The ADS can be accessed at: <http://adswwww.harvard.edu/>

## 1. Introduction

The ADS Abstract Service (Eichhorn 1997, Eichhorn et al. 1998a, Kurtz & Eichhorn 1998 ) provides access to over 1.1 million references and abstracts and over 700,000 scanned journal article pages (as of April 1998). This database covers most of the astronomical literature. One of the very important parts of the ADS database is the collection of links to other information sources that are associated with the references in the ADS.

This linking between different information providers is an extremely important resource for the working astronomer. Figure 1 shows schematically how this network links the different resources. This Astronomy-wide collaboration is called Urania (Universal Research Archive of Networked Information in Astronomy). The collection of links in the ADS allows other information providers access to all resources in the Urania collaboration through one link to the ADS, without the need for each of these groups to collect and maintain all these links.

## 2. Links in the ADS

The ADS Abstract Service allows the users to search the database of references and abstracts (Eichhorn et al. 1998b). The result of a query is returned as a list of references, sorted by how well the reference matches the query (Eichhorn 1997).

Figure 2 shows a page with the results of an ADS query. The results list contains the title and author list, the bibliographic code, publication date, and score (Eichhorn et al. 1995). In addition to this basic information it contains a

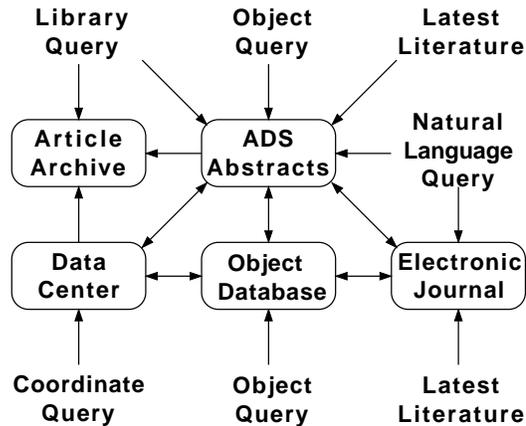


Figure 1. Connectivity Network in the Astronomical Digital Library collaboration called Urania

list of links with each reference, anchored to individual letters. Table 1 shows the list of different links available, together with the number of times each link is contained in the database.

The ‘O’ and ‘A’ links point to different versions of the abstract for the reference. The ‘A’ link points to abstracts that were re-written by the Scientific and Technical Information project (STI) at NASA. The ‘O’ link points to the original author abstract. We get these author abstracts either directly from the journal, from conference proceedings editors, or from the authors themselves.

The ‘E’, ‘F’, and ‘G’ links point to different versions of the full article. The ‘E’ links point to on-line HTML versions that make use of the World Wide Web (WWW) technology with links, animations, tables, etc. The access to these links is generally restricted to subscribers of the journals. The ‘F’ links point to articles in image format. These can be either Postscript or PDF versions, or both. These versions also may be restricted to subscribers. The ‘G’ links point to articles that were scanned in the ADS project (Eichhorn 1997). Access to these articles is available without restriction world-wide.

The ‘D’, ‘P’, ‘S’, and ‘N’ links point to different data sets associated with the article. The ‘S’ and ‘N’ links point to lists of astronomical objects mentioned in the article. They are maintained by the SIMBAD project at the Centre de Données astronomiques de Strasbourg (CDS), France and the NASA Extragalactic Database (NED) project at the Jet Propulsion Laboratory (JPL) in

**Astrophysics Data System (ADS)**  
**Query Results from Astronomy Database**  
 Retrieved 30 abstracts, starting with number 1. Total number selected: 24092

Bibcode	Score	Date	Available Items
Authors	Title	Access Control	Help
<a href="#">1997ApJS...113..333M</a>	0.560	12/1997	<a href="#">Q</a> <a href="#">E</a> <a href="#">E</a> <a href="#">E</a> <a href="#">E</a> <a href="#">E</a> <a href="#">E</a>
MATONECK, DAVID M.; FESEN, ROBERT A.; BLAIS, WILLIAM P.; LONG, KNOX S.	An Optical Search for Supernova Remnants in NGC 2403		
<a href="#">1984ApJ...279..157M</a>	0.344	04/1984	<a href="#">A</a> <a href="#">G</a> <a href="#">E</a> <a href="#">C</a> <a href="#">S</a> <a href="#">N</a>
MIDDLEBITCH, J.; KRISTIAN, J.	A search for young, luminous optical pulsars in extragalactic supernova remnants		
<a href="#">1995A&amp;A...305..527C</a>	0.342	01/1995	<a href="#">Q</a> <a href="#">Q</a> <a href="#">E</a> <a href="#">C</a> <a href="#">S</a>
CABRIT, S.; GUILLOTEAU, S.; ANDRE, P.; BERTOUT, C.; MONTMERLE, T.; SCHUSTER, K.	Plateau de Bure observations of HL Tauri: outflow motions in a remnant circumstellar envelope.		
<a href="#">1993A&amp;A...277..453D</a>	0.330	10/1993	<a href="#">G</a> <a href="#">D</a> <a href="#">E</a> <a href="#">C</a> <a href="#">S</a>
DAVID, P.; LESQUEREN, A. M.; SWAGNANAM, P.	AN OH SATELLITE LINE MASER SURVEY OF COOL IRAS SOURCES AND CIRCUMSTELLAR ENVELOPE EVOLUTION		
<a href="#">1993A&amp;A...275..265D</a>	0.315	06/1993	<a href="#">Q</a> <a href="#">G</a> <a href="#">E</a> <a href="#">C</a> <a href="#">S</a> <a href="#">N</a>
DICKEL, J. R.; MILNE, D. K.; JUNKER, N.; KLEIN, U.	N63A - A SUPERNOVA REMNANT WITHIN AN HII REGION		
<a href="#">1989ComAp...13..117I</a>	0.315	00/1989	
IKEUCHI, S.; YOSHIOKA, S.	TEMPORAL AND SPATIAL STRUCTURES OF AN INTERSTELLAR MEDIUM DOMINATED BY SUPERNOVA REMNANTS		

Figure 2. Results list from ADS abstract query

Pasadena, CA, USA respectively. The ‘D’ and ‘P’ links point to on-line data sets associated with the reference (the ‘P’ links point to the Planetary Data System (PDS), the ‘D’ links point to several different data centers).

‘R’ and ‘C’ links point to the list of references in the article and the list of publications that cite the article, respectively.

‘M’ links point to on-line document delivery systems that allow the user to order copies of individual articles. ‘L’ links point to entries for books in the Library of Congress’ on-line library. ‘T’ links provide access to the table of contents for conference proceedings volumes.

The bibliographic codes mentioned above are a vital part of the Astronomy Digital Library and the Urania collaboration. They allow different information

Link	Description	Number
A	NASA/STI abstract	170,00
C	Citations for article	71,000
D	Data available	4,700
E	Electronic on-line article	25,000
F	Full text of article	14,000
G	Scanned journal articles	77,000
I	Author information	100
L	Library catalog entry	18,000
M	Mail ordering of articles	3,000
N	NED objects	27,000
O	Original author abstract	49,000
P	PDS on-line data	5
R	References in article	79,000
S	SIMBAD objects	98,000
T	Table of contents	1,000

Table 1. Links types and their numbers in the ADS database

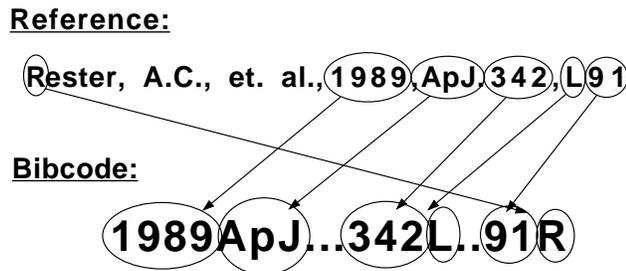


Figure 3. Composition of Bibliographic Codes

providers to link to individual articles without having to coordinate the link information for each article. These codes can be calculated automatically from a regular journal reference and can be easily understood by the user. They are by now used by most information providers in Astronomy. Figure 3 shows how these codes are constructed. A more detailed description is available on-line at:

[http://adsdoc.harvard.edu/abs\\_doc/bibcodes\\_help.html](http://adsdoc.harvard.edu/abs_doc/bibcodes_help.html)

A list of journal abbreviations used in these codes is available at:

[http://adsdoc.harvard.edu/abs\\_doc/journal\\_abbr.html](http://adsdoc.harvard.edu/abs_doc/journal_abbr.html)

These bibliographic codes (or bibcodes for short) provide the glue for the linking between the different partners in the Urania collaboration. Most systems provide so-called “name resolvers”. These name resolvers accept a request with

a bibcode as identifier and forward the request to the internal address where the data items for that reference are stored. This system makes it very easy for one data system to link to another without having to coordinate link information for each reference. Every data provider can calculate the necessary bibcodes without consulting other servers. The ADS provides a bibcode verification utility that allows the system that generates the bibcode to check whether the code is available in the ADS.

Number of Abstract Service Users	17,000
Number of Queries	370,000
Number of References retrieved	6,100,000
Number of Abstracts retrieved	300,000
Percentage of Author Queries	65%
Percentage of Text Queries	30%
Percentage of Title Queries	25%
Percentage of Object Queries	5%
Percentage of Multiple Field Queries	18%
Number of Article Service Users	8,000
Number of Queries	150,000
Number of Pages retrieved	380,000

Table 2. Monthly usage statistics of the ADS database for March, 1998

Link code	Link type	Usage
A,O	Abstracts	300,427
G	Scanned Articles	34,362
F	Full Text Articles	12,557
E	Electronic Articles	9,100
C	Citations	4,121
S,N	Object Lists	2,593
R	References	2,336
D,P	Data	830
M	Mail Order	183
L	Library Entries	133
T	Table of Contents	63
I	Author information	13

Table 3. Monthly usage of links in the ADS database for March, 1998

### 3. Linking of Mirror Sites

The ADS as well as many other information providers have multiple sites with the same information to improve access for users in different parts of the world. The ADS currently has mirror sites at the CDS in Strasbourg, France and at NAO in Tokyo, Japan. We are in the process of setting up a third mirror in Chile. Other information providers also maintain multiple mirror sites. This has two different implications for the ADS.

The first one is to synchronize our data and software at our mirror sites to always provide the same information. The second is to manage our links so they can point to different mirror sites of our data and the data of other data centers.

We have solved the first issue by developing automated procedures that synchronize the data and software on our mirror sites. The mirroring operation is started from a password protected web form. The data manager selects which part of the system needs to be mirrored and to which mirror site it needs to be transferred; the rest is done automatically through the mirroring scripts. The mirroring scripts use the same configuration files as the server software and can therefore run on any of our sites without modification (just as the server software which is configured entirely through these configuration files).

The second issue cannot be resolved by the ADS alone, since the decision as to which mirror site of a particular data source to use depends on the connectivity between the user and the different mirror sites of the data source. The mirror selection therefore has to be done by the user. We provide a choice of mirror sites in our preference settings form. The user can select any of the mirror sites available. This selection is then used whenever a link to this data source is built by the ADS software. All choices for the different mirror sites of the various information providers are defined in a configuration file. This configuration file is distinct for each ADS mirror site, so that the defaults for each data source are set appropriately for each of our mirrors. This provides a very flexible scheme for handling links to various mirrors of our collaborating data centers.

### 4. Access Statistics

The ADS services are accessed by astronomers world-wide. Table 2 lists some access statistics for March, 1998. The number of users of the Abstract Service per month is 17,000. This is about the number of astronomers world-wide. This means that most working astronomers are using this service at least once per month. The number of references retrieved per month of 6.1 million means that each reference on average is retrieved almost six times per month.

The usage of the different query fields (see table 2) shows that the majority of queries ask for articles based on author names (65%). Text queries are next most frequent with 30%, followed by title queries with 25%. Object queries are used in 5% of the queries. Queries with search criteria in multiple fields make up less than 20% of the queries. This statistic clearly shows that having references without abstracts in the database is still a very valuable resource, since they can be found in 90% of the queries that our users issue.

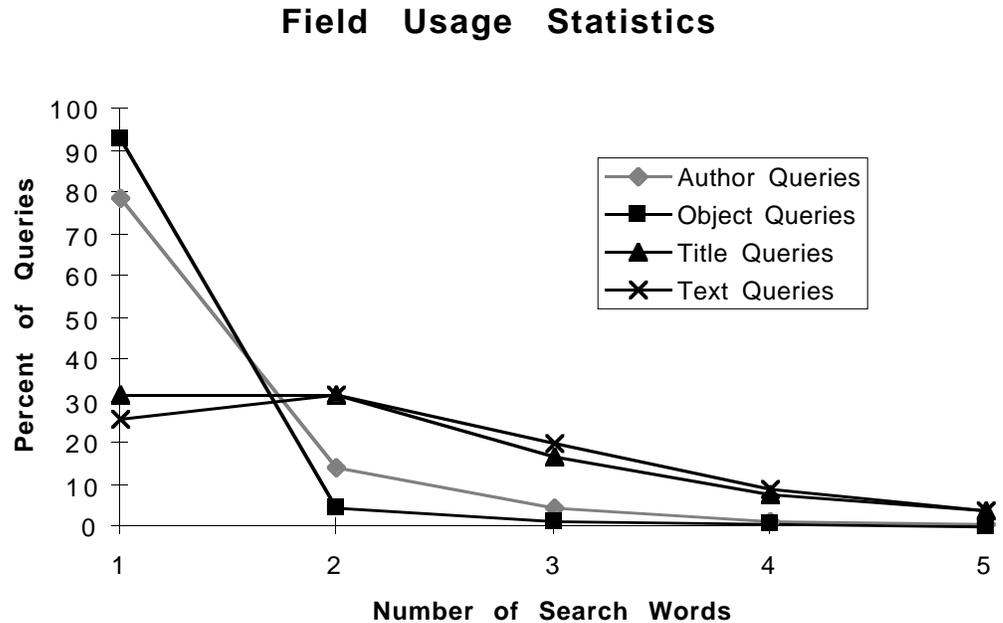


Figure 4. Usage of the different query fields in the ADS abstract service

The query fields fall into two categories as far as the number of search words is concerned. Author queries and object queries are used overwhelmingly with only one search term (author name or object name respectively). Searches with two authors make up about 15% of the author queries, searches with two objects only about 5% of the object queries. On the other hand, title and text searches are done frequently with more than one word. Title and text searches with one or two words make up about 30% each of these queries, searches with three words make up about 20% each.

Kurtz & Eichhorn 1998 show some more detailed access statistics for the *Astronomical Journal* for articles published between 1944 and 1997.

Table 3 shows how often each of the types of links was used in March, 1998. Most frequently used, of course, are the abstracts. The next most frequently used link type, as expected, are the full articles. Between on-line HTML articles, article images on other sites and scanned articles on the ADS server, 55,000 articles were called up in March, 1998. This amounts to over 2,500 articles per working day on average! The references and citations are called up over 2,000 and 4,000 times respectively. This too is a frequently used information source. The data links are used less frequently as expected. This does not mean that these links are not important. The capability to directly access data from a reference is an important part of the overall system.

## 5. Summary

Linking to different on-line resources has become a vital part of most on-line information providers' tools. The ADS has become a critical part of this system of links by collecting and managing links to all information providers in the field of Astronomy. This allows other data centers to access this system of links through one link to the ADS. This frees the data centers from the need to find and maintain links to other resources that their users may need. The users are the ultimate beneficiaries of this system of interlinked astronomical resources. They can easily access different information providers without having to do separate searches at the individual centers. The acceptance of this system has been overwhelming as the usage statistics of the ADS show.

**Acknowledgments.** This work was funded by NASA under grant NCCW 00254.

## References

- Eichhorn, G., Accomazzi, A., Grant, C. S., Kurtz, M. J. 1998, The Astrophysics Data System, these proceedings, 277
- Eichhorn, G., Accomazzi, A., Grant, C. S., Kurtz, M. J., & Murray, S. S. 1998, New Capabilities of the ADS Abstract and Article Service, in *Astronomical Data Analysis Software and Systems VII*, ASP Conference Series, Vol. 145, R. Albrecht, R.N. Hook and H.A. Bushouse, eds., 378
- Eichhorn, G. 1997, The Digital Library of the Astrophysics Data System, *Ap&SS*, 247, 189
- Eichhorn, G., Accomazzi, A., Grant, C. S., Kurtz, M. J., & Murray, S. S. 1995, Access to the Astrophysics Science Information and Abstract System, *Vistas in Astronomy*, 39, 217
- Kurtz, M. J. & Eichhorn, G. 1998, The Historical Literature of Astronomy, via ADS, these proceedings, 293