ASTRONOMY LIBRARIES – YOUR GATEWAY TO INFORMATION

Uta Grothkopf[1]

Abstract. This chapter reviews services offered by astronomy libraries for assisting astronomers in their research. Special attention is given to the two most important information tools in astronomy: the NASA Astrophysics Data System (ADS) and the arXiv e-print server. Some less known features will be presented to lead to more efficient use of these tools. The core astronomy journals are explained, along with the open access concept and how it interacts with existing journals. The chapter also provides an introduction to bibliometric studies and shows why publication and citation statistics are important for both researchers and observatories. Finally, examples of cooperation between librarians and astronomers are presented in the context of working groups and conferences.

1 Introduction

The world of libraries is changing fast. The times when they were simply repositories of printed documents are long gone, because preserving the astronomy legacy and being curators of historic documents is important, but is not the only challenge librarians meet today. In the internet era, librarians deliver services directly to the scienstist's desktop, either by providing access to scientific journals and databases through assistance in using specific information tools or simply by answering reference questions via email or messaging systems. Astronomy libraries are a good example of this changing paradigm, because the vast majority of information resources are available electronically, and most astronomers are used to, and actually prefer, online resources over printed formats.

To use information efficiently, students and researchers need to acquire some basic skills. These can be divided into three parts.

- *Information literacy:* We have a surplus of information rather than a shortage. Scientists have to be able to locate the texts that are important for their research, otherwise they may waste time with information they do not actually need. Effective

---

[1] European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany, esolib@eso.org

use of finding aids like library catalogs and reference works is called information literacy.

- *Computer literacy:* A certain understanding of the hardware, software, and network technology is needed to access information tools from the office, from home, or from abroad. This knowledge is called computer literacy.

- *Critical thinking skills:* Regardless of the medium (print or electronic, local or networked), no search result should be accepted without questioning its validity, correctness and completeness. Such critical thinking skills include common sense, as well as techniques for fully understanding a search result and putting it into context.

The combination of information literacy, computer literacy, and critical thinking skills is defined as information fluency, as explained in Viele (2008). Librarians will help develop these techniques and thus make the most of the time spent on literature and information searches.

2   What librarians can do for you

In today's interconnected world, a large fraction of the scholarly resources in astronomy are available online. Latest preprints are retrieved from the arXiv (astro-ph) e-print server, journal articles can be obtained via the NASA Astrophysics Data System (ADS) abstract service, and communication with colleagues of course mostly takes place by email. When asked what information services they use in their daily work, scientists often do not immediately think of their library. Yet, most astronomers do use libraries, even though they may not always realize it.

Start thinking of libraries not only as collections of printed books and journals, but also as gateways to electronic publications. Every time you download a recent journal article from the ADS, most probably you are allowed to do so because your library has paid the necessary subscription fee, negotiated access rights, and does the troubleshooting every time access is interrupted.

Librarians also compile web pages with link collections that provide access to electronic books, reference works, and databases. They help scientists discover new information retrieval systems and assist them in their research. To do this, librarians form professional networks to overcome the limited local resources and make use of joint collections, knowledge, and ideas provided by astronomy librarians worldwide.

In many organizations, librarians are in charge of institutional repositories. They hold introductory sessions to explain how to deposit manuscripts in the repository, make sure all necessary metadata get submitted, and are often experts in copyright and Open Access (OA). They also maintain databases of all papers published by the organization's scientists or, in the case of observatories, of those papers that use observational data from specific facilities. From these databases, statistics are derived and made available to management, governing bodies, and the public.

The library mission traditionally has been to fulfill the information needs of their users by selecting, collecting, preserving, and providing access to relevant resources.

While many changes have occurred regarding the tools used in library management, this mission is still valid. In the digital age, personalized service, tailored to the needs of the library users, distinguishes libraries from software tools, see Grothkopf (2003).

Increasingly, research libraries act as meeting places. They provide space for both informal and formal meetings, ranging from the "science coffee" that takes place every morning to regular discussion groups and formally organized talks and presentations. Instead of the stereotypical quiet reading room, libraries often host collaborating scientists who appreciate the relaxed and inspiring atmosphere in the library.

Similar to the virtual "Web 2.0" concept, the traditional library has been replaced, or rather supplemented, by the "Library 2.0" that offers web pages with library blogs, RSS feeds, and interactive tools for contacting the librarians. Of course, especially in small, specialized libraries like those that often can be found in observatories, the most interactive way of consulting librarians is still to simply meet and talk to them. They are happy to provide research assistance, help you with finding quality information resources and suggest all kinds of tips and tricks you may not have heard of.

3 Two astronomy gorillas

Without doubt the majority of astronomers cover almost all their needs for scholarly information using only two resources: the NASA Astrophysics Data System (ADS) and the arXiv.org e-print server. ADS and arXiv clearly are the "gorillas" among their peers as the largest, most important, most visible resources. And they are both free of charge to the individual astronomer, hence available to everybody, no matter whether affiliated with an institution or not, or regardless how wealthy that institution is.

In the following, we briefly introduce both services and point out some features for using them efficiently.

3.1 NASA Astrophysics Data System (ADS)

The NASA ADS abstract service is operated by the Smithsonian Astrophysical Observatory (SAO) under a NASA grant. The main site is at `http://adsabs.harvard.edu`[1], with several mirror sites located around the world. It is the largest digital library in astronomy, providing links to more than 7 million records in astronomy, astrophysics, and physics, and it is used almost daily by the majority of astronomers. Access to subscription-based publications (including the two or three most recent years of the core astronomy journals) is granted only if the user or his/her institution subscribes to that publication. For the years 1975 through approximately 1995, the abstracts database contains data from a variety of sources; since 1995, most abstracts have been provided directly by publishers or journal editors. These references should be 100% complete. With less completeness depending on the subject area, but still at a high rate, ADS also provides information about older literature and, in many cases, offers links to scanned articles back to volume 1 of the most important astronomy journals. In addition to published literature,

---

[1] ADS main search screen: `http://adsabs.harvard.edu/abstract_service.html`

ADS provides information about preprints posted on arXiv.org. Preprints can be retrieved via the ADS the same day they become available on arXiv, so that there is no delay in finding the most recent postings.

ADS receives information about publications from a variety of sources, most notably directly from publishers on or even before the publication date, as well as from the arXiv preprint database. ADS users can query the database for immediate results or subscribe to the myADS service to receive regular notifications about new publications that correspond to search criteria set by the user.

3.1.1   Some search tips

The ADS main search screen is mostly self-explanatory. However, there are some features users do not necessarily know about. Some of them are explained below, and they have been available in the ADS since January 2009. Since the ADS is continuously being developed and improved, further capabilities can be expected in future.

First author searches: To limit searches to first authors alone, use ˆ before the author's name. The resulting list will only include papers with that person as first author.

Filters: By scrolling down on the main search screen, you will reach the Filters section with a variety of options. For instance, to retrieve only publications from refereed journals, select "All refereed articles" in the first Filters section ("Select References From"). By using the "Select/deselect publications" option, users can enter journal codes of publications that are to be included in (or excluded from) the results. For a list of journal abbreviations, see the ADS Frequently Asked Questions page[2].

Citations: A very useful addition to the bibliographic record is the information about citations. To see the number of citing papers, go to the "Sorting" section towards the end of the screen and select "Sort by citation count" (or "Sort by normalized citation count" to normalize by the number of authors). Then enter your search. Alternatively, you can first enter your search and then use the "Sort options" pull-down menu in the upper right corner to sort by citations. On the results screen, you will see the number of citations above the title of each entry. It will be displayed with three decimal places (e.g., 32.000) which are only important for the normalized citation count when citation fractions below 1 can occur.

For any given result set, you can get the list of citations for all or selected articles of the result list. The citation lists can include or exclude self-cites, and they can be limited to citing papers in refereed journals. In order to use these features, first enter a search on the main search screen. Make sure the whole set of results is displayed; if necessary, the number of items to be returned can be increased (default: 200). On the results page, select the records for which you want to see the citing articles, or scroll down below the last record and choose "Select All Records". Then scroll down further until you find the

---

[2]`http://adsabs.harvard.edu/abs_doc/journals1.html`

option "Get citation lists for selected articles" or "Get refereed citation lists for selected articles". To exclude self-citations, check this option.

A detailed description of the creation and use of citations in the ADS has been given in Accomazzi et al. (2007).

Export search results: Next to the Sorting section on the main ADS page, you will find the Format area, which governs the output format of search results. Instead of the default short list format, a variety of formats can be selected that allow easy importing into bibliographic tools like EndNote and ProCite. To add records to a Zotero library (see also Section 9 on organizing your own library), users can simply click on the paper icon that shows up in the URL field of the browser window. ADS records can also be displayed in a format ready for cutting and pasting into TeX manuscripts, e.g., BIBTEX or AASTeX. Reference lists already formatted for specific journals can be generated, for instance for *Icarus, MNRAS,* and *Solar Physics.*

Historical literature: In addition to current publications, the ADS also provides access to historical literature. Together with the John G. Wolbach Library at the Harvard-Smithsonian Center for Astrophysics, the ADS digitizes microfilms from selected historical publications in astronomy. Although this compilation still lacks pagination so that one cannot go directly to a specific page, it still is very useful, in particular when looking for content that is otherwise difficult to find, such as observatory publications, reports, and annals. Lists of historical scans available so far are provided by the ADS[3].

Data links: The D links available for many records allow users to access the astronomical data used in these papers. This information is typically provided by the data center or observatory where the data are available. For instance, papers using data from telescopes of the European Southern Observatory (ESO) are linked by the ESO librarians to the programs that generated the data, and the D link for these records will take users directly to the ESO Archive where the data can be requested.

The "Select References In" option in the Filters area lets users limit queries to papers using data from a specific observatory, for instance HST, ESO, Gemini, Subaru, and Keck.

Catching errors: Some of the content available through the ADS is provided by libraries, even though the individual links do not reveal that information. In particular, librarians supply original bibliographic records, corrections to existing (especially historical) entries, and information about data links for their specific observatory.

3.1.2   Result lists: bibcodes, citation counts, and letter links

ADS result lists with brief entries display a variety of information above the authors and title. These are the bibcode of the record, the citation count, publication month and year, as well as several letter links, depending on availability.

---

[3] `adsabs.harvard.edu/historical.html`   `adsabs.harvard.edu/journals\_service.html`

Bibcode: In astronomy, bibcodes are used to uniquely identify publications; they are assigned by the ADS. A bibcode consists of 19 characters. The first four represent the publication year, followed by up to five characters for the journal or book abbreviation, up to four digits for the volume, one optional character to indicate a special section (e.g., Letters), up to four digits for the starting page, and finally the first letter of the family name of the first author. Unused characters/digits are filled with dots. A typical bibcode will be for instance 2005ApJ...635..260S, representing the article by Savaglio et al., published in *ApJ*, vol. 635, starting on page 260.

Citation count: For the citation count of an article to be displayed, the respective selections in the Sorting section must be made, see above.

Publication month and year: The publication month and year are shown in the format MM/YYYY. This information is useful for instance when citations are compare, because it makes a difference how many months ago a paper was published.

Letter links: ADS entries provide links to additional information, symbolized by capital letters on results pages. Some of the information is stored on the ADS server, other links take users to services that reside elsewhere. The number and kind of links vary. The most important letter links are: A = abstract, C = citations (other articles citing the paper), D = data links, E = article fulltext in HTML, F = article fulltext in PDF, N = list of objects mentioned in the article from NED (NASA/IPAC Extragalactic Database[4]), O = associated articles (e.g., errata), R = references (articles cited by a paper), S = list of objects mentioned in the article from SIMBAD database[5], X = links to articles in the arXiv database.

The letters are usually shown in blue. Fulltext links (PDF and HTML) of subscription-based articles will only allow access to the paper if you or your institution subscribe to the journal. However, E and F links of some articles will display in green, indicating that these articles are open access, i.e., available to everybody. Whether open access is granted and how soon after publication, depends on the journal. For an explanation of open access, see Section 6.

3.2   arXiv.org e-print archive

The other large service in astronomy is the arXiv.org e-print repository[6]. The main server is located at the Cornell University Library. It provides open access to over half a million e-prints in physics, mathematics, computer science, quantitative biology, and statistics. For astronomers, the most important arXiv section is astro-ph, an e-print archive that dates back to April 1992. Like the ADS, also arXiv has mirror sites in various countries.

Originally developed by Paul Ginsparg when he was still affiliated with the Los Alamos National Laboratory (LANL), arXiv is now owned, operated, and mostly funded

---

[4]http://nedwww.ipac.caltech.edu/
[5]http://simbad.ustrasbg.fr/simbad
[6]http://arxiv.org

by Cornell University.

3.2.1   Manuscripts uploaded by authors

ArXiv/astro-ph is a collection of manuscripts posted by authors. There is no obligation for any author to make papers available, but the majority of astronomers post their papers on the e-print archive (Schwarz & Kennicutt (2004)). The repository can be accessed via the web, or one can subscribe to email notifications and RSS feeds.

When submitting a paper, authors are asked to grant arXiv.org the non-exclusive right to distribute the article and certify that the work is either in the public domain or available under a Creative Commons (CC) license[7]. These licenses provide copyright holders with a mechanism for (partly or fully) waiving their rights in order to allow sharing of information, but CC license holders will still be credited if others use their works. The advantage is that obstacles typically arising from copyright laws can be avoided. Authors have to make certain that the license they grant to arXiv does not conflict with a journal agreement they may have signed[8].

To ensure that arXiv content is relevant to current research, an endorsement system was introduced in 2004. Since then, papers of first-time submitters have to be endorsed by existing members of the arXiv community before they are accepted by the e-print server. In January 2009, astro-ph was split into six sub-categories. Submissions have to be assigned to at least one of these subcategories. The split is meant to make announcements of new papers more manageable for those interested only in subsets of astro-ph[9].

As many authors post their papers before acceptance by a journal, several versions of the same paper can be posted while the refereeing process continues. The author decides whether every iteration is posted as a modified version. Typically, the text of the final posted version is identical to the published paper. The layout will most probably be different as the original typeset layout is copyrighted by the publisher, and most publishers will also require that you use and cite the refereed and published version of the paper, not simply the online arXiv version.

3.2.2   Finding your way through arXiv/astro-ph

On the main arXiv.org screen, astro-ph is listed as the first option in the physics section. Users can browse the entire database, go directly to entries made available today or during the recent week, or search astro-ph by author name, title, or abstract words and other categories.

For each entry, the arXiv number is shown, as well as title, author, and author-defined comments, e.g., number of pages and figures, and status of the manuscript (submitted, accepted, revised version, etc.).

The arXiv papers should be cited in the format arXiv:YYMM.NNNNv# [category], for instance arXiv:0803.1234v1 [astro-ph], to indicate version 1 of paper number 1234,

---

[7]http://creativecommons.org/license/
[8]http://arxiv.org/help/license
[9]http://arxiv.org/new/#jan2009

posted in March 2008 on the astro-ph server.

When citing manuscripts that have merely been submitted to journals and are not yet accepted, bear in mind that they are highly likely to undergo (sometimes substantial) changes before they are actually published. Indeed, scientists typically still cite the published version of papers whenever possible. Future changes may be one of the reasons, as well as the fact that the final reference reveals the journal where the paper will be published, so it conveys some of its reputation, e.g., refereed journal versus conference proceedings. This effect is desired when compiling a publication list for the curriculum vitae or home page.

The vast majority of astronomers post their manuscripts on astro-ph; for instance, an estimated 80% of the papers we look at in the context of the ESO Telescope Bibliography (see Section 9.3) are submitted to the e-print server. Schwarz & Kennicutt (2004) find that, on average, *ApJ* papers posted on astro-ph are cited more than twice as often as those not posted. Does it matter where in the daily list of new postings a manuscript is positioned? Dietrich (2008a; 2008b) finds that papers appearing at or near the top of astro-ph listings receive significantly more citations than those on lower positions. Because submissions are listed in the order they arrive at the astro-ph server, with each new list starting with papers received at 4 pm (US Eastern time), authors do have some control over the rank of their articles. It remains to be seen whether Dietrich's findings are considered an unwanted drawback to the arXiv service by the managers at Cornell so that it leads to modifications in the organization of the ranking system.

4   Databases beyond ADS and astro-ph

With ADS and astro-ph being so dominant and relevant for a large number of queries, other services are often neglected, even though they may provide additional and sometimes unique information.

4.1   SPIRES-HEP

SPIRES-HEP[10] is the High-Energy Physics database of the Stanford Physics Information Retrieval System (later renamed Stanford *Public* Information Retrieval System). It was originally developed at the Stanford Linear Accelerator Center and is now a joint project of SLAC, DESY (Deutsches Elektronen-Synchrotron), and FNAL (Fermi National Accelerator Laboratory, Fermilab), managed by the SLAC library.

SPIRES-HEP provides bibliographic information about high-energy physics literature, including journal articles, theses, reports, conference papers, books, and others, with links to the fulltexts of these items. Among others, arXiv is one of the sources from where fulltexts of papers referenced in SPIRES-HEP can be obtained.

---

[10]http://www.slac.stanford.edu/spires/hep/

4.2 Google Scholar

Google Scholar (GS)[11] is a free research tool for scholarly literature and citations. While it is known that Google Scholar covers journal articles as well as conference proceedings, books, reports, and other media, and that even full-text repositories are searched, GS keeps users in the dark about exactly which sources it is using.

Similar to the main Google search engine, Google Scholar also does a very good job of correcting misspellings and suggesting different spellings including name variations. It covers a wide range of subjects, but seems to be generally stronger in the fields of social sciences, arts and humanities, and engineering (see for instance Harzing & van der Wal 2008). One drawback is that result lists can contain duplicates.

4.3 Scirus

Another literature search tool is Scirus[12]. It is maintained and provided free of charge by Elsevier Science, one of the largest scientific publishers. As of January 2009, over 480 million items are indexed, including refereed and unrefereed publications, repositories, science-related websites from the American Physical Society (APS), arXiv, the Institute of Physics Publishing (IOPP), NASA, and many other sources. Scirus offers a sophisticated advanced search interface that allows users to filter by date range, information type, and subject area (including astronomy). However, this tool only helps for retrieving publications and other information resources; citations are not provided.

4.4 ISI Web of Knowledge and Web of Science (WoS)

The ISI (Institute for Scientific Information) Web of Knowledge[13] is a platform that provides access to a variety of databases (e.g., Inspec, Medline, Biosis) and tools for retrieving and analyzing research articles. Some backfiles, including citations, date back as far as 1900.

The citation database included in the ISI Web of Knowledge is called Web of Science (WoS)[14]. It covers over 9,000 publications in the sciences, social sciences, and arts and humanities, but is limited to journals, so that citations from books, reports, and other publications not included in the WoS are not counted. WoS is used by many organizations to analyze their scientists' research output, but it has been repeatedly criticized for omitting citations from books and for focusing mostly on English-language titles from the US and Europe.

In the past, the Journal Citation Report (JCR), which is prepared annually by the Web of Knowledge and which presents journal rankings, has been criticized for reporting flawed journal impact factors because of difficulties in handling possibly non-unique

---

[11]`http://scholar.google.com`
[12]`http://www.scirus.com`
[13]`http://www.isiwebofknowledge.com`
[14]`http://isiwebofknowledge.com/products_tools/multidisciplinary/webofscience/`

journal abbreviations. In astronomy, the *Astrophysical Journal* and *Astronomy & Astrophysics* have been particularly affected by this mistake (Sandqvist 2004). In the meantime, the error has been corrected and the acccurate journal impact factors published.

Unlike SPIRES, Google Scholar, and Scirus, WoS is not free. It is a commercial system, marketed by Thomson Reuters, that needs a subscription.

4.5 Scopus

Described as the "largest abstract and citation database of research literature and quality web sources" on their web site, Scopus[15] has become a serious competitor for WoS. As is the case with WoS, Scopus requires a subscription. The tool was developed and is marketed by science publisher Elsevier.

Scopus offers information from over 16,000 peer-reviewed journals, as well as conference proceedings, book series, scientific web pages, and repositories, including arXiv. Full-text articles can be accessed seamlessly, provided the necessary subscriptions are maintained. Scopus aims for complete coverage of their records from 1996 onwards, but historical material has also been added to the system, going back as far as 1823 in some cases. Citation analysis is provided from 1996 onwards.

One strength of Scopus are the innovative tools for analyzing publication and citation statistics, including the *h*-index and unique graphs.

The scope, coverage, and approach of Google Scholar, WoS, and Scopus are quite different, with WoS and Scopus providing substantial factual information, various search and ranking features, help pages, and guidance on how to use the services best. They act as an information hub (Jasco 2005). In contrast, Google Scholar applies the Google PageRank algorithm to display search results, which typically leads to the most popular hits being displayed at the top. Advanced search and display options are limited in comparison with WoS and Scopus.

Differences are becoming even more obvious when looking at citations. Meho & Yang (2007) find that the overlap in citations between Scopus and WoS is rather small, namely only 58%, and only 31% between Google Scholar and the union of WoS and Scopus. This emphasizes that none of these databases provides an absolute answer regarding completeness of search results and, specifically, numbers of citations. The three services complement each other, but none of them should be regarded as the perfect search system.

The main advice clearly is: when using publication and citation databases, never switch off your own thinking. Always evaluate all resources and put them in a context with other results.

5   Core astronomy journals

Every science has journals that are essential to that discipline. In astronomy, the number of these most important journals is small: only six titles are typically referred to as the

---

[15]http://www.scopus.com

"core journals." These are *Astronomy & Astrophysics (A&A), Astronomical Journal (AJ), Astrophysical Journal* and its *Supplement Series (ApJ/ApJS), Monthly Notices of the Royal Astronomical Society (MNRAS)*, and the *Publications of the Astronomical Society of the Pacific (PASP)*. These journals alone produced on average 36% of all refereed journal articles listed in the ADS for the years 2000–2007.

All core journals in astronomy are published on behalf of organizations or learned societies. *A&A* is published on behalf of the European Southern Observatory (ESO), *AJ* and *ApJ/ApJS* by the American Astronomical Society (AAS), the *Monthly Notices* on behalf of the Royal Astronomical Society (RAS), and *PASP* by the Astronomical Society of the Pacific (ASP). This means that, even if the journals are produced and marketed by commercial publishers, their (partly understandable) economical interests are somewhat leveled out by the societies who often influence or even control the subscription price.

The core astronomy journals are subscription-based; i.e., the latest issues are only accessible with a valid subscription. However, all of them apply the "delayed open access" model. After two or three years, articles are made available to everybody (for a description of open access, see Section 6).

All of them are currently still available in print, as well as electronic format, even though some publishing models are already changing. For instance, the default subscription of the *ApJ Letters* includes electronic-only access as of 2009, and also the basic subscription price for *A&A* will soon only cover the electronic version; in both cases, the print version needs to be explicitly ordered (and paid for). These are first moves towards e-only publishing in astronomy, and it will only be a matter of time when the other core journals follow.

While librarians are not directly involved in the production of the journals, they pay the subscription fee, negotiate licenses (which define who can access the publication, and under which conditions), make journals available through web pages and library catalogs, and troubleshoot in case of access problems.

A&A, AJ, ApJ/S, MNRAS, and *PASP* are available through individual subscriptions that do not require customers to purchase a whole package of journals. Content from all titles can be found through the ADS or via astro-ph, provided that the authors submitted their manuscripts to the arXiv e-print server. In addition to these well-established finding tools, publishers are also starting to develop their own platforms to offer easy access to all their journals, plus additional information like commenting systems, readers' communities, etc. For instance, users of IOPscience[16] will be able to search latest and historical research articles published by IOP. Finding tools range from the established author/title/keyword searches to emerging technologies, such as tag clouds. Retrieved content also includes e-prints from arXiv (presented via an IOP-developed e-print interface) and certain web pages.

This trend away from static journals towards "virtual journals" or article databases can be expected to grow in future. It remains to be seen whether publishers' services will be able to compete with the "master search engine" provided by the ADS.

---

[16]http://iopscience.iop.org

6 Open Access

6.1 The OA concept

A brief description of open access publishing is provided by Wikipedia: "publication of material in such a way that it is available to all potential users without financial or other barriers"[17].

The open access movement goes back to the 1960s, but has gained momentum mostly since the last decade of the past century. Among other reasons, it was prompted by the constantly rising subscription prices of scholarly journals that has led researchers, librarians, funding agencies, publishers. and others into discussion about how scientific publications can be made available in a better way. The main focus is on peer-reviewed, publicly funded literature. The idea is to remove some of the existing barriers, most notably price barriers (subscriptions, licensing fees, pay-per-view charges) and permission barriers (copyright and licensing restrictions, Suber 2007) and make them accessible to everybody. Of course "everybody" is a relative term, as some impediments continue to exist, for instance access limited to members affiliated with an institution, language barriers, accessibility barriers, etc.

There are various forms of open access:

- Self-archiving on authors' home pages: author self-archiving is often called the "green road" to OA. Most publishers (also commercial ones) grant their authors the right to make manuscripts, including the final version, available on the author's web page. If the final typeset version may be used, the publishers typically demand that the bibliographic details are given so that readers are pointed to the publisher's journal. The SHERPA RoMEO website[18] lists publishers' copyright and self-archiving policies.

- Self-archiving in a repository: these can be institutional or subject-based repositories. Many institutions have set up repositories where all affiliated researchers deposit their articles. Some scientific disciplines maintain central subject-based repositories. In astronomy, such a repository is the astro-ph e-print server.

  Self-archiving by scientists neglects two very important aspects: retrieval via a central access point, and preservation. Repositories are more reliable in this regard.

- Publishing in OA journals: typically called "open access publishing", this refers to journals that make their content available without subscription. In today's publishing landscape, many different flavors of OA are applied by publishers, ranging from delayed access (with free access to content after a given time) to institutional membership (whereby all papers are immediately available for those affiliated with the member institute), and on to the "author pays" model (i.e., authors pay a certain fee to the journal, readers don't).

---

[17]`http://en.wikipedia.org/wiki/Open_access_publishing`, accessed Jan. 31, 2009
[18]`http://www.sherpa.ac.uk/romeo.php`

Even though removing barriers related to (sometimes unaffordably high) prices is one of the key incentives of the open access movement, it must be stated that serious OA advocates do not claim that the *production* of open access publications is without costs. The production of OA journals may overall be less expensive than conventionally published literature; however, it still demands funding that needs to be provided by someone. The word "free" (of charge) can therefore only refer to the readers of open access publications, not to the publication process at large.

One of the advantages open access provides for authors is better accessibility because, theoretically, every interested researcher can obtain the publications. Whether or not this actually leads to more citations is still being investigated. For a discussion of this topic in the case of *The Astrophysical Journal*, see Kurtz & Henneken (2007).

According to the key findings of the third ALPSP (Association of Learned and Professional Society Publishers) survey of Scholarly Publishing Practice[19], the proportion of publishers that offer optional open access (authors pay a fee) has grown from 9% in 2005 to 30% in 2008. However, the report concludes that "the take-up of the author pays open access option is exceedingly low." This statement is not surprising, given that (i) the fee authors have to pay for optional OA is sometimes as high as 3,000 Euros and (ii) that, especially in sciences like astronomy the majority of research literature is available either from the arXiv e-print server or directly from the journals through "delayed open access." As a result, astronomers may not be particularly attracted by a publishing model that requires further payments from them.

A description of the state of open access in astronomy can be found in Grothkopf & Erdmann (2008).

6.2   OA finding aids

Several finding aids can be recommended for retrieving open access publications.

6.2.1   Directory of Open Access Journals (DOAJ)

The idea of creating a Directory of Open Access Journals[20] goes back to the First Nordic Conference on Scholarly Communication held in Lund and Copenhagen in 2002[21]. The DOAJ Head Office is located at the Lund University Libraries. As of January 2009, the directory includes over 3,800 scientific and scholarly journals, of which more than 1,300 are searchable at article level. Open access journals are defined as peer-reviewed journals that "use a funding model that does not charge readers or their institutions for access" and that make all content freely available without delay.

---

[19]http://www.alpsp.org/ngen_public/article.asp?pfs=0&did=47&aid=27749&oaid=-1
[20]http://www.doaj.org
[21]http://www.lub.lu.se/ncsc2002/

6.2.2 OAIster

OAIster[22] is a union catalog of digital resources, compiled by harvesting descriptive metadata using the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). With a current staff of only three, the OAIster team has put together an impressive catalog of digital resources, including digitized books and articles, original ("born-digital") texts, audio files, images, movies, and datasets. These items are often located in the so-called deep web, hidden from search engines by firewalls. OAIster reveals these resources by using OAI-PMH and makes them available via a searchable interface. In December 2008, OAIster provided access to more than 19 million records from over 1,000 data providers.

6.2.3 ScientificCommons

The aim of the ScientificCommons project[23] is to "provide the most comprehensive and freely available access to scientific knowledge on the internet". Like OAIster, ScientificCommons uses the OAI-PMH protocol to retrieve information from repositories. According to their website, ScientificCommons had indexed more than 13 million scientific publications as of January 2007. Author names are extracted across institutions and archives so that distributed publications can be found via one common interface. ScientificCommons also extracts professional relations between authors to make their research development transparent to the public. Personalization services include an option for authors to organize their publications directly with ScientificCommons.org in order to create their own researcher profiles.

ScientificCommons is a project of the University of St. Gallen (Switzerland), and is hosted and developed at the Institute for Media and Communications Management.

6.2.4 Directory of Open Access Repositories (OpenDOAR)

OpenDOAR[24] is a compilation of information about and links to academic open access repositories. It aims at increasing their visibility and retrievability by providing end-users with a search interface to locate specific archives by geographic location, subject, content type, and more. Physics and Astronomy are currently represented with 45 repositories.

From its start at the end of 2005, the size of OpenDOAR has increased quite steeply to reach its current coverage of approx. 1,300 repositories. Almost half of the included repositories are located in Europe, approx. 30% in North America, and around 20% in Asia, Australasia, and South America combined. OpenDOAR is maintained by the Unversity of Nottingham in the UK.

7 Problems with online documents

The internet and the advent of electronic publications have enormously modified how scientists use the literature. Especially in astronomy, we are used to finding (seemingly)

---

[22]http://www.oaister.org
[23]http://www.ScientificCommons.org
[24]http://www.opendoar.org

all research publications on the internet in electronic format. The ease of using online publications and their immediate availability should not let us forget that there are some severe obstacles on the way to the e-only scholarly society.

7.1 Long-term access

Ever since the advent of electronic publications, librarians have communicated the challenges in long-term access to the research community. Two major topics have always been of concern: archiving and preservation.

Archiving refers to questions such as:

- what material (scientific articles, electronic communication, databases, etc.) and which resources should be archived? (web pages only from certified providers? "everything"?)

- who should be in charge of archiving? (individual data providers? national libraries and archives? publishers?)

- where should the archived content be stored (private web pages, libraries, data centers, commercial archive providers, etc.) and how many mirror sites should exist?

Preservation of digital information resources encompasses assuring global perpetual access; provision of tools to read and interpret the data; migrating storage media to newer technologies, including all associated metadata, access rights, provenance information and, most importantly, the links within the document and those reaching out to other media items. Reference systems must be in place to make certain that the information can be found. The authenticity (identity) of electronic publications needs to be easily recognizable in order to determine which version of the document is being used, and their links and connections with other digital objects must be transferred to each newer storage medium. Finally, mechanisms are needed to guarantee the integrity (intactness) of digital information.

7.2 Completeness of digitized material

In order to make historical literature, i.e., material originally published on paper, available online, numerous scanning projects are underway at observatories, universities, the ADS, and other data providers. How complete are the results? At the minimum, all projects scan the scientific content, i.e., the articles published in the journals, magazines, conference proceedings, books, or other items that are being digitized. But what happens to other parts, such as errata, news sections, advertisements, letters to the editor, front and back matter, obituaries? Depending on the treatment received during and after digitization, such material may be lost or destroyed.

7.3 Deleted items

In the print environment, the path of publications is quite well-defined. After production at the publisher, it is sent to libraries where the material is displayed and later bound into

volumes and placed on shelves, available for future access whenever needed. Nothing gets lost or is changed, the scholarly content is archived the way it was originally published.

In the digital environment, things are more short-lived. For instance, what happens if a journal publisher finds out that an article plagiarized someone else's work, or if the content, if acted upon, poses a serious health risk? In some cases, it might be considered necessary to remove the article from the publisher's website, even though it had been formally published (see for instance the Elsevier policy on article withdrawal[25]).

Other examples are Wikipedia pages of ceased projects or comments and contributions posted to mailing lists and blogs that may be removed. While such items are intrinsically more ephemeral, they show that the availability of online content should not be taken for granted.

7.4 Broken links

In today's use of web resources, it often seems to be sufficient to indicate specific pages through Uniform Resource Locators (URLs). While many of them reveal which organization or company hosts the source (e.g., en.wikipdia.org) or what specific subdirectories are on that server (e.g., .../wiki/Url), they can also simply consist of figures and then not tell users anything (at least not at first sight) about their origin and host.

The real danger, however, is that URLs point to one exact location. Should the location change or any part of the string be modified, users will not be able to retrieve the intended web page, but will be sent to an error page.

To avoid such broken links, a better naming system has been developed: Digital Object Identifiers (DOIs[26]). DOIs are names, not locations, and provide a unique, permanent digital identifier to objects. Mostly used for scientific publications, it has been widely applied by publishers. By the end of 2008, approx. 40 million DOI names had been assigned.

A DOI follows a defined structure and may look like this: 10.1065/abc123defg, where 10.1065 is the prefix. The first two digits are the directory code; 10 indicates the publishing sector. 1065 represents a given publisher; abc123defg is the suffix. It is chosen by the publisher and has to be unique within the realm of the prefix. DOIs can be resolved through browsers, for instance at the website dx.doi.org. The site name and the DOI are seperated by a forward slash[27].

The system is managed by a membership consortium called the International DOI Foundation. The official link registration agency for scholarly and professional publications is CrossRef[28], a citation-linking network that covers millions of articles and other content items. In order to have DOIs assigned to their publications, publishers have to become CrossRef members and pay an annual fee.

---

[25]http://www.elsevier.com/wps/find/intro.cws_home/Article Withdrawal
[26]http://www.doi.org
[27]for instance, http://dx.doi.org/10.1065/abc123defg
[28]http://www.crossref.org/

8   Everything is on the internet... really?

Compared to other sciences, astronomy traditionally has been on the forefront of the move towards electronic publications. In 1997, only two years after the start of electronic publishing, almost half of the peer-reviewed literature in astronomy was available in electronic format (Boyce 1998). This may lead to believing that by now, *everything* has been made accessible online. This assumption is understandable, but wrong. Information resources that are not available on the internet often are forgotten, neglected, ignored.

Information not easy to find includes material such as:

- special supplements accompanying (in particular historic) publications, such as photographic plates;

- literature in radio astronomy; this refers in particular to historical publications, but also to recent literature published in engineering journals;

- observatory publications, especially those older than two or three years;

- books and conference proceedings, in particular if published more than approx. five years ago;

- non-English language literature.

When looking for literature, it is crucial to use various retrieval tools and not rely on "quick & dirty" internet searches that may tempt you to believe that the result is already complete, while in reality it is not.

9   Organizing a personal library

Having access to the literature is one part of using information, but keeping your own library of papers that are of interest may be referred to in your own publications is another. An additional difficulty arises from journals typically applying their own style for citations. This means that authors need to follow certain guidelines when they compile reference lists for their manuscripts. These guidelines are different for each journal.

For many years, various systems have been around to help scientists with this task. These were mostly commercial programs that had to be purchased and installed on a computer, so-called reference management software. Such software stores information about (scientific) articles, and the bibliographic citations can then be output in various formats to be used in reference lists.

More recently, open source software systems have been made available that are available to scientists at no cost or for a small fee. The majority of these programs are still used locally on the user's desktop, but some have a web complement or are even entirely web-based.

Standard features that most reference management systems fulfill include:

- create, import, export records;

- assign notes, keywords, or tags;

- create collections for specific topics or purposes;

- store copies of the articles (in PDF format) for off-line reading.

Of special interest here is in which format records can be exported. Astronomers usually compile reference lists in one of the following two ways.

- LaTeX: in many cases the list of references is the last section of a manuscript. References are manually added as soon as they are cited in the text. The LaTeX reference list is directly integrated in the manuscript

- BibTex: a BibTex reference list (.bib file) requires a bibliography style file outside of the manuscript which calls the reference that is cited in the manuscript.

Many reference management systems are able to handle BibTex references (e.g., export records in BibTex format), while only a few have a direct link (export feature) for LaTeX references. If references need to be exported from systems without direct LaTeX integration, it is recommended to export them first to a 'plain' export format (e.g., a clipboard) and copy and paste them from there into the LaTeX reference list. A good comparison of reference management software can be found in Wikipedia[29].

In the following, we briefly explain four reference management systems and outline some specific characteristics. These four systems are JabRef, Papers, Zotero, and Mendeley.

- **JabRef:** The JabRef[30] reference manager is available as freeware under the GNU General Public License. It is platform-independent so that it runs equally well on Windows, Linux, and Mac OS. JabRef's main features include the ability to

    – import, organize, annotate, and archive

  records. Existing BibTex libraries (.bib files) can be imported, and the developers claim that it is also possible to fetch entries from the arXiv eprint website. PDFs can be read offline so that users are independent of internet access. Exporting records requires custom export filters, i.e., layout files that reside outside of JabRef. These must be written by users and can be shared via SourceForge.net. JabRef is built around BibTex and is a good tool for users who are comfortable using BibTex.

  Users can create various collections, for instance, to group records by topic or to keep all records pertaining to specific reference lists in one place. In addition, individual records can be highlighted for easier retrieval.

---

[29]http://en.wikipedia.org/wiki/Comparison_of_reference_management_software
[30]http://jabref.sourceforge.net/

- **Papers:** Papers[31] is a reference management software that has been developed specifically for the Apple Macintosh operating system. It is not free of charge; a single-user license costs EUR 29 as of mid-2009. Papers is integrated with the iPhone so that articles can be synchronized.

  The software allows users to

    - find, import, organize, annotate, and archive

  libraries. As is the case with JabRef, import plug-ins exist for BibTex files and other formats. A specialty of Papers are several built-in search engines that provide direct access to articles databases, including the ADS and arXiv. Without having to leave Papers, users can search the repositories, save repeating searches, and use some of their search features, like limiting them to first author searches or other specific fields. Retrieved records can be imported with one keystroke. The search history is accessible from within Papers so that searches can be re-used easily.

  Papers intends to encourage users to read articles on-screen by showing PDFs in full-screen mode, providing the opportunity to annotate them, and by remembering the last reading position in the file. In addition to bibliographic information about articles, Papers also allows users to store email addresses of author names and passwords for access to fulltexts of journals.

  Exporters exist for BibTex, Word2008, and CSV (comma-separated values) formats. Unfortunately, the system does not provide an option to add keywords to records or tag them, which is quite surprising for a software with otherwise high functionality.

- **Zotero:** Zotero[32] is a plug-in for the FIrefox 3.0 (or higher) browser. It can be used on Windows, Mac, or Linux machines. The main features include options to

    - find, download, organize, tag, and archive

  records. Existing libraries can be imported from intermediate formats, e.g., BibTex. The main specialty of Zotero, however, is the ease with which records can be captured directly from the web. Simply by clicking on the little notepad or paper icon in the URL-field, web page content can be imported and then be processed further, e.g. by assigning records to collections that have been created or by attaching files (PDFs, screenshots, etc.) to records. Related papers can be linked to each other.

  Libraries (collections) can be exported for instance to BibTex or Rich Text Format. Alternatively, a bibliography of selected records can be created using any of the journal styles included in the large style gallery. The second option will provide records formatted according to that particular journal's citation style. Unfortunately, no astronomy-related journals are currently included in the list of journals,

---

[31] http://mekentosj.com/papers/
[32] http://www.zotero.org

so that astronomers will probably simply end up creating a bibliography of the records they want to export, copy it to the so-called clipboard, and then format it manually.

The Zotero window can be collapsed entirely in the Firefox browser, in which case only the Zotero icon is visible in the lower right corner. When in use, the program will either occupy the lower third of the browser window or it can be opened to fill the whole window. Zotero is organized in three columns, with the left one holding the various collections, the middle one titles, authors, and publication years of the individual papers, and the right column showing details of the currently edited record. Article PDFs can be stored with the records for offline use, and users can assign tags and notes to records. In addition to published literature, Zotero can also take snapshots of web pages, which can then be viewed offline, annotated, and marked using a highlighter tool or little post-it stickers.

In order to visualize records in the collections, Zotero offers an option of viewing a timeline. Timelines are organized on three levels; filters can be assigned to limit the number of records shown, and highlighting specific keywords allows users to view publishing trends.

- **Mendeley:** Mendeley[33] is a reference management tool for the desktop (Windows, Mac, Linux) with an online counterpart. The program is free of charge, but users need to sign up in order to download the software. The current version (as of mid-2010) is still beta (0.9), indicating that Mendeley is still being developed. The program intends to provide a system with which users can

  - manage, share, and discover

research papers. Records can be imported directly from the ADS, arXiv, Google Scholar, and other databases. Exporting records from Mendeley is currently possible in Word and OpenOffice Writer, using more than 100 citation styles. Once again, astronomy journal styles are not available.

A special feature of Mendeley is its ability to automatically extract information from PDFs through drag and drop. From files dropped into Mendeley, the system will import bibliographic details along with DOIs, arXiv IDs, and other information. PDFs can be annotated using a highlighter tool or sticky notes. Mendeley's powerful full-text search capability retrieves search terms not only in author names, titles, and keywords, but in the entire text of articles.

The program is very much geared towards establishing collaborations among colleagues and tries to act as a social networking tool. By using the web part of the system, users can form networks and invite others to share their papers and discover new ones. Research trends are visualized, for instance, by displaying the most popular authors and topics and through statistics about the papers in the collection.

---

[33] http://www.mendeley.com

All systems described above have some advantages and some disadvantages. JabRef is platform-independent, but its use is not very intuitive, because the program is built around BibTex and therefore will be used mostly by those who are familiar with BibTex. The second program, Papers is more user-friendly, and it searches the ADS and arXiv, but is only available for the Mac and is not available entirely free of charge. Zotero, the Firefox plug-in, conveniently captures metadata directly from the web, thus allowing easy import of records of articles and web pages, but the style gallery for exporting records is not very suitable for astronomers so that manual formatting of references is necessary. Mendeley provides a remarkable full-text search capability aimed at building communities among scientists. Right now, the functionality is still somewhat limited, but the developers seem to be moving fast in order to incorporate more features.

In the end, the decision to use a particular reference management system will depend on personal preferences and on which system matches most of a user's selection criteria.

10   Bibliometric studies

Publicly funded organizations typically have to report back to their management and funding agencies about how their resources were spent and whether they provide the expected output. For scientific entities, the output is increasingly measured by bibliometric methods. Often librarians are in charge of these studies.

Bibliometrics can be defined as a set of methods used for publication and citation analysis to explore the impact in the respective field. Observatories tend to use such metrics to evaluate the research output based on data from their telescopes and instruments, and the acceptance of these facilities among the user community.

Some commonly used bibliometric methods are listed in Table 1. All of them have both advantages and disadvantages. For instance, counting the number of publications reveals the productivity of an entity, but gives no information about the impact and importance of these publications. Citation analysis, assuming that important articles are cited more often by other authors, does reflect the impact among peers. However, citation counts can be inflated, for instance by erroneously counting citations of different scientists who have the same name and initials, because of friends and colleagues who systematically cite each other, or even because of articles that only point out incorrect findings in the cited work (Meho 2007). Calculating the mean or median number of citations per paper has the advantage of allowing comparison (bearing in mind the difficulties intrinsic to such comparisons, see Section 9.1) of publications from different scientists or from various facilities, but high citations to only a few articles can increase the average value, which may wrongly suggest that the majority of papers included in the result set have received high cites.

	Good	Bad
# Publications	Productivity	No impact
# Citations	Impact	Can be inflated
Mean/median cites per paper	Allow comparison of different eras	Reward low productivity
"High-Impact Papers"	Show trends	Favor "hot topics"
h-index	Productivity + impact	Determined by years of operation

Table 1: Advantages and disadvantages of metrics.

Another measure are so-called High-Impact Papers (HIPs), a term coined originally by ISI (now Thomson Reuters) and later applied to publications in astronomy (Meylan, Madrid & Macchetto 2004). An HIP is defined as a paper that belongs to the 200 highest cited refereed papers in a given year. Once this subset is defined, e.g., by using the ADS Abstract Service, papers using observational data are analyzed for the observing facilities that provided the data. This method shows interesting trends, in particular if applied over several years. However, it is very time-intensive to compile and may favor "hot topics" that receive a lot of momentary attention.

A few years ago, Hirsch (2005) proposed a new method, the so-called *h*-index. It aims at quantifying a researcher's scientific research output in a more balanced way by combining numbers of publications and numbers of citations, hence productivity and impact. The idea seems to be simple and yet more effective than many other metrics: the *h*-index is defined as the number of publications that have received at least *h* citations. Using the ADS or other citation databases (e.g., Web of Science, Scopus), the value can be computed quite easily. First, submit a query to retrieve all papers for which you would like to calculate the *h*-index. For instance, such a query can be for all publications by a specific author, using data from a particular telescope, or all scientists affiliated with a specific organization. The set of results needs to be ordered by decreasing number of citations, so that the paper with the highest number of citations resides at the top of the list, the one with the lowest citation number at the bottom. At that point, *h* can be found where the rank number of the paper (starting with rank number 1 at the top) is higher than or equal to the number of citations the paper has received.

As might be suspected, *h* computed as described cannot be the same for authors that have been publishing for many years and for those who just entered research; naturally, the senior researcher will have published many more articles that will have gathered many citations in the course of the years. Similarly, *h* will also not be comparable across disciplines or even across subfields of disciplines; e.g., papers published in the subject area of planetary astronomy will on average have different *h*-indices than, for instance, UV astronomy publications. To make comparisons possible, at least to some extent, Hirsch introduced a second measure, the so-called *m*-parameter, which is calculated by dividing

h by the number of years. For instance, this can be the number of years of scientific activity of researchers or the number of years of operation of facilities: $m = h/t$.

10.1   Caveats of citation analysis

Regardless of which bibliometric method is chosen, there are some general problems one should keep in mind.

- *Incompleteness:* by now, several web-based services have been developed that provide citation information. Some of them (e.g., Web of Science, Scopus) require a subscription, others (e.g., ADS Abstract Service, Google Scholar) are free-of-charge. The sources and technologies the services use to obtain and analyze the data vary, and citation statistics can be vastly different. Therefore, it cannot be emphasized enough that one single citation database will never give the full picture.

- *Incorrectness:* the sources of incorrect citation information can be manifold. A simple, but frequent, one is that authors make mistakes in their lists of references. Accordingly, these references cannot be resolved by the citation databases and are not assigned to the intended cited paper, i.e., the citation will be ignored. Another reason for errors lies in different abbreviations used for the same journal. References to each of them may then be assigned to seemingly different journals, and citations may be counted in parallel rather than cumulatively.

- *Citing behavior:* it should be remembered that many authors prefer to cite well-established scientists, those that are cited by everybody. There may be many reasons for this behavior: it may arise from a wish to "go with the flow" of what other authors do; it may save time to use reference lists from papers written previously; authors may have a wish to show that they read the "classics" in the field, or perhaps they want to show "solidarity" with friends and close colleagues. Regardless of the exact motivation, this creates a bias towards highly cited scientists and puts young and less connected authors at a disadvantage.

- *Multi-author papers:* authors who publish in large collaborations[34] by definition create a professional network of many colleagues who will (presumably) cite this paper in their future publications. Obviously, this leads to a much larger basis of potential quotes than papers published by single authors or small groups of co-authors will have. A way to handle references to such many-authored papers would be to assign only normalized (proportional) citation credits, i.e., 1 divided by the number of authors.

One more word of caution. It is very bad style to quote without actually having read the article, even though everyone in the field seems to quote it. One also would risk attributing something to the cited author that he or she did not actually write.

---

[34] In a recent paper, Crabtree (2008) found that the average number of authors has increased from 2.5 in 1980 to almost 7 in 2006. In the context of the ESO Telescope Bibliography, we noticed a 2008 Nature paper with no fewer than 93 authors.

10.2   Alternative measures

Citation analysis is an important part of bibliometrics, but it has one severe disadvantage: citations do not measure actual use: i.e., how often were papers consulted with or without subsequent inclusion in the final list of references?

A new measure to better reflect the use of publications has been introduced by the ADS. Instead of citations, readership information is being collected. A *read* is counted as each time a user requests information via one of the ADS letter links. In 2002, 50% of the requested items were abstracts, followed by one of the fulltext versions (38%) and citation lists (8%) (Kurtz et al. 2003). Using their database of logs, the ADS team can identify relations between the private act of reading a paper and the public one of actually citing it. The reads history is available for ADS entries from the record's abstract page.

A similar approach is pursued by Citebase[35]. This experimental citation database looks at *downloads* of papers from arXiv.org. As of January 2009, only the UK-site is monitored, but the other arXiv mirror sites are ignored, including the main server in the U.S. This means that the number of downloads per e-print displayed is notoriously too low. It will be interesting to follow the development of this project.

On a more general level, access and download statistics can also help for evaluating the use of journals and their overall acceptance within the community of scientists. Libraries occasionally consult such statistics to determine which journal subscription should be continued and which may no longer be necessary.

10.3   Telescope bibliographies

At many observatories, librarians are in charge of maintaining databases of papers written by staff or other astronomers that use data from the observatory's facilities. These databases are typically compiled by scanning the important astronomical literature (manually or semi-automatically), trying to locate references to instruments and telescopes that were used to gather the data.

There are many reasons for maintaining telescope bibliographies, for instance, to measure the scientific impact of telescopes and instruments. The performance of existing facilities can also provide important guidelines for future facilities. Funding authorities often request bibliometric studies because they help them to understand the overall productivity and impact of the funded organization.

Following the observatory's policies for acknowledging telescope use is also important for the authors. Some funding agencies demand that their support be acknowledged, and also many observatories ask (or even oblige) authors to mention use of the facilities in a footnote[36]. Mentioning facilities in the footnote will also help compilers of telescope bibliographies to keep track of the scientific papers published with their data and may influence observing proposal committees and other groups for granting future observing

---

[35] http://www.citebase.org
[36] For instance, the policy for publishing with ESO data can be found at http://www.eso.org/sci/observing/policies/publications.html

time. Last, not least, publishing with data from renowned facilities is also advantageous for authors as it increases their visibility and helps them to be recognized by colleagues.

A list of telescope bibliographies from major ground-based and space-based observatories is available from the ESO library[37]. In addition, the ADS provides access to several keyword lists, among them many observatories. On the main ADS search screen, scroll down to the Filters section and choose "Select References In / All of the following groups" and select the observatory of interest. Note, however, that the selection criteria, the years covered, and the range of completeness will differ. To find out the exact coverage, click on the facility's name, which will take you to a short explanation. As explained before, the D link of ADS records will lead to the data center or archive from where the data used in the respective paper can be retrieved.

11   Cooperations of librarians and astronomers

Traditionally, librarians have undertaken the role of mediator between information providers and scientists. Many librarians regularly inform the astronomers at their institute about new developments from publishers and database hosts, and report back any concerns, questions, as well as suggestions for improvements. You will know the librarian at your home institute, and if you are looking for a librarian at another institution, the Directory of Astronomy Librarians and Libraries[38] may be of use.

On a larger scale, there are some more formal ways to enhance communication between the various groups, in particular, an IAU Working Group and the LISA conferences.

11.1   IAU Commission 5: Working Group Libraries (WG Lib)

The Working Group Libraries is associated with the IAU commission 5 "Documentation and Astronomical Data." It aims at better cooperation between astronomers and librarians. The working group was officially recognized for the first time in 1990, but librarians were involved in Commission 5 activities even before that. Information about past and present activities can be found in the *Reports on Astronomy* (published every three years in the *Transactions of the IAU*), as well as on the working group's web pages[39].

During IAU General Assemblies, the Working Group often hosts so-called business meetings. These sessions are open to everybody interested and typically focus on data access and preservation, science metrics, and the role of librarians in fostering research.

11.2   Library and Information Services in Astronomy (LISA)

In 1988, the first conference on Library and Information Services in Astronomy was held in Washington, DC, at the U.S. Naval Observatory. Since then, LISA has been developed into a series of conferences, and four further meetings took place, each attended by

---

[37] http://www.eso.org/libraries/publicationlists.html
[38] http://www.eso.org/libraries/addresses/addresses.html
[39] http://www.eso.org/libraries/IAU-WGLib/

more than 100 participants. LISA conferences provide a unique meeting point for librarians, astronomers, publishers, and computer specialists from around the world. The wide geographic distribution of participants is indeed one of its main characteristics.

Thanks to the relatively small number of core journals, in many cases coupled with quite generous funding, astronomy libraries have often had access to evolving technologies earlier than other subject areas, and have applied them early on in their day-to-day work. In addition to changing technology, the professional role of librarians has undergone considerable development. This is reflected in the topics covered at conferences. Today, LISA is a forum to exchange experiences, view topics from different angles, and gather information about emerging fields of interest in astronomy libraries.

A description of the history of LISA meetings, their logistics, and topics can be found in Corbin & Grothkopf (2006); a web page provides links to LISA conferences, as well as proceedings of past meetings[40].

12 Conclusions

The publication paradigm continues to shift from printed material to electronic formats for scientific literature, leading to an evolved concept of library services and information access. Librarians are taking on new and diversified roles. Observatory librarians fulfill the information needs of astronomers and engineers and develop digital services to give greater access to scientific content. They support their library users when conducting research and provide help so that existing tools can be used efficiently.

In this chapter, we have introduced the varied research assistance libraries provide. Some tips and tricks were provided to help especially young astronomers to retrieve information from the ADS digital library and the arXiv (astro-ph) e-print repository. An overview was given of the most important astronomy journals and the concept of open access (OA) publishing. Because of the widespread habit of astronomers of posting manuscripts on astro-ph, OA affects astronomy and astrophysics less than many other subject areas, yet it should not be assumed that nothing remains to be done.

Scientists have gotten very used to the idea that all necessary information is available (typically free-of-charge) on the internet. However, there are some severe problems regarding long-term access, completeness of digitized material, and broken links, to name just a few. Librarians and information providers are working towards more reliable systems, but it is essential that scientists are also aware of these problems.

Increasingly, librarians maintain telescope bibliographies, i.e., databases of articles that use observational data. From these databases, publication and citation statistics are derived to provide information about the scientific output of specific facilities to management and funding organizations. Telescope bibliographies conclude the "life cycle" of observations because they close the loop from observing proposals to observations and to the published results, and from there back to the underlying data.

Finally, we have shown examples of close cooperation between librarians and astronomers that increase communication and, ultimately, lead to more efficient research.

---

[40]http://www.eso.org/libraries/list.html

Acknowledgments

References

Accomazzi, A. et al., 2007: Creation and use of citations in the ADS, in: *"Library and Information Services in Astronomy V"*, eds. S. Ricketts, C. Birdie, E. Isaksson, San Francisco, CA, Astronomical Society of the Pacific, ASP conference series 377, 93, arXiv:cs/0610011v1 [cs.DL], `http://arxiv.org/pdf/cs/0610011v1`

Boyce, P.B. 1998: Electronic publishing in astronomy, in: *"The impact of electronic publishing on the academic community"*, ed. I. Butterworth, London, Portland Press `http://www.portlandpress.com/pp/books/online/tiepac/session1/ch3.htm`

Corbin, B.G. & Grothkopf, U. 2006: LISA – the Library and Information Services in Astronomy conferences, in: *"Organizations and Strategies in Astronomy"*, vol. 7, A. Heck (ed.), Kluwer Academic Publishers, Dordrecht, 285 `http://www.eso.org/libraries/articles/lisaconferences.pdf`

Crabtree,D.R. 2008: Scientific productivity and impact of large telescopes, in: *"Observatory operations: strategies, processes, and systems II"*, eds: R.J. Brissenden & D.R. Silva, SPIE conference proceedings 7016, pp. 70161A-70161A-10

Dietrich, J.P. 2008a: The importance of being first: position dependent citation rates on arXiv:astro-ph, PASP, 120, 224 `http://www.journals.uchicago.edu/doi/pdf/10.1086/527522`

Dietrich, J.P. 2008b: Disentangling visibility and self-promotion bias in the arXiv:astro-ph positional citation effect, PASP, 120, 801 `http://www.journals.uchicago.edu/doi/pdf/10.1086/589836`

Grothkopf, U. 2003: From books to bytes: changes in the ESO Libraries over the past decade, The ESO Messenger, no. 112 (June 2003), 51 `http://www.eso.org/sci/libraries/articles/books2bytes.pdf`

Grothkopf, U. & Erdmann, C. 2008: Open access – state of the art, IAU Information Bulletin, 102, 64, `http://www.eso.org/libraries/IAU-WGLib/IAU_IB102p64.pdf`

Harzing, A.W.K. & van der Wal, R. 2008: Google Scholar: the democratization of citation analysis? Ethics in Science and Environmental Politics, 8, 62 `http://www.harzing.com/download/gsdemo.pdf`

Hirsch, J.E. 2005: An index to quantify an individual's scientific research output, Proc. Natl. Acad. Sci. USA, 102, 16, 569, arXiv:physics/0508025v5 [physics.soc-ph] `http://arxiv.org/pdf/physics/0508025v5`

Jasco, P. 2005: As we may search – comparison of major features of the *Web of Science*, *Scopus*, and *Google Scholar* citation-based and citation-enhanced databases, Current Science, 89, 1537, `http://www.ias.ac.in/currsci/nov102005/1537.pdf`

Kurtz, M.J., Eichhorn, G., Accomazzi, A., *et al.* 2003: The NASA Astrophysics Data System: obsolescence of reads and cites, in: *"Library and Information Services in Astronomy IV (LISA IV)"*, eds.: B.G. Corbin, E.P. Bryson & M. Wolf, U.S. Naval Observatory, Washington, DC, 223, `http://www.eso.org/sci/libraries/lisa4/Kurtz.pdf`

Kurtz, M.J. & Henneken, E.A. 2007: Open access does not increase citations for research articles from The Astrophysical Journal, arXiv:0709.0896v1 [cs.DL] `http://arxiv.org/pdf/0709.0896v1`

Mahoney, T.J. 2007: Correspondence to the Editors of "The Observatory", Obs., 127, 401, reprinted: IAU Information Bulletin, 2008, 102, 70 `http://www.eso.org/sci/libraries/IAU-WGLib/IAU_IB102p70.pdf`

Meho, L. 2007: The rise and rise of citation analysis, Physics World, January 2007, 32 `http://eprints.rclis.org/archive/00008340/01/PhysicsWorld.pdf`

Meho, L. & Yang, K. 2007: A new era in citation and bibliometric analyses: Web of Science, Scopus, and Google Scholar, subm. to Journal of the American Society for Information Science and Technology, 58, arXiv:cs/0612132v1 [cs.DL] `http://arxiv.org/pdf/cs/0612132v1`

Meylan, G., Madrid, J. P. & Macchetto, D. 2004: HST science metrics, PASP, 116, 790, `http://www.journals.uchicago.edu/doi/pdf/10.1086/423227`

Sandqvist, A. 2004: The A&A experience with impact factors, in: *"Organizations and Strategies in Astronomy"*, vol. 5. A. Heck (ed.), Kluwer Academic Publishers, Dordrecht, 197. Preprint at `http://www.astro.su.se/~aage/OSA5.pdf`

Schwarz, G.J. & Kennicutt, R.C. 2004: Demographic and citation trends in astrophysical journal papers and preprints, BAAS, 36, 1654, arXiv:astro-ph/0411275v1, `http://arxiv.org/pdf/astro-ph/0411275v1`

Suber, P. 2007, Open access overview: focusing on open access to peer-reviewed research articles and their preprints, `http://www.earlham.edu/~peters/fos/overview.htm`

Viele, P. 2008: Information fluency and physics curriculum: faculty/librarian collaboration, APS Forum on Education Newsletter, Summer 2008 `http://www.aps.org/units/fed/newsletters/summer2008/viele.cfm`