BITS AND BYTES AND STILL A LOT OF PAPER: ASTRONOMY LIBRARIES AND LIBRARIANS IN THE AGE OF ELECTRONIC PUBLISHING

This article was originally published in Astrophysics and Space Science, **247**: *155-174, 1997;* ©*Kluwer Academic Publishers*

U. GROTHKOPF
European Southern Observatory
85748 Garching, Germany
esolib@eso.org

1. Introduction

The future has already begun. The information superhighway, hypermedia, digital libraries and electronic publishing are not vague concepts anymore that might be awaiting us beyond the year 2000. Instead, they are already here. Each day our electronic mailbox is flooded with announcements about new sites on the World Wide Web that vie for our attention. More and more information resources are easily accessible and need to be checked out. Modern communication technology has brought us so close together that all the information available anywhere on this planet seems to be at our fingertips. This is the Information Age.

For us librarians, these are not only extremely fascinating, but also very challenging, times. Librarians have been information providers for centuries, and nowadays we have the opportunity to use tools that allow us to provide an even quicker, more complete and sometimes more sophisticated service to our users. Databases and reference sources can be queried via the Internet, library catalogs are available remotely, newsgroups and mailing lists provide a wonderful opportunity to discuss mutual concerns quickly, and electronic mail allows us to contact colleagues who might be able to help us solve a problem. And all this can be done within a minimum of time.

Technology will take us one step further very soon. It is no longer only references and information *about* documents that is provided electronically, but also full texts of publications that have become available in digitized form. If we look for information about a particular topic today, we can start for instance by querying a library catalog on the World Wide Web (WWW).

The results we receive might consist of some formal bibliographic data like authors, title, and publication date. Some contents-related information (keywords, thesaurus descriptors, or the contents table) may be included as well. Plus, we may have the opportunity to access the abstract or even the entire text, as well as graphics, tables, video sequences, or an embedded software program that can be run locally. The references at the end of the publication are active links. One mouseclick, and we move on to the referenced paper without even noticing where exactly we are going on the network. Another mouseclick, and a search engine finds other publications dealing with similar topics that also might be of interest to us. Particularly useful documents can be printed or downloaded onto our own computer for later use. The search strategy we chose to find these publications can be stored and reused whenever needed.

Access to information is becoming more and more seamless. The year 1997 is a milestone in astronomical publishing, as some of the most important publishers recently launched electronic versions of their journals or intend to do so in the course of the year. However, various barriers currently impede immediate access to every piece of information. In the following, we will look at the many advantages of electronic publications, but will also focus on some of the questions and unresolved problems.

Change in general, and technological change in particular, has its price. Librarians are worried about some of the current developments and therefore speak up if short-sighted decisions are being taken. Neil Postman, a North American communications theorist, stated in his book "Technopoly" that

> [...] it is a mistake to suppose that any technological innovation has a one-sided effect. Every technology is both a burden and a blessing; not either-or, but this-and-that. (Postman, 1992, pp. 4-5)

It is necessary to be aware of both the positive and the negative implications in order to make the most efficient use of technology.

2. From Traditional to Digital Libraries

The tools used by librarians in their daily work have changed vastly during recent years. Today, hardly any library is equipped exactly as it was only a few years ago. In addition to traditional means like card catalogs and microfiche readers, most libraries now also offer an online public access catalog (OPAC), public PCs equipped with CD-ROM drives, scanners, or public terminals connected to the Internet. An increasing number of libraries are building homepages on the World Wide Web from where users have access to a variety of services without physically entering a library.

Many libraries are in transit from the traditional towards the digital library. We witness a shift from libraries offering information about (electronic and print) information towards providing access to full texts of documents. Not only recent publications, but also many historical library holdings are being digitized (see e.g., Corbin and Coletti, 1995). These electronic collections allow users from everywhere at any time to consult the material without doing any harm to fragile documents.

Despite numerous digitization projects, electronic media by no means are dominant compared to print material. There is still a lot of paper in our libraries, and we expect this to be the case for a long time to come. The paper-based library will co-exist with the digital library for the foreseeable future, because electronic publications are not developing at the expense of print media, but in addition to them.

The notion of library has long expanded beyond the physical building of the library. Our services always included access to sources that are physically located outside the library. Over the course of the years, librarians have collaborated in many ways. Central cataloging, union lists of journals, cooperative collection development and interlibrary loan are only a few examples of resource sharing. Forced by decreasing budgets, many libraries have redefined their acquisitions policy from purchasing documents "just in case" to "just in time", since no library can afford to purchase every item that might be needed by one user one day. Through collaboration and reciprocal services among libraries, we can provide a much larger range of resources to our users and fulfill their information needs quicker, cheaper, and more completely than one library alone would be able to do.

While projects that aim at helping each other might be seen as a nicety during prosperous years and become a necessity in times of economical restraints, they play an ever more essential role in the electronic environment. James Michael suggested a blueprint for the library without walls that consisted of five elements (Michael, 1994):

1. interconnectivity – connecting to a network
2. interoperability – the ability of one computer to talk to another
3. integration – of internal and external resources into one single user interface
4. intermediation – reference services, navigational help and instruction provided by librarians
5. interdependency – because one single library cannot own all the resources that might ever be needed by users

This last item, interdependency, is the final step for the "Global Digital Library" to become reality. In the electronic environment, even more than in the traditional paper-based world, no library can (or may) store all the documents to which it provides access. Digital libraries are only possible if reliable partners cooperate on a long-term basis. Authors, libraries, publishers, archives – the concept of one player in the electronic publishing sector as a self-sufficient entity has been overcome for good. The digital library indeed brings us closer together than ever.

3. Electronic Publications: Impact on Library Functions

3.1. ELECTRONIC VERSUS PAPER-BASED PUBLICATIONS

Without doubt, electronic publications in general, and electronic journals in particular, provide some advantages compared to paper-based documents. Electronic documents typically are delivered with a powerful mechanism to search easily through them. Navigation tools allow readers to jump to particular sections, e.g., to references or graphics, and from there to the relevant section in the body of the text. "Forward referencing" provides links to articles which were published later, but cite the original article. Corrections can be included without difficulty. References can be linked to abstracting services, from where abstracts or full texts of cited papers can be obtained, and similar publications can be retrieved based on the original article. Electronic versions of journals usually are available in advance of printed versions, and in addition users can browse the contents tables of forthcoming issues. Journal issues need not be shipped, thus avoiding delays due to mailing systems. Electronic documents can be accessed from anywhere at any time and by as many simultaneous users as needed.

New tools will go beyond these features soon. Graphics will be available together with the underlying data, allowing readers to modify them according to their needs; video or sound sequences as well as computer programs can be embedded; journal issues can be linked to data archives, image libraries, laboratory measurements, or software collections (Boyce, 1996), or users may be able to launch demonstrations or other virtual experiments.

On the other hand, electronic publications, at least if used with today's technology, are bad for browsing and nearly unacceptable for reading online. Interesting documents are likely to be printed, probably even several times, at the users' sites, thus shifting printing costs from the publishers to the scientists' institutions. Further costs arise because special tools, both hardware and software, may be required for downloading, reading, and using electronic publications. Unless a text-only version of electronic publications is provided, terminals need to have graphical capabilities. Specific versions of Web browsers need to be installed in order to correctly display the documents. Appropriate software might be necessary to store or print electronic publications. Bandwidth is unlikely to keep up with demand at least in the near future (Butler, 1996); if network capacities are not sufficiently high, access to documents can be painfully slow or impossible. Even with powerful connections, loading of individual documents on the Web can be extremely time-intensive, for instance if the providers include extensive graphics. Many network users try to work around this problem by simply switching off the display of graphics in their Web browsers.

Costs for electronic publications are not limited to purchasing equipment. The myth of all information on the Internet being free-of-charge has not yet been destroyed. In the first days of the net, many documents were maintained by individuals who spent their private time and personal efforts into setting up network resources. This voluntary attitude has led to the widespread, but wrong assumption that electronically delivered information doesn't cost anything. The truth is that the Internet is just the communication channel; valuable content must be paid for as much as before. Commercial publishers are about to leave their experimental phase during which access to their publications often was allowed without any charge, for testing purposes, and now are establishing subscription rates. Combined subscriptions for print copy plus access to the electronic version of journals typically are more expensive than print-only subscriptions. While the reasons for price increases may be understood by librarians, it can be very difficult to explain them to those who decide on the size of library budgets. As a consequence, we have to fight even harder for our budgets now.

3.2. PROCESSING ELECTRONIC PUBLICATIONS IN THE LIBRARY

Boyce and Dalterio (1996) pointed out that electronic publishing means different things to different people. With regard to the classic publication model author–publisher–library–reader, they identified six phases of electronic publishing: author preparation and submission of the electronic manuscript, peer review, copy editing and typesetting, database preparation, production and distribution, and archiving. To librarians, access to and distribution of electronic publications as well as archiving (which will be dealt with in section 3.5.) are of obvious importance.

While book-like electronic media like CD-ROMs and diskettes are processed in libraries very similarly to print media, remotely stored electronic publications bring upon us changes that effect all phases of collection development. Selection criteria for purchasing material probably remain more or less as before (Nisonger, 1996), but it might be more difficult to obtain evaluative information (e.g., book reviews) about electronic material. In January 1997, the first journal devoted to providing such evaluative information was launched[1] and may find a lot of approval among acquisitions librarians.

[1]Electronic Resources Review, http://www.mcb.co.uk/liblink/err/jourhome.htm

Once the decision to purchase (or lease) electronic publications has been made, there are often technological implications in trying to access the documents. Special software might be necessary that needs to be properly installed on local PCs or made available through the network. The library needs to provide public terminals and printers to ensure that patrons visiting the library can access the material. Because of the extensive use of the World Wide Web in the scientific community, it is hoped that a standardized way of providing electronic documents (i.e., in SGML (Standard Generalized Markup Language) or HTML (Hypertext Markup Language) format) will evolve that will overcome the various and confusingly different software packages currently used by publishers.

Cataloging needs to include all the bibliographic details catalog records for paper-based documents contain, plus information about access procedures, electronic format, size of files, hardware and software requirements and other technical details. Most libraries that have automated their catalog use the MARC (MAchine Readable Cataloging) formats for bibliographic representation. MARC currently is being enhanced by a subset of fields that cater to additional descriptions of electronic publications, including the electronic address of the document. Library catalogs with a hyperlinked World Wide Web interface can then provide access to electronic publications from within the catalog just by clicking on the address.

Access control can be taken care of by the publisher through IP (Internet Protocol) addresses, or it can be left as the responsibility of the subscriber (i.e., the library) to inform authorized readers about the user ID and password. The latter solution is not preferred by librarians, as it is nearly impossible to distribute a password to the library's user community and at the same time avoid any possible misuse by persons not entitled to access.

Costs of electronic resources will be difficult to calculate from now on. Publishers have come up with a large variety of new pricing strategies (see Brown and Duda, 1996); "pay per view" and "flat fees plus additional charges according to usage" are just two examples. Pricing models are still in flux and changes to current solutions are inevitable. In addition, different libraries may have to pay different prices for journal subscriptions. Some publishers have started to offer special conditions and reduced prices to consortia of libraries. Participants need to pay only a portion of the total subscription costs, but gain access to all the publisher's journals. For large university library systems, this model might be advantageous, but for small individual libraries it can be difficult or impossible to participate in a consortium for administrative, organisational or political reasons.

Instead of owning the material, there is a clear shift towards providing access to electronic documents (electronic journals in particular) just for a given time. Libraries are leasing electronic material rather than purchasing it. The necessary contracts, so-called license agreements, raise another set of questions (see e.g., Soete, 1997): What exactly are the rights and responsibilities of the parties of the contract? Are there limitations regarding simultaneous accesses, downloading or printing of material? What are the technical requirements? Is the library entitled to access the user statistics in order to evaluate usage or does the publisher keep these data secret?

One of the most critical parts of license agreements concerns copyright. Publishers from many countries currently allow libraries to take photocopies of their print publications based on "fair use", i.e., if it is for personal, not commercial, use. This naturally includes interlibrary loan, but ILL can be restricted or even prohibited by license agreements. A typical license currently favors the information provider over the information user (Okerson, 1996a). If the restrictive clauses are not eliminated from the contracts, they will disable traditional and necessary cooperation among libraries. A positive exception is the text of the American Astronomical Society (AAS) contract. The AAS has included a librarian in their Publications Board for several years and the license agreement they have set up for their electronic journals is exemplary in its fairness and thoughtfulness[2]. The problem of negotiating contracts for electronic documents has led to extensive discussion among librarians. Documents that guide librarians through the jungle of licensing, for example the "Liblicense" resource[3] are necessary and very welcome. A newly established mailing list devoted to the topic of electronic content licensing, liblicense-l[4], received some 800 subscriptions within only a few days of its inception which clearly shows that discussion and advice on this topic are sought by many librarians.

It is the nature of leased material that it remains physically with the provider after the contract has ceased. Librarians must be aware that they may be left without anything when a license agreement terminates. This situation is unlike the print environment where a library at least owns the journal volumes that were published while a subscription lasted. Some journal publishers deliver backups of electronic journals on CD-ROMs which stay in the library even after cancellation of the subscription. In addition to the problem of whether or not CD-ROMs will be a viable long-term storage medium, they are not capable of presenting one of the major advantages of electronic over print journals, i.e., they are not interlinked with other

[2] The Astrophysical Journal and Astrophysical Journal Supplement Series Institutional Site License, URL http://www.journals.uchicago.edu/ApJ/sitelicense.html

[3] Liblicense: Licensing digital information. Yale University Library, URL http://www.library.yale.edu/~Llicense/index.shtml

[4] To subscribe send a message to listproc@pantheon.yale.edu, leave the subject line blank and type in the body of the message: subscribe LIBLICENSE-L Firstname Lastname

journals or databases. Backup on CD-ROM therefore is not an appropriate alternative. The only fair solution would be to allow libraries access to those volumes they subscribed to.

3.3. ELECTRONIC BOOKS AND NON-SERIAL DOCUMENTS

Everybody who enjoys the calming sound of turning book pages most probably dislikes the thought of having to do with a computer mouseclick instead. Today's technology is advanced, but not yet advanced enough for most of us to really want to read electronic books. The general market seems to reflect this, since it has not developed as steadily as it was predicted only a few years ago. The 1996 Frankfurt Book Fair, one of the world's largest international book fairs, saw a smaller percentage of multimedia products than anticipated (Baier, 1996). Klaus-Dieter Lehmann, Director of the Deutsche Bibliothek stated in an interview that only 2,000 to 3,000 out of 300,000 new acquisitions per year are digital, everything else still is on paper (Lehmann, 1996).

Electronic books do have some advantages though. They allow users to quickly locate specific information through built-in search functions, and information can be updated and kept accurate easily. Electronic format therefore is a valid alternative to print media, especially with regard to directories, dictionaries, encyclopedias, and reference works.

In astronomy, extensive use of the capabilities of e-books has been made in regard to user manuals for telescopes or individual instruments. These days the latest information often can only be found on the Web. For conference announcements and programs, it has become standard for users to obtain up-to-date information from a Web page. Likewise, more and more conference proceedings can be found on the Internet (e.g., "Weaving the Astronomy Web (WAW)"[5], "Library and Information Services in Astronomy II (LISA II)"[6] or the "Astronomical Data Analysis Software and Systems (ADASS)"[7] conference proceedings). A novelty in astronomy was the book "Information & on-line data in astronomy" (Egret and Albrecht, 1995); in addition to the paper version, an electronic complement is maintained on the World Wide Web, from where chapter abstracts are available. Hypertext links mentioned in the book are provided as clickable links.

On the other hand, electronic books and documents have specific disadvantages. With regard to networked documents, version control poses a distinct problem. We are used to informing our users about the edition of a book or the publication year of a thesis, and we know that the document

[5] WAW, http://cdsweb.u-strasbg.fr/waw/proceedings.html
[6] LISA II, http://www.eso.org/lisa-ii/lisa-ii.html
[7] ADASS, http://iraf.noao.edu/ADASS/adass.html

has not been changed since it arrived on paper in our library. In contrast to this, information about publication dates and versions of documents that are accessed over networks are extremely short-lived. The responsibility for these documents remains with the author or publisher and thus they can be changed by them, if deemed appropriate. The dynamic nature of networked documents leads to the dilemma of up-to-date information versus access to the original version. Von Ungern-Sternberg and Lindquist (1995) pointed out that another concern is the authenticity or integrity of electronic documents. Once a document is stored on a computer, it can be relatively easily manipulated by additions, deletions or changes. "A number of 'digital signatures' or 'electronic seals' based on cryptographic techniques" may help to ensure authenticity. If electronic books are delivered offline (e.g., on diskettes, CD-ROMs or their successors), the difficulties of version control and validation of authenticity diminish. However, offline publications cannot be updated as easily as networked documents, and they don't exploit the capability of interlinking with external resources. Instead, hypertext links are used to navigate within the document or in order to make use of built-in multimedia effects like sound, video, and simulations.

Although an unlimited number of copies can be downloaded or printed from electronic books, assuming the electronic version is free-of-charge or the access fee has been paid, nevertheless electronic books can become unavailable. Out-of-print material on paper almost always can be located in some way, most importantly through interlibrary loan, and therefore hardly ever disappears entirely. Online documents, in contrast, might be withdrawn by the authors or publishers and may never be found again.

On the Internet, it is astonishingly easy to become a "publisher". Everybody equipped with a networked terminal and a server has the opportunity to distribute information globally without any difficulty. While this can be seen as moving towards more democracy in a general context, it is a threat in scientific publishing, because quality standards are likely to decrease. Algorithms that try to match users' requests with available documents do not solve this problem, because "prioritizing according to relevance won't give you a true measure of quality" (Schwartz, 1996). Commercial publishers and learned societies therefore argue that a careful editing process is necessary to assure that only those documents which would pass the quality control for paper-based publication are published electronically. Unfortunately, such editing is one of the major cost factors of journal production which will not disappear even if the journal is available exclusively in electronic format.

3.4. ELECTRONIC JOURNALS

We are still in the early days of electronic journals. This phase brings upon us a variety of formats and delivery methods. Formats include bitmaps, PostScript, PDF (Portable Document Format) as well as ASCII (American Standard Code for Information Interchange), SGML and HTML. Electronic journals might be delivered on CD-ROM, by electronic mail, or through the networks. Networked journals require that users actively access the publishers' sites; for convenience of the subscribers, new issues often are announced through e-mail notifications. Although publishers are still experimenting with various formats, the extensive use of the World Wide Web certainly will lead to a future of electronic journals that is network-based (Wusteman, 1996).

The increase in the number of so-called electronic journals is breathtaking. In the introduction to the 1996 edition of the "Directory of Electronic Journals, Newsletters and Academic Discussion Lists", Ann Okerson noted that "in this sixth edition, the number of journal and newsletter titles (nearly 1700) has more than doubled since last year's and multiplied by over 15 times since the first edition (there were 110 listings in 1991, 240 in 1993, and nearly 700 last year)" (Okerson, 1996b). Most of these journals are not "true electronic journals" though. In a widely distributed survey article by Steve Hitchcock, Leslie Carr and Wendy Hall (1996), the authors point out that currently most of the electronic journals are rather electronic versions of paper journals than fully operational e-journals; that is, they are confined to presenting bitmapped images or electronic formats intended for printing, like PostScript or PDF. In order to fully exploit the capabilities and advantages of electronic journals, they must be able to interlink sources and therefore need to go beyond print-oriented format as soon as possible.

The acceptance of electronic journals among the scientific community currently is not as high as the large number of electronic versions seems to indicate. According to Ann Schaffner (1994), in order for e-journals to be accepted by authors, it is necessary that they fulfill at least the basic functions of paper-based journals, which she defines as

- Building a collective knowledge base
- Communicating information
- Validating the quality of research
- Distributing rewards
- Building scientific communities

It is obvious that these criteria are met by electronic versions of established print journals, but print-oriented formats only will be helpful in achieving acceptance during the transition phase from paper to electronic. Readers of a (paper) journal might accept more easily the electronic edition if it has the same "look and feel". Currently many scientists rely on paper, at least for reading full text articles. Interviews with chemists at Cornell University in 1995 showed that the possibility of creating a print copy determined to a large extent the scientists' acceptance of electronic journals (Stewart, 1996). However, in the long run "authors are unlikely to accept 'electronic clones' of paper journals" (Ginsparg, 1996).

Accessing electronic publications over the Internet can be unacceptably slow, and incorrect or out-of-date electronic addresses can turn locating documents into an extremely time-consuming task. Reliable links are the backbone of electronic services. A study by Harter and Kim (1996) revealed that 55% (71 out of 129) of the links to e-journal archives the authors tried to access did not work at first try. If electronic journals are to be effectively used, the reliability of electronic references must stabilize.

Harter and Kim also analyzed to what extent authors of e-journal articles (in whom one can assume a sympathetic attitude towards electronic media) are citing electronic journals and other online resources (electronic personal papers, newsgroup postings, electronic preprints etc.). Their findings showed that only 1.9% (83 out of a total of 4,317) of the citations in peer-reviewed articles published in 1995 referred to electronic sources. Indexing services can help increasing the acceptance of electronic publications but they are just beginning to include e-journals in their databases. Electronic journals must struggle to meet the criteria for inclusion applied by abstract services in order to be eligible.

One of the most important problems of electronic publications is archiving. Unless a stable system for future retrieval of publications is in place, authors will continue to be reluctant to publish in electronic journals.

3.5. ARCHIVING

In order to guarantee continued access to information, publications have been archived throughout time. Traditionally, preservation has been a task of libraries, which they have performed reliably through the centuries. Up to now, even small specialized libraries were able to provide a highly valuable repository of contemporary and historical publications. Their holdings could be unique within their geographical region.

We have seen many media tested for and applied to archiving, for example paper, microfilm, microfiche, CD-ROM, and magnetic tape. Of all these, paper still is the preferred storage medium for many librarians, since it is the only format that can be used without any special viewing devices and therefore is independent of current and future technology. In addition, no other storage medium up to now has survived for as long as paper.

Electronic publications turn this ordered situation upside-down. They unfold their most valuable features when they are interlinked with other documents, archives, and databases. Thus, if we regard electronic publications not as self-contained, off-line documents, but as parts of a large system, we must conclude that they need an archiving system that is based on a global concept and goes beyond our current model.

A new physical storage medium (e.g., magnetic tape, networked on computers) as well as an appropriate data format (e.g., SGML) have to be selected (Wusteman, 1996). It is obvious that current technology soon will be replaced by newer versions or entirely new systems, therefore both the physical storage medium and the data format must be sufficiently flexible to be transferred to the next technological generation. This transfer is being referred to as "technology refreshing" or, as an even broader concept, "technology migration" (Garrett and Waters, 1996).

The rapid pace at which technology becomes obsolete makes archiving an extremely expensive undertaking. Libraries' budgets are highly unlikely to rise with the necessary speed (in fact, they are unlikely to rise at all). Librarians may find themselves being forced to abandon archiving, one of their traditional tasks. New volunteers appear on the scene, some eager to take over. In addition to non-profit organizations, commercial institutions and subscription agencies participate in the debate. Many librarians are skeptical about whether the interest of commercial organizations in archiving will go beyond immediate economical consideration. We also wonder whether we will have to pay twice for journals in the end – once for the license, and again for access to back issues.

In contrast to commerce-oriented solutions, the Task Force on Archiving of Digital Information, established by the Commission on Preservation and Access and the Research Libraries Group, suggests in its final report that a national system of digital archives should be developed (Garrett and Waters, 1996). Digital archives in their sense are "held together in a national archival system through the operation of two essential mechanisms": the repositories must meet or exceed criteria set by a program for archival certification, and a "critical fail-safe mechanism" must be in place that can carry out rescue functions when needed. This report focuses on the United States and although such a well-organized, nation-wide archival system will be difficult, maybe even impossible, to achieve for many countries, the Task Force envisions a possible expansion on an international level.

International standards for archiving electronic publications in conjunction with a stable and mature network environment would be the ideal solution, but it will take some time before the necessary infrastructure will be in place. In the meantime, we must find other solutions (see e.g., Neavill and Sheblé, 1995). We have reached a critical point in time, because technology has become so advanced that publishers are tempted to abandon the traditional archiving medium before actually being able to come up with another long-term solution that is appropriate for the digital age. If the mere availability of technology leads to short-sighted solutions, it will have expensive, if not fatal implications for future access to scientific literature. In this regard, the opportunities given to us by technology clearly are a danger. In order to ensure uninterrupted access to publications, archiving must continue on paper until the question of infinite preservation is fully clarified.

4. Indexing and Retrieving Electronic Publications

Locating information on the Internet seems to be so easy. Every publication on the network has its proper address, and all we need to do in order to retrieve it is to point our Web browser to that address.

From their daily work, librarians know that the reality is different. In paper-based publications, we have come across many strange or wrong citations over the course of the years. If this is the case in the relatively standardized and agreed-upon print environment, it is unlikely that this situation will improve when it comes to electronic publications. The current naming system for networked documents relies on a chain of words, letters, or figures, subdivided by commas, dots, slashes, hyphens or other characters. The so-called Uniform Resource Locators (URL) are extremely liable to errors. One lower-case letter instead of a capital one will result in an error message.

Librarians and publishers are experimenting with better naming systems. Uniform Resource Locators are already developing into Uniform Resource Names (URN), which allow a unique name to be assigned to the publication that will not be changed even if the document is moved to another computer. A name resolver will keep track of the actual location of the requested file and translate the name into the correct URL. The American Astronomical Society makes use of this approach in their electronic journals (Warnock and Fullton, 1996). OCLC (Online Computer Library Center, Inc.) has been testing Persistent Uniform Resource Locators (PURLs)[8] that also point to an intermediate resolution service instead of directly to the location of an Internet resource.

These mechanisms are not yet commonly used, and librarians have to be imaginative in order to locate requested documents when users present to them strange network addresses and weird names of electronic publications. Our profession demands that we always try more than just one approach in order to locate information. If the most obvious way does not lead to

[8]PURL, http://purl.oclc.org/

satisfactory results, we think of alternatives. This attitude is being extended to Internet resources where usually several search strategies are possible.

For many years, librarians have been the experts within their institutes with regard to searching commercial databases. They often even held a monopoly, because they were the only ones who knew which retrieval language had to be used for which database and how required information could be obtained most cost-effectively. It is obvious that searching the seemingly chaotic Internet is completely different from well-organized bibliographic or full-text databases and therefore requires new methods. The most commonly used resource discovery tools are search engines that allow searching the Internet for words or phrases. The drawback is that enquiries typically result in a huge number of documents with extremely high noise and little precision. Projects have been set up in order to improve search results. What is needed is information about information, so-called metadata. OCLC and NCSA (National Center for Supercomputing Applications) are two of the main initiaters who proposed the Dublin Core Metadata elements set which is "intended to describe the essential features of electronic documents that support resource discovery" (Weibel *et al.*, 1996). The 13 Dublin Core elements include, for instance, information about the subject, title, authors, form, etc. and can be generated by authors of Web documents without extensive training. "Indexing the Internet" sounds like a mission impossible but the achievements so far are very promising. If some remaining problems can be solved, it is quite possible that the Dublin Core will be implemented in one of the future versions of HTML. Mapping between MARC and the Dublin Core also is in preparation so that data interchange between catalog records and SGML documents will be possible (Dempsey and Weibel, 1996, and Weibel, 1995)[9].

This approach to information retrieval still involves human mediation. In a recent article, Bruce Schatz (1997) explains how new concepts may allow users to search through distributed repositories across the net. He foresees that automatic indexing and "vocabulary switching" (i.e., terminology being automatically "translated" into the appropriate vocabulary of different subject domains) finally will allow us to "effectively utilize the whole of scientific information." Although such systems are still under development, one of the first applications can be seen already in the Illinois Digital Library Initiative[10] in which the AAS is also about to participate.

[9]For more information on metadata see also: IFLA, Digital Libraries: Metadata Resources, URL http://www.nlc-bnc.ca/ifla/II/metadata.htm and Library of Congress: Metadata, Dublin Core and USMARC: a review of current efforts, Discussion Paper No. 99, URL gopher://marvel.loc.gov/00/.listarch/usmarc/dp99.doc

[10]Digital Library Initiative project at the University of Illinois at Urbana-Champaign, http://dli.grainger.uiuc.edu/

5. The Role of Librarians in the Electronic Environment

The mission of astronomy librarians typically is to provide access to selected information resources according to the needs of our users, to collect and preserve material, to organise it, and to deliver it to the requester. This mission statement refers to all available media and document formats and therefore is not limited to particular technologies or specific software applications. Librarians are system-independent information specialists, aiming at fulfilling the information needs of library users. Our experience indicates that user needs and expectations are very similar in the paper-based and in the electronic environment:

> Users' needs will continue to be what they long have been. Users will want information reliably locatable [...] (and) easily accessible [...]. In the electronic environment the need for access tools will be more evident [...]. Users will expect information to be available that was placed in the library's care a long time ago; and they will expect that the integrity of the information they get from the library will be assured. (Graham, 1995)

As in the paper-based environment, digital services must be "planned, implemented, and supported" (Hastings and Tennant, 1996), and instead of being the "gatekeepers" to material, librarians now follow new directions which were identified by Ching-chih Chen (1994) as follows:

- from library-centered to information-centered
- from the library as an institution to the library as an information provider, and the librarians as a skilled information specialist functioning in an all-related information environment
- from using new technology for the automation of library functions to utilizing technology for the enhancement of information access and delivery of items not physically contained within the four walls of the library
- from library networking for information provision to area networking for all types of information source providers

Chen continues to add another important aspect, namely "from access to selectivity". It is not lack of information we have to cope with, but an information overload that needs to be mastered. "End-user searching" is a frequently used term today, since (at least seemingly) user-friendly retrieval technologies are evolving. However, recent studies revealed that "the preferred method of receiving information by many of our clientele is person to person" and that "service is personal" (see Enyart and Smith, 1996). A report about the University of California Digital Library draws the same conclusion: "Experience indicates that, despite the availability of intelligent systems, increasing remote access will also increase demands for service, both online and face-to-face" (University of California Library Council, 1996).

Since the beginning of our profession, librarians have been in direct contact with users seeking information. We have learned to understand what they need, not what they say they need (which can be considerably different). Our services are personalized and targeted for our clientele, and up to now no "interface agents" and "personal filters" as described by Nicholas Negroponte (1996, pp. 149-159) are in place that are able to substitute person-to-person mediation.

> Technology for producing and distributing information is useless without some way to locate, filter, organize and summarize it. A new profession of "information managers" will have to combine the skills of computer scientists, librarians, publishers and database experts to help us discover and manage information. These human agents will work with software agents that specialize in manipulating information – offspring of indexing programs such as Archie, Veronica and various "World Wide Web crawlers" that aid Internet navigators today. (Varian, 1995)

What Hal Varian overlooks is the fact that this profession exists already. The "human agents" he mentions are the librarians of today, no matter whether they call us information navigators, information officers, cybrarians, or librarians, a term most of us still prefer (Ojala, 1993).

In the electronic environment, our role comprises at least three major working areas: We offer services for those users who want to be guided to the most suitable information resources; we provide research assistance for those who prefer to conduct searches themselves and only turn to librarians in case their repertoire of search and retrieval techniques did not lead to satisfying results; and we closely collaborate with Information Technology (IT) departments that design new and enhance existing systems. These working areas encompass the following aspects:

1. Information Access Provider:
   We provide access to the most important information resources on the network, making use of current technologies (WWW or its successor). Electronic resources must be organized in a logical, easily understandable manner, integrating documents and services that belong logically together. The purpose of value-added services like subject-oriented clearinghouses is "not only to save the researcher time and effort in searching for appropriate sources in the vastly unordered, unstructured Internet, but also to provide him or her with a pre-assessed, semi-ordered, annotated list of sites with activatable links" (Rusch-

Feja, 1997) which match or supplement the targeted group's information needs.

2. Research Assistant:
   The second aspect of our role encompasses identifying, locating and obtaining publications not owned or leased by the library. No resource, be it electronic or on paper, can be called "complete" today, and if we don't find particular information on the Internet, this does not mean that it doesn't exist. It is the librarian's duty to know which additional sources can be queried to make a search as complete as possible. Publications "not imbedded in a formal journal context" (Rusch-Feja, 1997) as well as electronic equivalents to today's "grey literature" need to be retrieved, which requires the librarian's experience in locating information as well as technological skills. Research assistance in this sense also includes helping our users to become familiar with handling new technologies. Various methods for providing user support for networked library services can be developed, for instance distance support (by telephone or e-mail), printed or online manuals, and on-screen instructions (Mackenzie Owen and Wiercx, 1996). Face-to-face end-user training already has become an important part of our work.

3. Collaborative System Designer:
   Librarians are in direct contact with users of information retrieval systems. Often users report difficulties to us which they encountered while using a system, or we recognize what needs to be redesigned while we are explaining a system to our users. When IT departments design new databases, application programs and user interfaces for us, we must be able to explain precisely to them "how language works and how to use layout, typography and design principles" (Moore, 1996) in order to provide the required functionality. The information flood can only be mastered with appropriate tools that are capable of matching user needs with the available information, no matter whether these tools will be operated by librarians or users in the end.

Access to global networks and the availability of information to scientists through these networks have changed vastly how science is done today. Despite the many advantages of electronic publishing, there are still many problems to be solved. The information overload results in a need for specialists who are experienced in retrieving information and therefore can guide end-users through the jungle of available resources and technologies. In the Information Age, professional librarians are likely to be more in demand than ever.

6. Conclusion

As we are heading towards the 21st century, we are witnessing an information revolution. The importance and value of information, delivered in a timely manner to the requester, is steadily growing. New technologies have changed dramatically the ways in which information can be obtained.

Astronomy librarians are in the fortunate situation of having tools in their hands that provide search and navigation capabilities of which we wouldn't have dreamed only a few years ago. On the other hand, the current situation also confronts us with a large number of questions. Librarians are not interested exclusively in accessing the very latest (pre-)publication, but need to make sure that scientific documents will be available throughout time. Many librarians are alarmed by some of the directions currently suggested. Mature and standardized solutions for retrieving and accessing and in particular for archiving electronic publications must be in place before the current, paper-based system can be abandoned.

Librarians traditionally have been on the forefront of making information available to users. They will continue to be so, but the means and methods used in order to achieve this aim are undergoing rapid changes. The role of librarians in the new publishing scenario has been expanded, it involves new tools and technologies, and reflects changed requirements and new end-user behavior.

Acknowledgements

I am very thankful to Ellen Bouton (National Radio Astronomy Observatory, Charlottesville), Marlene Cummins (University of Toronto Astronomy Library, Toronto), and Sarah Stevens-Rayburn (Space Telescope Science Institute, Baltimore) who provided a world of help through thoughtful comments and suggestions, as well as through careful proofreading (that was how I learned that the four of us can be called "fuss-buckets"). Special thanks go to Carolina Noreña Janssen (Museo de Ciencias Naturales, Madrid) for extensive discussions and helpful comments.

References

Baier, H. (1996) Frankfurter Buchmesse: Seismograph für Trends, *Buch und Bibliothek* **Vol. 48 no. 12**, pp. 895-896

Boyce, P.B. and Dalterio, H. (1996) Electronic publishing of scientific journals, *Physics Today* **Vol. 49 no. 1**, pp. 42-47

Boyce, P.B. (1996) Building a peer-reviewed scientific journal on the Internet, *Computers in Physics* **Vol. 10 no. 3**, pp. 216-221

Brown, E.W. and Duda, A.L. (1996) Electronic publishing programs in science and technology. Part 1: The journals, *Issues in Science and Technology Librarianship* **Fall**, URL http://www.library.ucsb.edu/istl/96-fall/brown-duda.html

Butler, D. (1996) Internet growth - boom or bust? *Nature* **Vol. 380**, 4 April, p. 378

Chen, C. (1994) Information SuperHighway and the Digital Global Library: realities and challenges, *Microcomputers for Information Management* **Vol. 11 no. 3**, pp. 143-155

Corbin, B.G. and Coletti, D.J. (1995) Digitization of historical astronomical literature, *Vistas in Astronomy* **Vol. 39 no. 2**, pp. 127-135. Electronic version at URL http://www.eso.org/lisa-ii/lisaii-abstracts.html#abs11

Dempsey, L. and Weibel, S.L. (1996) The Warwick Metadata Workshop: a framework for the deployment of resource description, *D-Lib Magazine* **July/August**, URL http://www.dlib.org/dlib/july96/07weibel.html

Egret, D. and Albrecht, M.A. (eds.) (1995) *Information & on-line data in astronomy.* Kluwer Academic Publishers, Dordrecht. Electronic complement at URL http://cdsweb.u-strasbg.fr/data-online.html

Enyart, M.G. and Smith, R.A. (1996) Reference services: more than information chauffeuring, *Special Libraries* **Vol. 87 no. 3**, pp. 156-162

Garrett, J. and Waters, D. (co-chairs) (1996) *Preserving digital information.* Report of the Task Force on Archiving of Digital Information commissioned by The Commission on Preservation and Access and The Research Libraries Group. URL http://www.rlg.org/ArchTF/tfadi.index.htm

Ginsparg, P. (1996) Winners and losers in the global research village, *Joint ICSU Press-UNESCO Expert Conference on Electronic Publishing in Science*, Paris, France, February 19-23, 1996. URL http://xxx.lanl.gov/blurb/pg96unesco.html

Graham, P.S. (1995) The digital research library: tasks and commitments. In: *Digital Libraries '95*, Austin, Texas, 11-12 June 1995. Electronic version: URL http://csdl.tamu.edu/DL95/papers/graham/graham.html

Harter, S.P. and Kim, H.J. (1996) Accessing electronic journals and other e-publications: an empirical study, *College & Research Libraries* **Vol. 57 no. 5**, pp. 440-456

Hastings, K. and Tennant, R. (1996) How to build a digital librarian, *D-Lib Magazine* **November**, URL http://www.dlib.org/dlib/november96/ucb/11hastings.html

Hitchcock, S., Carr, L. and Hall, W. (1996) A survey of STM online journals 1990-95: the calm before the storm. In: *Directory of Electronic Journals, Newsletters and Academic Discussion Lists*, D.W. Mogge (ed.), 6th ed., Association of Research Libraries, Washington, D.C., pp. 7-32. Electronic version: URL http://journals.ecs.soton.ac.uk/survey/survey.html

Lehmann, K.-D. (1996) Bücher haben Vorteile, *Der Spiegel* **No. 51**, pp. 184-189

Mackenzie Owen, J.S. and Wiercx, A. (1996) *Knowledge models for networked library services – EUR 16905*. Office for Official Publications of the European Communities, Luxembourg. (Libraries in the information society series)

Michael, J. (1994) Responding to the revolution: becoming the library without walls. In: *From A to Z39.50: A networking primer.* J.J. Michael and M. Hinnebusch. Mecklermedia, London, pp. 32-38

Moore, N. (1996) Creators, communicators and consolidators: the new information professional, *Managing Information* **Vol. 3 no. 6**, pp. 24-25

Neavill, G.B. and Sheblé, M.A. (1995) Archiving electronic journals, *Serials Review* **Vol. 21 no. 4**, pp. 13-21

Negroponte, N. (1996) *Being digital*. Vintage Books, New York.

Nisonger, T.E. (1996) Collection management issues for electronic journals, *IFLA Journal* **Vol. 22 no. 3**, pp. 233-239

Ojala, M. (1993) What will they call us in the future? *Special Libraries* **Vol. 84 no. 4**, pp. 226-229

Okerson, A. (1996a) Who owns digital works? *Scientific American* **Vol. 275, July**, pp. 64-68

Okerson, A.S. (1996b) Introduction. In: *Directory of Electronic Journals, Newsletters and Academic Discussion Lists*, D.W. Mogge (ed.), 6th ed., Association of Research Libraries, Washington, D.C., URL http://poe.acc.virginia.edu/~pm9k/libsci/96/intro.html

Postman, N. (1992) *Technopoly: the surrender of culture to technology.* A.A. Knopf, New York.

Rusch-Feja, D. (1997) Subject-oriented collection of information resources from the Internet: a clearinghouse concept to support scientists in a German research institute, *Libri*, in press

Schaffner, A.C. (1994) The future of scientific journals: lessons from the past, *Information Technology and Libraries* **Vol. 13**, pp. 239-247

Schatz, B.R. (1997) Information retrieval in digital libraries: bringing search to the net, *Science* **Vol. 275**, 17 January, pp. 327-334

Schwartz, R.A. (1996) A sea change for academic publishing, *ASEE Prism* **Vol. 5 no. 9**, pp. 12-16

Soete, G. (1997) Let there be light! Licensing electronic resources: state of the evolving art. Summary of proceedings, San Francisco, USA, December 8-9, 1996. Version as of January 6, 1997. URL http://arl.cni.org/scomm/sum.html

Stewart, L. (1996) User acceptance of electronic journals: interviews with chemists at Cornell University, *College & Research Libraries* **Vol. 57 no. 4**, pp. 339-349

University of California Library Council (1996) The University of California Digital Library: a framework for planning and strategic initiatives. Version as of October 31, 1996, updated November 19, 1996. URL http://sunsite.berkeley.edu/UCDL/title.html

Varian, H.R. (1995) The information economy. How much will two bits be worth in the digital marketplace? *Scientific American* **Vol. 273**, September, pp. 161-162

Von Ungern-Sternberg, S. and Lindquist, M.G. (1995) The impact of electronic journals on library functions, *Journal of Information Science* **Vol. 21 no. 5**, pp. 396-401

Warnock, A. and Fullton, J.M. (1996) The Electronic Astrophysical Journal: resource location and archive management. In: *Astronomical Data Analysis Software and Systems V*, G.H. Jacoby and J. Barnes (eds.), ASP Conference Series **Vol. 101**, pp. 589-592

Weibel, S.L. (1995) The World Wide Web and emerging Internet resource discovery standards for scholarly literature, *Library Trends* **Vol. 43**, pp. 627-644

Weibel, S., Godby, J., Miller, E. and Daniel, R. (1996) OCLC/NCSA metadata workshop report. http://www.oclc.org:5046/conferences/metadata/dublin_core_report.html

Wusteman, J. (1996) Electronic journal formats, *Program* **Vol. 30 no. 4**, pp. 319-343