

Electronic Publications and Collection Management – Issues to Consider

Uta Grothkopf

*European Southern Observatory, Karl-Schwarzschild-Str. 2,
D-85748 Garching, Germany, e-mail: esolib@eso.org*

Abstract. Electronic publications affect virtually all areas of library collection management. Previously, clearly distinguished work areas could be defined, including Acquisitions, Processing, Using/Making Available, and Archiving of publications. These areas must be modified in the digital environment to remain effective. Based on the traditional model of library collection management, the article gives an overview of issues librarians should consider when dealing with electronic publications.

1. Introduction

Traditionally, library collection management has been grouped into clearly separated work areas which can be identified as Acquisitions, Processing, Using and Making Available, as well as Archiving of publications. In the digital environment, the difference between the various task groups is disappearing, and work areas become increasingly interdependent. In order to stay effective, traditional tasks require some modifications in the electronic era.

2. Acquisitions

2.1. Information about new electronic publications

One of the most obvious problems with regard to electronic publications management is discovering what is available. With new Web sites being announced each day, it can be difficult to stay informed about available resources on the Internet. Therefore, annotated lists of new electronic publications and information resources are highly appreciated services. Very useful compilations are:

- NewJour, a list of new journals and newsletters on the Internet. NewJour can be found at <http://gort.ucsd.edu/newjour/>
- the Internet Resources Newsletter (<http://www.hw.ac.uk/libWWW/irn/>), a WWW newsletter devoted to describing and linking to new Internet resources.

Many journal publishers offer alerting services; they distribute information about new e-journals or e-journal issues by e-mail to subscribers of the alerting service. Similarly, publishers and Internet bookshops often announce the availability of new publications by customer-specified authors or subject areas.

Librarians' mailing lists are excellent fora for discussing new electronic publications and sources; mailing lists cover all areas of library and information sciences. Some Web sites provide a good overview of existing lists, for instance the "Library-Oriented Lists & Electronic Serials" (<http://www.wrlc.org/liblists/>) and "Tile.net" (<http://tile.net/lists/library2.html>).

2.2. Selection criteria

How do we decide which electronic publications shall be made available to our users? With regard to content, the selection criteria applied to print publications shall also be applied to electronic media; for scholarly (e-)journals, publishers continue to provide quality control through the well-established peer-review system. In addition, electronic publications also must be judged with regard to technical features and usability:

- *Technical implications*
Electronic publications should be usable with widely available software. If special hardware or proprietary software is required, additional costs for purchase and maintenance will arise, and the interoperability with other systems can be limited.
- *Application of technical standards*
Standards for setting up and maintaining information resources are evolving and should be applied to electronic publications in order to allow seamless interconnection with other information systems and to be prepared for technological changes. Areas that require standards include network addresses (standardized location description), metadata (to describe the information source), and archiving (to guarantee future access).
- *Accuracy and currency of maintenance*
Is the Web site to which we consider providing access well-maintained and updated regularly? Are links from this site to other resources alive and working, or are they outdated? Can publications made available on this site be expected to be available indefinitely or must users fear that authors/publishers will remove documents after a short time? For libraries, long-term commitment regarding maintenance of the offered publications is a critical selection criterion.
- *Usage conditions*
When selecting electronic publications, we should review the usage conditions carefully. Unacceptable restrictions can prevent libraries from buying or subscribing to a given publication.

2.3. Obtaining electronic publications

In the print environment, obtaining publications could be as easy as purchasing books and journals and storing them on shelves. In contrast, obtaining electronic publications is much more complex. Access and usage conditions can vary largely for different publications; details typically are specified in separate contracts, so-called license agreements. License agreements complement or replace national

and international copyright (Okerson 1997). A number of issues need to be considered:

Access conditions: Librarians must be aware that “obtaining electronic publications” usually means leasing access rights for a given time rather than obtaining physical items. At the end of a leasing period, libraries may be left without anything, maybe even without the right to access those volumes published during the leasing period.

Future access to electronic publications will depend on the way they are archived and made accessible to users. All solutions currently mentioned in license agreements should be regarded as temporary and ephemeral as it is impossible to predict which solutions may be valid in only a few years time. It is obvious that archiving should be done by institutions with long-term commitment to this issue, and librarians should advocate a stable and reliable solution.

Usage conditions: The notion of one copyright applicable to all users is no longer true. While readers usually know for which purposes print publications can be used, they are often not aware of certain restrictions that must be adhered to in the electronic environment. To add to the confusion, different rights can apply to users from different institutes, countries, or user groups. With electronic publications, we face a “disunion” of information users. As a consequence, the responsibility of the individual librarian rises considerably. If contract terms are unfavorable or – even worse – unacceptable, we must negotiate with publishers instead of being able to rely on general copyright. In order to not abandon important rights we had with print publications, all clauses of a license agreement should be read carefully before a contract is signed, even though this is a time-consuming task. For instance, attention should be paid to the following questions¹:

- How are “use” and “user” defined in the contract?
- Will librarians be held responsible for how users use (or mis-use) the publication?
- Can the publication be used for essential library functions, e.g., interlibrary loan?
- Is the “fair use” concept adhered to?
- Will the publication be archived and if so, who is responsible for archiving?
- Is future access to volumes published during the subscription period covered by the contract even if a subscription has been cancelled in the meantime?
- Will additional costs arise for access to back volumes?

Costs: There is no longer just one price per publication, but a large variety of pricing models. It can be difficult to keep track of all available options and to find out which pricing model is suitable for whom. The ongoing changes in the publishing market lead to more flexibility for all parties involved, and libraries may have an option to choose among various pricing models based on different access conditions (number of simultaneous users, total number of logins etc.). Access through library consortia seems to be an advantageous solution as it gives

¹An excellent collection of sample contracts and other helpful resources is available from the Yale University Library at <http://www.library.yale.edu/~license/index.shtml>

participating libraries access to all journals subscribed to by any of the libraries who are members of the consortium. However, some libraries may not be able to join consortia for political or organizational reasons; there is a danger that the users of these libraries will be excluded from access to certain publications.

It is hoped that the ongoing market changes will finally lead to price decreases. Unfortunately, just the opposite has occurred up to now as publishers generally raise the prices of the new (i.e., electronic) product compared to the old (print) medium. In addition, the costs involved in electronic publications are not limited to the obvious access fee, be it through subscription rates, on a pay-per-view basis, or through consortia. Printing costs typically are shifted to the individual institutes instead of being borne by the publishers. Charges involved in accessing ceased or cancelled publications cannot be determined as of today since solutions for archiving electronic publications including costs for future access still are very much in flux. Special tools (hard- and software) may be required to use the publication. A reliable and fast Internet connection has to be in place, and the general telecommunications infrastructure must be able to keep up with the rising demand.

Some, if not all, of these costs will have to be borne by libraries. Unfortunately, it remains an unanswered question how we can increase library budgets accordingly.

3. Processing electronic publications

3.1. Cataloging electronic publications

In order to use electronic publications, they must be retrievable through well-known information sources like library catalogs. It would be most efficient to use existing library procedures to create catalog records. As is the case with print publications, catalog entries for electronic documents must correspond to existing cataloging formats, for instance MACHine Readable Cataloging (MARC) and rules for entering data into the format, e.g., Anglo-American Cataloging Rules (AACR2). Many standard catalog formats and cataloging rules have been enhanced in the recent past in order to make them suitable for bibliographic records for items accessible through the Internet.

Electronic publications should contain the standard bibliographic details which typically are found on the title page of print publications (author, title, publication date, etc.) as well as descriptions of the content (keywords, thesaurus terms, etc.). In addition to formal and contents-related data, catalog entries for electronic documents must contain information regarding the network address (or network name) of the e-publication, the mode of access (via WWW, gopher, telnet, etc.), required tools (necessary hardware and software), computer file characteristics (size and format of files), information on access rights and restrictions as well as copyright information.

This "information about information" is called metadata. One of the most popular examples of metadata systems is the Dublin Core set of metadata. Currently, preparations are being made in order to map MARC and the Dublin Core so that metadata can be easily retrieved through standard library records (see Rusch-Feja 1998).

3.2. Locating electronic publications

Locating networked documents can be difficult, because network addresses can change quickly. Information about the location of an electronic publication therefore must be available even if the publication has been moved from the original computer to a new location. Currently, most Internet resources are located using Uniform Resource Locators (URL) which unfortunately are very unreliable as even one misspelt character in the address will prevent users from access. A better solution are Uniform Resource Names (URN) which allow permanent, location-independent names to be assigned to electronic documents. Such a system is based on a central directory and a database to locate the objects contained in the directory. Name resolvers allow users to retrieve the current network address of the document from the central directory of addresses. Address changes are only noted in the directory, not in the name itself.

4. Using and making available electronic publications

4.1. Announcing and presenting electronic publications

Typically, print publications still receive more attention in our libraries than their electronic counterparts. How can we announce, present and promote electronic publications? For those libraries equipped with a Web-based catalog, an obvious solution is to use the MARC 856 tag (see Olson 1995) to include a publication's network address. Users will then be able to directly "click through" to an e-publication after having retrieved the catalog entry. The library may also decide to create a Web page that lists electronic books and journals, preferably by title, author, or editor rather than by publisher as most users hardly ever remember which publisher issues which publication. The availability of newly acquired resources can be announced in the "News" section of the library's Web pages or distributed to users by e-mail.

We are used to routing particular journal issues among interested users. In a similar way, selected information about new electronic journal issues can be distributed, for instance by sending journals' contents tables via e-mail, provided that electronic distribution of contents tables is permitted by the publisher.

In this context, we should mention that the most interesting publications are useless if they are not accessed by users. Today, users expect that information resources are interconnected, and that the links are stable and fast. Should these expectations not be met by a publication, scientists may not be willing to use it. In order to monitor the actual usage, librarians should have access to statistics collected by publishers. Small numbers may indicate that users do not need the resource, or that they are not yet sufficiently aware of its existence; large numbers can help librarians to justify expenditure on a given resource as well as to argue for more funding.

4.2. Identifying e-publications

In order to be used effectively, electronic documents need to be uniquely identifiable. Recently, various international scientific publishers and cooperations agreed on implementing the so-called Digital Object Identifier (<http://www.doi.org>). The DOI intends to link users and owners of information in order to "facilitate

automated digital commerce”, but it is obvious that it represents the interests of the information owners rather than those of information users. (Information users do not have a strong, unified lobby yet.) The DOI is based on a name resolver that points to the current location of requested documents. In addition, it provides links to associated information, for instance access rights and prices. DOI are a first step towards fully functional Electronic Copyright Management Systems (ECMS) which will allow copyright holders (the authors or publishers) to automatically tag copyrighted works and closely monitor reuse of material transmitted interactively on digital networks (Gervais 1997).

In contrast, information users are more concerned about the availability of means to prove the authenticity (identity) and integrity (correctness) of documents. Digital information can be changed and corrupted easily, by accident or deliberately. Mechanisms are necessary to prove the integrity of documents, for instance whether they are still intact after they were moved to a new storage medium. Users must be assured that what they see is what they requested; techniques that facilitate tracing who created, used or modified a document, like digital watermarking and digital fingerprinting, are becoming more and more important.

4.3. Searching, viewing and printing electronic publications

The electronic medium offers considerable advantages over print publications with regard to search and retrieval of documents, immediate access to full texts, and hyperlinks to other relevant information resources. But “electronic only” usually is not sufficient, and users expect electronic publications to be available in a variety of formats for different purposes. HTML or any subsequent hypertext language is required for online browsing, searching and navigating. It has been mentioned already that electronic publications should be usable with standard hardware and software, avoiding the necessity of installing special programs on each user’s desktop. After having retrieved relevant documents, many users prefer to print them, and a suitable format must be available for this purpose. PDF (Portable Data Format) and PostScript are two currently popular print formats.

4.4. Citing electronic publications

Information found in electronic resources must be referenced as accurately as information from printed documents, and librarians often are asked how electronic publications are to be cited. A widely agreed-upon citation system for electronic publications is not yet available; the following sites provide useful suggestions:

- The IFLA Collection of Citation Guides serves as a clearinghouse by providing access to further sites of interest (<http://www.nlc-bnc.ca/ifla/I/training/citation/citing.htm>)
- ISO Standard 690-2 on “Information and documentation – Bibliographic references” deals specifically with information and documentation (<http://www.nlc-bnc.ca/iso/tc46sc9/standard/690-2e.htm>)
- “Bibliographic formats for citing electronic information” is based on the book “Electronic styles” by Xia Li and Nancy B. Crane (<http://www.uvm.edu/~ncrane/estyles/>). The text contains helpful rules

for citing different types of electronic information resources and provides a large number of examples.

5. Archiving electronic publications

The topic of archiving electronic publications still bears more questions than answers. Up to now, it has not been decided who is or will be responsible for archiving, what exactly shall be archived, and where the archived material shall be stored. It can be assumed that solutions will vary largely in different countries.

Electronic publications may contain text, graphics, tables, sound, movies, or other types of data. They are not limited to the information contained, but are to be understood as interlinked entities which require (and are required by) other distributed sources. As electronic media can be changed more easily than documents on paper, they are more vulnerable than print publications. Archival data for electronic publications therefore must include the following information (Garrett & Waters 1996):

- content of electronic publications (i.e., the information contained)
- fixity (content stored in a medium as opposed to continuously updated “dynamic” documents)
- reference (reliable systems for locating and citing)
- provenance (a record of the document’s origin and chain of custody)
- context (links within a document and/or between various sources)

Many institutions are trying to find appropriate answers to the question of how electronic documents can be prepared today for the (still unknown) technology of tomorrow. In this context, we need to address the uncertain durability and future availability of current storage media like CD-ROM and magnetic tape and the respective reading devices. Deterioration of the storage media is a distinct problem, but an even greater danger lies in the foreseeable technological obsolescence of the equipment needed to use their content. As evolving technologies do not necessarily replace previous ones, libraries theoretically have to provide all kinds of reading devices in parallel – an unfulfillable task. Migrating publications to newer storage media will be the only suitable solution. Archiving and migrating data stored in proprietary systems can turn out to be extremely difficult, as it may be necessary to preserve not only the intellectual content, but the entire search and retrieval system in order to maintain usability (European Communities 1997). In contrast, migration will be easier if documents are stored in system-independent data formats.

It is crucial to all libraries as well as information users that access to electronic publications is guaranteed over time. As archiving electronic publications is a very expensive undertaking, small libraries probably will not have sufficient funds to continue archiving publications as they have done in the print environment. National libraries are possible candidates for becoming archiving institutions, probably in collaboration with the publishers. For instance, national libraries could act as mirror sites as long as a given publisher still maintains the electronic publication. In case the publisher disappears, the national library could solely take over responsibility in order to guarantee future access.

6. Conclusion

Electronic publications affect all areas of library management, and traditional collection management needs to be enhanced to remain effective. There is no single rule or method for dealing with electronic publications that can be applied to all libraries; each librarian has to find the solution that suits local users best. The electronic environment has initiated a much more market-oriented approach in evaluating information resources than we have been used to up to now; librarians usually can choose among a variety of access conditions when dealing with e-publications. This new flexibility leads to a remarkably increased responsibility for the individual librarian.

While it is important to find individual solutions, we must also overcome our isolation, and seek dialog with our colleagues in order to learn about other approaches to coping with a rapidly changing information environment. Not only can we profit from the solutions our colleagues have already applied, but we will also be able to join forces and hold a stronger position during negotiations with vendors and publishers. Joint efforts will have more impact on implementing user-friendly access conditions.

Many issues in electronic publishing are still open to discussion. Often publishers have not yet abandoned the phase of experimenting with possible solutions, and most of them are willing to cooperate with and listen to librarians. Dealing with electronic publications needs thoughtful and fair solutions; only these will be advantageous for the entire community of information providers, mediators, and users.

References

- European Communities, 1997, Guidelines on best practices for using electronic information. Updated and enl. ed., Luxembourg: European Communities, 1997. [Online] Available: <http://www2.echo.lu/dlm/en/gdlines.html>; updated 31 March 1998 [July 18, 1998]
- Garrett, J. & Waters, D., 1996, Preserving Digital Information: Final Report and Recommendations, [Online]. Available: <http://www.rlg.org/ArchTF/> [July 1, 1998]
- Gervais, Daniel G., 1997, Electronic Copyright Management Systems (ECMS) in a network environment, [Online]. Available: <http://www.copyright.com/> (choose News / CCC in the News / Articles) [July 1, 1998]
- Okerson, A., 1997, Copyright or Contract? *Library Journal*, 122(14), 136-139
- Olson, Nancy B., 1995, Cataloging Internet resources: a manual and practical guide, Dublin, OH: OCLC Online Computer Library Center, [Online], ftp://ftp.rsch.oclc.org/pub/internet_cataloging_project/Manual.txt [July 1, 1998]
- Rusch-Feja, D., 1998, Metadata: standards for retrieving WWW documents (and other digitized and non-digitized resources), these proceedings, 157