National Library of Canada | Bibliothèque nationale du Canada

Electronic Publications Pilot Project (EPPP)

Final Report

Electronic Publications Pilot Project Team
and
Electronic Collections Committee

June 1996

TABLE OF CONTENTS

Editors' Note

On 22 November 1995, the National Library of Canada Executive Committee approved in principle the Electronic Publications Pilot Project (EPPP) Report. By April 1996, there were some 64 electronic titles in the Library's EPPP collection. While primarily of interest to the National Library, it is hoped that this report will be of use to the Canadian library community in general.

Based on the report, an overview plan to mainstream the handling and preservation of these Canadian electronic publications is being implemented.

Executive Summary

From June 1994 to July 1995, the National Library of Canada (NLC) conducted a pilot project to acquire, catalogue, preserve and provide access to a small number of Canadian electronic journals and other representative publications available on the Internet. This project was called the Electronic Publications Pilot Project (EPPP).

The EPPP operated on existing NLC hardware using software, already licensed by the NLC, and freeware.

By July 1995, the EPPP collection consisted of 25 titles with another 21 titles in progress. These titles are mounted on the NLC InfoServer and are accessible internally and externally on the NLC gopher and an EPPP home page that is part of the NLC Web site (see Universal Resource Locators: http://www.nlc-bnc.ca/eppp/e3pe.htm). Catalogue records for the EPPP titles are available in the Access AMICUS database.

The EPPP researched a number of issues including acquisition and long-term preservation. Findings concerning these issues are included in this report.

Electronic publications should be acquired on a selective basis. Regarding access to these acquisitions, the NLC intends to adopt a policy in which all items acquired for the NLC collection would be accessible on-site and, in general, to off-site users as well. The NLC will continue to explore copyright issues.

Mounting individual electronic titles on the NLC server was very time-consuming. Staff found that it was more efficient to “mirror” WWW publications automatically by periodically copying the whole publication site rather than selected files. The report recommends further development of automated procedures for acquiring e-publications and for validating the content of the NLC copy against subsequent versions available at other network sites.

Many problems related to file handling for serials acquisition and control were identified during the project. The serials control process used for the EPPP must be streamlined and automated before the NLC can expect to handle significant volumes of material.

The Universal Resource Locators (URL) for each title were included in the catalogue record of titles acquired during the EPPP. This link should be improved with the eventual goal of having a hypertext link between the URL in the bibliographic record and the item in the NLC electronic collection. Initiatives currently underway to incorporate descriptive data about an electronic document

(i.e., “metadata”) as part of the overall document structure may provide future cataloguing efficiencies, such as “electronic Cataloguing-in-Publication”. The NLC will work with publishers and standards groups to define Canadian library requirements for metadata in e-publications.

The EPPP found that the lack of appropriate standards to be a major obstacle to the long-term accessibility and preservation of electronic publications. The NLC will monitor the development and promote the adoption of e-publication standards. When an e-publication is available in several formats, the NLC will normally select the “standard” (or format neutral) versions of that publication for acquisition and long-term preservation. In cases where the only available version is in a proprietary format, the NLC would convert the e-publication to a standard format, if possible.

The EPPP found that none of the current electronic media is acceptable as a long-term archival medium. For short-term preservation, the NLC will store its e-publication collection online on magnetic disk. For mid- and longer-term preservation, hierarchical storage management (HSM) systems will be further studied to determine whether or not they can satisfy the NLC’s storage and preservation requirements for electronic publications. A tracking system should be developed to monitor the NLC’s electronic collection so that items threatened by hardware or software obsolescence can be converted to new preservation formats.

The work of acquiring, cataloguing and managing a collection of electronic publications should continue with the existing pilot level EPPP technical system. It is expected that more NLC personnel will work with electronic publications than is now the case.

The report includes requirements for an NLC Manage E-Publications application that should interface with AMICUS. Implementation of the production level Manage E-publications application would take one to two years depending on the priority and resources allocated to development.

The EPPP report also identifies the need for further work at the policy level to:

- define what constitutes published electronic Canadiana (in consultation with other agencies);
- develop selection and preservation criteria for the NLC electronic collection;
- incorporate electronic publications into the selection, collection management and preservation policies of the NLC.

The report also stresses the need for the NLC to initiate discussions with the Canadian library and publishing communities in order to develop a cooperative

national strategy for the long-term storage and preservation of electronic publications.

I. Introduction and Background

Established by Parliament in 1953, the National Library of Canada (NLC) is a federal institution located in Ottawa, whose main role is to acquire, preserve and promote the published heritage of Canada for all Canadians, both now and in the future. The Library serves as one of the nation's prime resources for research in Canadian Studies. It also plays a major role in fostering the development of library resources and services throughout the country, and in facilitating resource sharing among Canadian libraries.

The explosion in electronic publishing and networked information is creating major challenges for the NLC and Canadian libraries. The present shift toward an electronic information environment raises issues that will impact the way the NLC fulfills its heritage mandate. Identifying, locating, reporting, storing and preserving electronic publications raise issues to be addressed and solved in the near future. If these issues are not addressed, there will be major gaps in the Canadian heritage and research tools of the nation. Thus, it is emphasized that the NLC's preservation and promotion of the published heritage of Canada for all Canadians includes electronic publications (e-publications), and it means preservation in perpetuity of such publications.

Given the NLC's mandate to preserve the Canadian published heritage, acquisition of an e-publication by the NLC means that the NLC must obtain an electronic copy of the e-publication. For the NLC, simply acquiring the right to point or link to an e-publication stored at the e-publisher's site is insufficient to ensure preservation. Experience to date indicates that the preservation of e-publications cannot be left to e-publishers, just as the preservation of publications in other media cannot be left to publishers. The NLC must also acquire an electronic copy to check that the e-publication is in a form that is readable by standard software for current and future generations of readers and researchers.

Following the recommendations of an earlier study[1] of the NLC and e-publications, the NLC conducted a pilot project during the period June 1994 to July 1995 to acquire, store, preserve, catalogue and provide access and service for a small number of existing Canadian on-line electronic journals and other

[1] Thacker, Jane and Bruce Crisp. *Electronic publications and the National Library of Canada: discussion paper and recommendations.* EC93:20. April 1993. 46 pp. (Internal document distributed externally on request.)

representative publications available on the Internet. This project was called the Electronic Publications Pilot Project (EPPP).

For the purposes of this Project, an electronic publication was defined as follows:

> An electronic publication is a document resulting from the act of publishing in which the information is encoded, accessed, and made intelligible through the use of a computer. Electronic publications fall into two broad categories: those that are distributed in multiple copies on physical media, such as compact disks, diskettes, and magnetic tapes; and those that reside on a host computer and are accessible over a communications network. Electronic publications undergo the same formal preparation activities associated with traditional print publications.

The NLC is already addressing the acquisition and processing of e-publications on physical media. The EPPP and this report deal with another category of e-publications, that is, networked (or on-line) e-publications. Such publications are obtained over the Internet via protocols such as file transfer, electronic mail, or hypertext transfer. More specifically, the EPPP dealt with existing Canadian on-line serial e-publications on the Internet and e-publications written and published on the Internet by the NLC. The EPPP did not deal with existing (non-electronic) publications that are imaged and then made accessible on the Internet.

A huge range and number of documents and document types exist on the "Canadian" portion of the Internet. It would be impossible for any particular library or for Canadian libraries in general to acquire and preserve all these documents.[2] Yet, the NLC is charged with acquiring and preserving the Canadian published heritage, in perpetuity. Therefore, one of the first EPPP activities was to define the types of e-publications that the NLC should acquire and preserve. This was accomplished by stating what, for NLC purposes during the pilot, an e-publication was not. These exclusions are stated in terms of Internet delivery protocol and style of content. At least for this initial work with e-publications, the NLC excluded the following:

Electronic mail (e-mail)
Except for e-publications delivered to patrons and subscribers via e-mail, e-mail is not considered to be an e-publication. Thus, e-mail communications between individuals and groups, including discussion groups, listservs, and bulletin boards, are not considered to be e-publications.

[2] In the Internet world there are many institutions that are in the process of acquiring and storing certain types or groups of electronic documents. This is commonly called "archiving". In many cases, the objectives and scope of the archiving activity are ad-hoc, including the intended duration of electronic document preservation in the electronic archive and access to documents in the archive.

World Wide Web documents
An e-publication is not just any World Wide Web (also known as WWW or Web) home page or group of Web pages. For the NLC to treat Web documents as e-publications during the EPPP, the Web presentation, despite being in multimedia and hypertext, must have the characteristics of traditional publications such as undergoing formal preparation activities traditionally associated with the publishing process; and the Web presentation must not consist mainly of pointers to other Web locations. Although this is an "I'll-know-it-when-I-see-it" statement, this exclusion definition must suffice for the NLC's initial work with e-publications.

Gopher sites
As with Web documents, an e-publication is not just any Gopher site, or portion of a Gopher site. For the NLC to treat a Gopher document as an e-publication, it must have the characteristics of traditional publications such as undergoing formal preparation activities traditionally associated with the publishing process.

On-line databases
An on-line database is not considered to be an e-publication. However, an e-publication constructed or derived from a database, and so variable with each reading, may be an e-publication.

File Transfer Protocol (FTP)
Except for e-publications delivered to patrons and subscribers via FTP, FTP archives are not considered to be e-publications. Therefore, data files in FTP archives are not considered to be e-publications, although other objects and files in the archives may be e-publications.

Bulletin board system (BBS)
Bulletin board systems, which can be employed to support chat groups, are not considered to be e-publications.

For the purposes of this project a Canadian e-publication was defined as an e-publication "published" in Canada (i.e., the computer providing the e-publication is in Canada), and/or, an e-publication sponsored or produced by a Canadian company or individual.

The above specification of an e-publication for the EPPP is not sufficient for the long-term. A clear specification would help to define the mandates of various players dealing with electronic information. The NLC, in consultation with other agencies, is now working on the definitions of what constitutes a "published" electronic document and published electronic Canadiana.

Concerning the technical platform, the EPPP e-publication server had to operate on existing NLC hardware, and, due to a limited budget (see Appendix B), used either software already licensed by the NLC, or freeware. As a corollary, the EPPP excluded any e-publications for which the NLC could not obtain adequate client freeware for viewing the documents.

For this pilot, the EPPP excluded any e-publications for which the NLC could not obtain permission from the e-publisher to mount the title on the NLC server.

III. Methodology

Initiation

The EPPP project plan was developed in the summer of 1994. Staff from several areas of the NLC were assigned to work part-time on the project team. The project began in September 1994. Acceptance of the EPPP final report by the NLC Executive Committee formally completed the project, although operation of the EPPP server and activities will continue until it is replaced by a Manage E-Publications application.

In ramping-up the EPPP, it was quickly apparent that not all EPPP team members and resource people were at the same level of expertise with respect to the Internet and the use of Internet tools. The first major activity was, therefore, to organize and provide an Internet familiarization session for those people who required it. These sessions were well received and the experience gained was useful in expanding Internet training to other NLC personnel.

Acquisition of EPPP titles

According to the e-publication selection criteria for the EPPP, acquisitions personnel worked to identify titles for inclusion in the pilot and to obtain permission to mount the titles on the NLC server.

Various functional and organizational areas of the NLC, such as acquisitions, cataloguing, reference, and preservation, studied the handling of e-publications and prepared procedures for working with e-publications.

Owing to the low non-salary budget of the EPPP, freeware or existing NLC systems were employed wherever possible for systems support in all areas of the EPPP, including acquisition, document creation, cataloguing, and access to and presentation of e-publications.

EPPP systems

The EPPP server was first mounted on the NLC Communications Gateway, which is a Digital Equipment Corporation VAX 4100 operating under the Open VMS operating system; and then in June 1995 on the NLC InfoServer, which is a Digital Equipment Corporation Alpha AXP 2100 operating under the UNIX operating system.

The standard workstation at NLC is a 486SX under MS Windows 3.1 with 16Mb RAM and a 120 Mb hard disk. Fourteen workstations were upgraded to make them capable of more fully supporting access to and manipulation of multimedia e-publications. The upgrades typically consisted of CPU to DX2 66 MHZ, hard disk to 340 Mb, and the addition of sound cards and headphones.

In order to facilitate NLC publishing, both for production of the NLC Web site and for paper publications, part of the EPPP budget was used to purchase a scanner.

Various means of implementing facilities for naming, organizing, and storing e-publications on the NLC server were investigated. With the Communications Gateway VAX under Open VMS, the Microsoft Windows File Manager on the workstation and Network File Services were used to manage the files on the VAX. On the InfoServer under UNIX, an XWindows package on the workstation and a public domain file manager on the InfoServer were used to manage files. Technical services personnel, who acquired titles and built the file storage structure, were trained in the file-management tools.

A file-directory structure was created to store and manipulate the e-publications. On the InfoServer, public-domain, SQL-database software was employed to store the EPPP titles and descriptions, and to point to the directory structure where the physical files containing the e-publications are stored. E-publication Universal Resource Locators (URLs) are constructed on the fly from the database information.

Additional information on the EPPP technical system is at Appendix F.

Titles were made available for access on the NLC Gopher employing the existing structure in the NLC gopherspace.

For Web access, an NLC EPPP home page was created, and access via title and broad Dewey subject category was opened to the public in March 1995. The EPPP became part of the overall NLC Web structure in June 1995.

EPPP team members researched a number of issues during the project including copyright and long-term preservation. Conclusions concerning these issues are included in this report.

Throughout the project, team members recorded ideas and recommendations on operational procedures and on requirements for the NLC Manage E-Publications application. These requirements form a section of this report.

IV. Results

IV.1. General Issues

The issues discussed in this section cross organizational and functional lines.

IV.1.1. Copyright

In Canada, copyright is defined as[4]:

> "The sole right to produce or reproduce a work or any substantial part in any material form or to authorize such reproduction".

Fair dealing is a defence for users in cases of copyright infringement, in particular an infringement of the sole right to reproduce. The purposes for which the work was used in the infringement must be:

- private study
- research
- criticism
- review or
- newspaper summary.

Also, in cases of criticism, review or newspaper summary, the source and the author's name must be given.

In discussing NLC copyright policy options with respect to e-publishing, it was assumed that the NLC will acquire electronic copy of Canadian e-publications in order to preserve the published Canadian heritage. The NLC will also mount these publications on its server.

The NLC will keep track of copyright permissions and restrictions on titles in its collection of e-publications by recording the copyright status of each e-publication in the NLC collection in the AMICUS database, and, if it is in the free subset, the reason for its being there.

The Canadian Information Highway Advisory Council, Copyright Subcommittee recommended that browsing be considered an act of reproduction[5]. This recommendation was reconsidered at the full Council level to state:[6]

[4] Canada. *Copyright Act*. Section 3(1).

> "It should be left to the copyright owner to determine whether and when browsing should be permitted on the Information Highway ..."

Thus, interpretation of the Copyright Act, with respect to e-publications, and interpretation of fair dealing in this context require study and discussion. The NLC is in the process of obtaining the opinion of legal counsel on this issue.

EPPP economic and copyright models for e-publishing

The Canadian Internet electronic journals included in the EPPP reflect a variety of copyright practices. For purposes of this pilot, all copyright holders have given the NLC permission to copy, store and disseminate their publications. The economic model for these e-journals is predominantly free Internet distribution. However, in June 1995, the EPPP dealt with two examples of commercial e-journals.

Three titles in the EPPP have no copyright statement, although it should be understood that copyright exists, with or without statement. Two titles state copyright, date and owner. Another adds "all rights reserved". The following are typical of more extended copyright statements:

> *Big Dreams - copyright (c) Dare to Dream Enterprises. Permission is hereby given to print out and distribute unlimited exact copies as long as no fee is charged.*
>
> *Government Information in Canada/Information Gouvernementale au Canada and its contents may be freely copied and distributed on condition that the journal and its contents are not altered, journal and authors are credited, and there is no charge for the journal or its contents. Rights for reprinting and republication for profit are reserved to authors.*
>
> *International Teletimes is a publication of the Global Village Communication Society and is copyrighted ©1993 by the same. All articles are copyrighted by their respective authors; however, International Teletimes retains the right to reprint all material unless otherwise requested by the author. This magazine is free to be copied and distributed UNCHANGED so long as it is not sold for profit.*

---

[5] Information Highway Advisory Council. *Copyright and the Information Highway. Final Report of the Copyright Subcommittee*. Ottawa, Ontario. Information Highway Advsory Council Secretariat. March, 1995.

[6] Information Highway Advisory Council. *Connection Community Content The Challenge of the Information Highway Final Report*. Ottawa, Ontario. Information Highway Advsory Council Secretariat. September, 1995. p. 115.

Although these statements are taken from e-publications, the terminology reflects the print world. These statements are not explicit about electronic access and distribution.

A non-traditional version of the same intent is employed by *Cropduster*:

The EPPP is discussing this with interested parties to learn more about this model and to identify ways of meeting these concerns. Wimsey also initiated a discussion on the can.infohighway newsgroup.

Copyright policy options for NLC

A number of possible scenarios on how the NLC could handle access to its e-collection in terms of the copyright issue were considered by the EPPP, together with the pros and cons of each scenario.

The NLC position is that access to publications through libraries should be equitable and affordable, and include access to e-publications. The NLC will consult stakeholders and legal counsel in developing an access policy to maximize equitable and affordable access to items acquired for the NLC e-collection for Canadians.

As the processing of e-journals moves into the mainstream of NLC's operations, the tasks of efficient handling of copyright clearance and recording of access rights must be addressed. These are recorded as system-support requirements for the production level Manage E-publications system.

IV.1.2. Electronic Publication Standards

The lack of widespread adoption of standards for on-line e-publishing makes their collection and preservation, and provision of access very difficult. The intention of instituting appropriate standards would be to ensure that the widest possible audience had the necessary tools to access an e-publication.

The incorporation of standard identifier information in e-publications would facilitate the finding and retrieval of e-publications by the intended audience using standardized finding tools, and would serve to promote the e-publications that followed these standards.

It is much more probable that future generations of Canadians will be able to access e-publications that adhere to standards than to access e-publications that do not. This is because the software necessary to interpret the standard e-publications is more likely to be migrated to future generations of hardware. Therefore, e-publishers and authors concerned about preservation of their electronic works would be wise to understand and adopt appropriate e-publication standards.

Although the NLC is a publisher, it is certainly not a large publisher. The NLC has neither the resources nor the extensive expertise required to be a leader in

the development of e-publication standards. However, the NLC could facilitate and promulgate the adoption of e-publication standards by e-publishers.[7]

Standards that are likely to be significant for e-publishing include:

- electronic signatures for ensuring e-publication integrity;
- encryption and time stamps for commercial transactions;
- Standard Generalized Mark-up Language (SGML)[8] and Document Type Definitions (DTD) for identifying, creating, storing, and presenting e-publications;
- Hyper Text Mark-up Language (HTML)[9] for creating, storing, and presenting e-publications;
- Text Encoding Initiative (TEI)[10] which is a DTD originally developed for e-publications in the humanities.

IV.1.3. Electronic Publication Integrity

E-publications can be more easily changed and manipulated than works in other media, both by authorized and unauthorized personnel. Many e-publications change almost continuously, in an authorized fashion, over a given period of time.

There are cases where an institution, such as the NLC, may legitimately want to change an e-publication, for example, from storage in a proprietary standard such as PostScript (TM) or WordPerfect (TM) to an international standard, such as SGML, in part, to avoid issues of preservation of proprietary software. The issues are whether this changes the original e-publication in a way that would be unacceptable from a preservation perspective, and whether this would lead to infringement of an author's right of integrity.

The right of integrity, or an author's "moral right" with respect to a title, is defined in the *Copyright Act*[11] of Canada. The right of integrity is infringed only if the work is:

> "to the prejudice of the honour or reputation of the author,
> a) distorted, mutilated or otherwise modified; or
> b) used in association with a product, a service, cause or institution."

---

[7] NLC could use its international roles within the International Standards Organization Technical Committee 46 and the International Federation of Library Associations and Institutions to propose and participate in the international development of library requirements for the form and presentation of e-publications.

[8] For example, see URL: http://www.sil.org/sgml/sgml.html

[9] For example, see URL: http://www.w3.org/hypertext/WWW/MarkUp/MarkUp.html

[10] For example, see URL: http://etext.virginia.edu/TEI.html

[11] Canada. *Copyright Act*. See S. 14.1(1) for definition of the right of integrity, and S. 28.2 for definition of infringement of the right of integrity.

Thus, it appears that conversion of an e-publication to an international standard, all the while attempting to preserve the quality of the original, would not be an infringement of the right of integrity. In addition, since the conversion would be performed to ensure the feasibility of long-term preservation, such a conversion would normally be acceptable from a preservation viewpoint. It is noted, however, that the NLC will attempt to avoid conversions by allocating priority to acquiring, preserving and facilitating long-term access to "format neutral" (i.e., standard) versions of Canadian e-publications, if more than one electronic version of the publication exists.

The answer to ensuring respect of an author's moral rights to an e-publication will be through the widespread use of standards, such as digital signatures and encryption. In the meantime, institutions such as the NLC cannot be sure that the acquired version of e-publication is unchanged from the "original" copyrighted version.

The cases where e-publications are legitimately and continually changed by the copyright holder raise issues for acquisition and preservation. These issues are discussed in the Acquisitions, and Collection Management and Preservation chapters of this report.

IV.1.4. Versions of E-publications

It is widely believed that there will be more versions of publications in the electronic environment than in print. Versions are documents with identical, or nearly identical content, but different physical forms. For example, there might be a Postscript version of an e-publication, where the Postscript format is intrinsic to the document, plus a second version in HTML.

The issues are:

1) which version the NLC should retain as its "preservation version";
2) whether or not the NLC should keep multiple versions (if so, this would parallel the print world); and
3) whether or not the NLC should simply aim to preserve the version least likely to create long-term technical problems.

These issues are partially addressed above with the recommendation concerning e-publication standards. The NLC will, when more than one version of a publication exists, attempt to acquire those versions of publications which are "format neutral", or conform to standards. For example, an HTML version rather than a PostScript (TM) or a WordPerfect (TM) version would be retained for purposes of preservation and long-term access. To do otherwise might entail archiving proprietary software (and, perhaps, hardware) so that access could be

maintained. Retaining standard versions of publications would presumably pose far less technical problems for the NLC. In fact, if only a non-standard version of a Canadian e-publication is available to the NLC, the NLC should convert the e-publication to a standard format.

The versions question was also noted as an issue for developing acquisitions procedures, and is further discussed in that section. With the recommendation that all standard versions be acquired, staff gained valuable experience in handling different title formats in the NLC environment.

The question of whether multiple versions which contain the same information can be described in a single cataloguing record has been the subject of much debate on discussion lists such as VPIEJ-L. For EPPP titles, the descriptive cataloguer created a separate record for each version of the title, which was acquired. This practice may be changed to creating only one bibliographic record for all versions. If some of those acquired versions are rejected or converted to other formats for long-term preservation, the bibliographic records will have to be revised or deleted, unless the decision on preservation is part of the acquisition process and, therefore, made before cataloguing. Revisions to the bibliographic record would also be necessary if only one record were created.

IV.1.5. "Boundary" of an Electronic Publication

The boundary of an e-publication is an issue for e-publications and copyright because, with hypertext e-publications, it can be difficult to define where a work starts and ends and where a copyright starts and ends. This matter of boundaries also raises issues concerning acquisition, preservation and copyright management systems.

Hypertext e-publications deserve special focus because they are the main reason for eliminating known "archival" media such as paper and microfilm in the preservation of e-publications. Hypertext e-publications may exist on two or more (and sometimes on many) remote servers and Internet domains, with hypertext links being used to "assemble" the parts into a coherent whole. This does not seem to be an issue in the short term, when many hypertext links could reasonably be expected to function for a limited lifespan online. It becomes an important issue; however, in the long-term, if a publication consists partially (or wholly) of a series of hypertext links, many or none of which may be operational when the publication is accessed in the future.

It has been suggested that hypertext links are analogous to footnotes or bibliographic entries in a print medium; that is, a reference may go out of print and be unavailable for consultation in the future. However, this is an oversimplification. Hypertext links, rather than serving only a reference or

footnoting function, can also be intrinsic to the publication. Sometimes, they comprise too large a part of the publication to be dismissed simply as a blind lead in a bibliography.

What, then are the options? One could, copyright permitting, bring in the files associated with every hypertext link in a publication and store them with the files for the publication themselves. Or, one could choose to pull in only a portion of the links, such as those to Canadian sites or to Canadian publications elsewhere. One could also choose to regard a hypertext publication as analogous to a TV program guide, which points to program items that may or may not be archived somewhere, but may not be viewed beyond the limited lifespan of an issue of the guide. Despite the leads being blind, the "pointing tool" retains a certain social and historical value.

From a practical standpoint, it is often impossible to bring in all the linked objects of a hypertext e-publication. For example, how far does one go in retrieving the next set of links in the linked documents? Consider a case in which the loop of links is infinite. A compromise is essential.

It also remains to be determined whether an e-publication which consists of nothing but electronic links to other items on remote servers is itself a publication with intrinsic value. Directory type Web sites, consisting mainly of links, are unquestionably valuable. So far, it has been NLC netsurfer experience that the non-directory type of Web site, consisting mainly of links, is ephemeral and more analogous to a broadcast than to a publication. This situation should be monitored, and more objective criteria obtained for determining whether and when an electronic "publication", especially one composed mainly, or only, of hypertext links, should be treated as a publication for preservation.

After consideration of the difficulties involved in tracing hypertext links and acquiring the linked objects, the EPPP team took the position that a hypertext e-publication consists of the linked objects stored at one Internet domain. Links (i.e. Universal Resource Locators (URLs)) to objects on other domains are retained only as links. It is realized that the integrity of the off-domain URLs cannot be maintained or preserved. Therefore, at least for the pilot, the EPPP operated under the principle that information obtained via an off-domain server is not part of the e-publication in question. Assuming that an e-publisher has obtained copyright clearance for all the objects on the domain of the e-publication, limiting an e-publication to one domain also greatly reduces the complexity of hypertext copyright issues.

After gaining more experience in the handling of hypertext e-publications, other objective criteria should be defined to determine the boundary of an e-publication.

IV.1.6. NLC Organizational Roles With Respect to E-Publications

During the pilot operation, the roles and responsibilities of organizational areas within the NLC with respect to e-publications approximated existing organizational assignments. However, the short time-frame of the pilot and the limited number of titles in it prevented the exploration of some NLC organizational areas with e-publication roles and responsibilities. Nevertheless, it is anticipated that NLC roles and responsibilities will remain as assigned for publications in other media.

It is expected that Acquisitions and Bibliographic Services (ABS) will be responsible for carrying out technical services activities, such as acquiring, checking-in, and cataloguing e-publications; and will generate requirements for and be the owner of the system components which support these activities.

Research and Information Services (RIS) would be responsible for presenting/delivering e-publications to the public, collections management, reference and research employing e-publications, etc.; and would generate requirements for and be the owner of the system components which support these activities.

Information Technology Services (ITS) would be responsible for the development, maintenance, and operation of systems supporting the various functions for the delivery of e-publication services. It is noted that many or most of these functions can be supported through automation, including ascertaining the types of access and features clients require. After reviewing such feedback, RIS personnel would request system changes and allocate system maintenance priorities as is done with other systems supported by ITS.

If the NLC adopts the e-publication preservation approach recommended in the Preservation chapter of this paper, then the existing NLC informatics branch (ITS), would support the operation of the e-publication system. Information Resource Management (IRM) would be responsible for initiating and coordinating policies and strategies for e-publication collection development and preservation.

However, effective policies and procedures governing the major activities in managing and providing service from e-publications cannot be developed without NLC personnel being well-versed in relevant systems operations. Therefore, selected/appropriate NLC personnel from various areas of the Library need training in systems and procedures for e-publications

The full-time equivalent of some four people from ABS is currently working on handling e-publications in the EPPP. This level of personnel resource allocation

will be expanded even though a production level e-publication system has not yet been implemented. The EPPP system would continue until replaced by the production level e-publication system.

IV.2. Policies and Procedures

IV.2.1. Technical Services

The e-journals which were acquired through the EPPP are mounted on an NLC server. Technical services personnel in the NLC Acquisitions and Bibliographic Services Branch (ABS) access the journals on the server in order to do their work.

One of the major differences that the team discovered between working with printed publications and working with e-publications was that, unlike print publications which must be handled sequentially, tasks and procedures related to the processing of e-publications could be performed concurrently. Because the acquired publications were accessible on-line, a cataloguer could do descriptive work while another cataloguer did subject analysis, and a serials control staff member prepared to check-in the issue in the serials control module. This concurrent approach is a viable method of handling e-publications. The NLC e-mail facility is used extensively to expedite the process.

The first e-journal which was acquired, checked in and then catalogued was *InfoCyle*. This publication was the "test" title, used to determine whether or not the technical services procedures drafted and revised by the project team were

be spread to a greater percentage of the operational staff through ongoing technical training.

In order to handle e-publications at the NLC, personnel resources must be dedicated to acquiring and controlling e-publications as part of the Library's regular operations. The acquisition, control and collection management of electronic publications should be integrated into the operational workplans of the Library's various areas. However, with resource cuts at the NLC amounting to 27.5 percent of the base budget over four years, it is obvious that additional personnel resources will not be available to support the handling of e-publications. This means that resources for other priorities within the Library must be reallocated to accommodate e-publications within existing operations. The NLC must also set processing priorities for e-publications, as is done with other forms of material, but it is important that e-publications should compete for resources as part of each area's regular workload and not as an experiment or special project.

In terms of the regular handling of e-publications, it is expected that the largest impact on personnel resources would be in the technical services areas of the NLC. After e-publication technical-services processing is automated, about the same amount of personnel time per title would be required to process an e-publication as a non-e-publication -- with the exception of the following extra e-publication activities:

- identification and selection of e-titles for legal deposit, and,
- negotiation of acquisition and public access rights with e-publishers.

It is estimated that the full-time equivalent of two people in acquisitions should be reassigned to handle a current estimated workload of 500 to 1000 e-publication titles per year. Of course, the number of e-publication titles per year is expected to grow[12].

Wherever possible, systems solutions should be found to assist in streamlining the operations involved in acquiring and processing electronic publications. The Manage Electronic Publications Requirements chapter of this report provides application features that are intended to streamline the handling of e-publications to minimize the human resources that would be devoted to e-publications.

The number and quality of e-publications are growing rapidly. In order to preserve the Canadian published heritage, the NLC must establish policies and procedures to handle the technical services aspects of e-publications as part of the mainstream of technical services activities.

[12] There are more than 27 000 Web sites on the Internet and that number is doubling every 53 days (Sun Microsystems, April '95). A significant percentage of these sites are Canadian.

Dewey subject listing) on the server (e.g., on the NLC Gopher, or, on the WWW).

The acquisition staff member then notifies the descriptive cataloguer, the subject cataloguer and the Serials Record Section (SRS) staff member that the e-journal is now available under a specific menu option on the server. Any additional details about the e-journal (whether in electronic or print format) should also be forwarded to the descriptive cataloguer, the subject cataloguer and the SRS staff member at this time.

Acquisitions issues and problems

Titles which matched the EPPP title selection criteria were searched on the Internet; subscriptions were set up for new journals, and, gradually, a list of electronic titles, which were possible candidates for inclusion in the project, was drawn up. Then began the task of contacting the publishers of each title selected to obtain the formal right to mount the e-publication on the NLC server. By June 1995, clearance had been obtained for more than 45 titles.

The task of obtaining the right to mount titles on the NLC server proved a major, time-consuming undertaking, though the results have generally been positive. E-mail messages or phone calls, in some cases both, must be exchanged, questions from the publishers answered and so on. This has been and continues to be an ongoing and sometimes difficult task. Only two refusals were received during the pilot; both occurred because EPPP did not provide usage statistics for the publishers' specific titles. This e-publisher requirement has now been satisfied. Examples of issues and problems include the following:

- Some journals are available in several versions; for example, *CTHEORY* and *International Teletimes* are available in a WWW edition and an e-mail edition. *Surfaces* is available in Mac, MS-DOS and ASCII formats. They may also exist in different word processor versions, such as Microsoft Word (TM) or WordPerfect (TM). It has been recommended that if standard versions of an e-publication exist, only these standard versions should be acquired. Since more than one version was acquired for several titles, staff gained valuable experience in handling different title formats in the NLC environment.

- Roman alphabet, including French, diacritics did not display properly in some types of documents; this problem was related to using Internet SMTP e-mail for document transfer.

- At the beginning of the Project, the headers for issues received via e-mail were stored in separate files. These separate header files caused confusion among users and it was to merge headers with their respective issues.

- Subscribing and receiving journals automatically by e-mail on the NLC server has not created particular difficulties. But e-journals disseminated via the Web often have various component parts and hypertext links to other publications, which makes them difficult to assemble and acquire automatically. It has since been decided to "mirror" WWW publications and perform periodic copies of the whole WWW publication source rather than perform file by file transfers. This is among the recommendations in the Manage Electronic Publications Requirements chapter of this report.

- Further to the integrity general issue, e-publishers often change or delete back issues of an HTML e-publication when presenting a new view of the e-publication including the new current issue. If the current version of the e-publication is simply mirrored, then some of the retrospective information is lost. During the EPPP, this was addressed as follows:
    1. The situation is recognized and/or staff are notified of the retrospective change by the e-publisher. Automated means to help recognize this situation are requirements for the production level e-publications system.
    2. A new folder area is established for the e-title.
    3. The existing structure and contents of the e-publication are moved from the current area to the new area. Registry information describing the new area, including a “new” title for the retrospective information, is created.
    4. The current e-publication structure and contents are mirrored to the original area established for the title.

  One of the problems of this approach is that one e-journal may have several titles in the registry files with overlap in terms of issues and content of these titles. This might lead to confusion in access. However, in order to preserve all issues of the e-title, such an approach is necessary.

- Many of the problems associated with e-journals were related to file handling:
    - Since an “issue” of an e-title could consist of multiple files (one for each article or review, etc.), it was difficult to identify each file uniquely and yet remain within the eight-character limit imposed by the VAX VMS environment. (This constraint was removed when NLC’s e-journals were moved to the Alpha UNIX file server in June 1995.)
    - A related problem occurred when the issues/holdings for these titles had to be listed as separate files on the server. If obscure file names

had to be used to represent articles, reviews, or issues, they would cause location difficulties for users.

- Some publishers combined all their back issues into a "log". When a new issue was distributed, the old issue was added to the log. This approach raised several questions for staff. As the articles in the log were originally separate publications, should they be archived that way or, since the publisher had archived them in the log format, should the NLC archive them in accordance with the publisher's format? Should current issues be deleted on the NLC server when the log format for that month became available at the FTP site? Staff time was taken up in managing this update because it required the new log to be retrieved via FTP each time a new issue was available and the "old" issue to be deleted from the NLC server when the new one appeared. The best way to handle this operational question is being investigated.
- To save space on their servers, publishers often used various compression utilities to archive back issues of e-journals. To access these files and mount them on the VAX VMS server for viewing by users, staff were required to learn how to use several types of compression utilities.

Differences between the directory structure on the NLC server and the one on a server NLC was mirroring also caused problems. The publisher Duthie on the Wimsey server in British Columbia made access to three e-journals available through one title, not three. *The Reader*, *The Children's Reader* and *The Proofreader* were all situated in one directory called *The Reader*. In order to list these publications separately on the NLC server, three distinct titles, which all point to the same URL, were created.

Acquisitions recommendation

Assuming good HTML practices, such as the use of relative URLs[13] in an e-publication, are employed, automatically mirroring HTML publications saves time for acquisitions and serials control staff. It is expected that other time-saving practices will be discovered as other forms of e-publishing become widespread.

The e-publication acquisition procedures, such as mirroring HTML publications, should be automated to allow technical services personnel to complete the acquisition process without informatics personnel support.

[13] A relative URL provides storage location information relative to the full "root" URL of the e-publication rather than being a new full URL. Thus, an entire e-publication can be moved or duplicated to a new storage location and only the "root" URL has to be changed.

IV.2.1.3. Acquisitions Selection Criteria

1. In general, the *National Library of Canada Collection Guidelines*[15] should be applied when the NLC selects e-publications.

2. The e-publication should be Canadian, that is, published in Canada and/or, on a Canadian server, with a Canadian editor, or with major Canadian contributors (a definition of a Canadian server must be established). This criterion would be refined with experience so that, for example, the NLC could perhaps add Canadian content publications residing on foreign servers.

3. E-publications can cover all subjects, but with minimal resources, emphasis should be placed on the areas of special emphasis: government publications, literature, history, music and library science.

4. E-publications acquired by the NLC should include government e-publications, especially federal government publications.

5. A title is collected in all the media in which it is available (for example: print, CD-ROM) and in all standard networked versions.

6. If resources do not permit acquisition of all e-publications, the titles should be those which are not available in any other media so that at least one version is preserved.

7. E-journals should be complete journals, not subsets of a title and not advertisements or indexes of more complete titles in other media.

8. Both free and priced e-publications should be selected. In the near-term, the focus is being placed on Canadian government e-publications, and non-commercial e-publications.

9. Both monographic and serial e-publications should be selected.

10. If an e-publication is available in two or more standard networked versions, at least two standard networked versions should be selected.

IV.2.1.4. E-cataloguing in Publication, or, Metadata

Another way to implement selective legal deposit would be to position the NLC as a registration and naming authority within a system such as the proposed

[15] National Library of Canada. Collection Management Policy Team Working Group on Collection Guidelines for Canadiana. *National Library of Canada Collection Guidelines*. 1990.

Uniform Resource Name (URN) being developed by the Internet Engineering Task Force. As in the current Cataloguing-In-Publication (CIP) and International-Standard-Serial-Number-(ISSN) programs, Canadian publishers would contact the NLC to register the existence of a network publication and obtain a standard identifier for the work. As part of this process, they would submit descriptive data about the item (similar to the data provided by the CIP information form, the proposed Uniform Resource Characteristic, the Text Encoding Initiative header, and/or the data covered by the Government Information Locator Service). The NLC would then review this data and determine whether it wanted to require deposit of the entire item. In return, the item would be included in some form of value-added directory/register coordinated or maintained by the NLC for networked information resources. This directory/register could be linked to the publisher's original site(s) for the item as well as to a permanent storage site that would be publicly accessible under certain conditions (e.g. after a given period).

The NLC will analyze and define its requirements for the identifying data, i.e., metadata, that should be published as an integral part of e-publications. The considerable work that has already been done in this area[16][17] should be analyzed with and by Canadian e-publishers and standards groups to determine specific Canadian and library requirements.[18]

[16] For example, see URL:
http://www.oclc.org:5047/oclc/research/conferences/metadata/dublin_core_report.html for a description of discussion and analysis leading to the "Dublin core" metadata proposals.
[17] Treasury Board of Canada Secretariat. *The Internet - A Guide to Internet Use in the Federal Government.* Appendix C: Proposed Format and Presentation Elements, Proposed template for Canadian federal documents. 1995.
URL: http://www.tbs-sct.gc.ca/tb/pubs/in/ine.html
[18] Through its international role within ISO Technical Committee 46 and IFLA, NLC could propose the international development of library requirements for e-document metadata.

Assuming that:

- it is agreed by e-publishers and NLC partners that an e-publication directory/register is highly desirable to facilitate access to e-publications and to aid in the systematic handling of e-publications; and,
- maintenance of an e-publication directory/register would be a byproduct of the NLC's handling e-publications and would not involve any additional personnel resources;

then, the NLC should work towards establishing an e-publication directory/register consisting of standard metadata about the individual e-publications.

IV.2.1.5. Serials Control

Procedures for check-in of serial e-publications

When the serials control staff member receives notification from the acquisitions staff member that a new e-journal is available on the server, he/she accesses the server, verifies the e-journal, then follows standard check-in practice to set up the item in the Dynix Serials Control Module, establishing a publication pattern, etc. (Obviously, it is not possible to attach a barcode to the issue of the e-journal.)

For current issues received regularly via a subscription to an e-mail list, each time the e-journal is published, either the issue itself, or, a publication notice is received. In both cases, serials control personnel are automatically notified of the receipt of another issue, and access the server to verify the new issue. When the issue itself is received, the serials control staff member checks it in on the Dynix Serials Control Module in the usual way. When only a publication notice is received, the serials control staff member signs on to the NLC server, directly accesses the host server from which the e-publication can be retrieved, and copies the appropriate file from the host server to the NLC server.

For HTML serial publications for which the NLC does not have automatic mirroring privileges, the host server is accessed regularly (e.g. monthly) and current file(s) are copied to the NLC server.

For HTML serial publications for which the NLC does have automatic mirroring privileges, the serials control staff member activates the mirroring program on an "as required" basis.

On an ongoing basis, as each issue of the e-journal is received, the serials control staff member manually inputs the volume and issue number of the received item on the Web menu.

Overdue issues are detected during the claiming process in the Dynix Serials Control Module and the serials control staff member sends an e-mail message to the publisher of the e-journal to determine why an issue is late or if there a problem, such as change in frequency, cancellation, etc.

Serials control observations

Once the e-journal title has been acquired and catalogued, the task of tracking and controlling its regular receipt is the responsibility of the Serials Records

For example, when reviewing the error log on the NLC server (unmatched titles), should the serials control staff member discover that the title of an e-journal has changed, then serials control provides acquisitions with details of the change, so that the original order/request can be modified, or, a new order/request record created. (The acquisitions staff member may need to access the issue with the title change on the NLC server to complete this work.)

If any other staff members notice a title change or new issues, the information should be passed to the serials control staff member who will start the process for handling the changes. This includes notifying both the descriptive cataloguer and the subject cataloguer.

Serials control recommendation

As in other technical services areas, extensive automation and streamlining are required before the NLC can handle a significant volume of e-journals. Examples of required features include automatic checking of HTML host servers to determine whether or not to capture the new “issue” of an HTML e-publication, and automated checking of e-mail subscriptions status.

IV.2.1.6. Cataloguing

Cataloguing procedures

Draft descriptive cataloguing guidelines had been developed before the start of the pilot project. These are based on the Anglo-American Cataloguing Rules, OCLC's Guidelines for Bibliographic Description of Internet Resources (1993) and Machine Readable Representation of Bibliographic Information (MARBI) Committee discussion papers. Following these guidelines, bibliographic records for titles being used in the pilot project began appearing in the NLC’s AMICUS database and in the NLC's in-house Dynix On-line Public Access Catalogue (OPAC).

These guidelines call for the use of the new MARC field 856 and its sub-fields (the new MARC tag for information related to e-publications). Because NLC systems could not handle field 856 at the beginning of the project, it was decided to use the field 500 (general note) until the 856 field could be included in the NLC cataloguing system AMICUS. Field 856 was subsequently validated for use in AMICUS. Therefore, this field is now used instead of field 500 on all new records for e-publications. In addition, all records created before this validation have been revised to change the 500s to 856s.

Universal Resource Locators (URLs) of EPPP titles are recorded in the catalogue record on both the cataloguing module on AMICUS and in the OPAC

maintained on Dynix. During the pilot, a cut-and-paste link from the URL record to the WWW/Gopher servers is used for access to electronic documents. This cut and paste will be replaced with an improved link after the pilot.

Publishers of several e-titles included in the pilot project had applied for ISSNs before the EPPP started. ISSNs are now being assigned to the remaining titles and titles are being registered in the ISSN Network database.

Subject analysis personnel prepare and input subject headings (including French or English equivalents), Dewey classification numbers, and any necessary subject-related notes and fixed fields. The subject cataloguer also provides subject access to the journals mounted on the Web server. The Dewey Decimal Classification Schedule is used to provide access to the e-journals by subject category,. The Schedule was reviewed and edited in both English and French and the schedule information was prepared so that it could be mounted on the Web server. Because the subject cataloguer is able to add up to three Dewey Class numbers to identify each e-journal title in the project, it is possible to provide broader subject access to the end user.

An example of a typical catalogue record for an e-serial created during the EPPP is given at Appendix G.

Cataloguing observations

Comments on the bibliographic record for the first e-journal catalogued for the Project - *InfoCycle* - were received from several areas in the library, e.g. staff in RIS provided some useful insights, and some revisions to the record were subsequently made. Overall, significant practice in using the new tag and its sub-fields was achieved.

Subject cataloguers experienced difficulty in doing subject analysis in an on-line environment. Flipping back and forth between the files for issues and articles was quite time-consuming; in fact, the cataloguing of on-line titles took much longer than the cataloguing of print equivalents.

Once the handling of e-publications moves into the mainstream of NLC processing, all Canadian e-titles collected by NLC will be included in Canada's national bibliography *Canadiana*.

Inclusion of e-titles in Canadiana, and the resulting publicity for the title, will help to encourage e-publishers to follow the legal deposit regulations and send their e-publications to the NLC.

limited duration of the EPPP, these recommendations are based on research and analysis. Further literature searches, dialogue with those already archiving e-journals, and NLC experiments will determine their feasibility, and indeed, whether they are "real" preservation issues in a networked environment.

Authenticity and completeness of e-publications

The "NLC local copy" (since most of the NLC e-journals will presumably be on servers elsewhere) must be in sync with the "official copy".

E-publication acquisition must be accompanied by automatic validation and audit of the content. This may involve:

- obtaining commitment from publishers that updated versions will be brought to the attention of the NLC;
- hashing or digital time-stamping the overall document (all component parts together) as it arrives, or, hashing or digital time stamping each component part;
- keeping track of the "periodicity" of changes to an e-publication and taking snapshots accordingly;
- developing automated mechanisms whereby acquisitions staff can compare "issues" or documents to determine if the "copy in hand" is the same as another "copy".

E-publication security encompasses both protection of the contents from deliberate or accidental change and back-up mechanisms and disaster contingency. Current mechanisms for backing-up, storing back-ups, and recovering/reconstructing the NLC databases, and emergency policies and procedures currently in place for the NLC LAN servers, bibliographic databases, etc. should partially address this issue. Methods, such as hashing or digital time-stamping, as outlined above, should also protect the issues from being changed (as should the current mechanisms, which determine who is authorized to access and change records).

E-publication compression

The concept of an original document is almost unanimously rejected in the online, networked electronic environment. However, the information should remain as true to its original intellectual content as possible.

One issue in the fidelity of the file is that of file compression. Choices in the compression and decompression of a file containing an e-publication may yield an e-publication that is different from the original prior to compression. This

could be an issue for preservation. The decisions for NLC are whether to compress previously uncompressed e-publications for storage, and whether to store e-publications in compressed mode, when they are received as compressed files.

Because certain compression algorithms lead to some loss of information from the original, and because this loss of "faithfulness" to the original is not acceptable for preservation purposes, the NLC should avoid compressing files for long-term storage whenever possible.

In a situation where a publisher ships an e-publication in a compressed format only, all consumers of the e-publication would suffer a similar loss of detail. The publisher made the decision to compress the file and risk information loss, and the NLC is acquiring and preserving a copy of the e-publication which is "equal in quality to the highest priced edition" of the work, as required under Legal Deposit.[20] However, when a networked electronic publication arrives compressed, it should be stored in uncompressed format because the preservation of compression/decompression technologies presents unnecessary additional complexity. The main reason for compressing data is to save in data-transmission cost and time; storage savings are not appreciable.

Even given the above recommendation, the NLC may want to accept some data loss through compression, in certain cases. For example, audio or video digital objects can be very large, and compression/decompression can be such that data loss is undetectable by the human eye or ear. Such cases must be identified and agreed upon using accepted criteria. Some may have to be judged individually. It would be impossible to avoid compression entirely. Extreme proof of this is that the entire run of 16 volumes of *PC Magazine*, including all pictorial matter stored in a high-resolution format, would need a quantity of terabytes of storage space to contain it in electronic form. And that is just one serial run.

The NLC will compress for preservation only those files and objects for which there is no detectable loss of information ("no detectable loss" must be defined through experimentation).

As with any other software techniques needed to read e-publications, only standard compression/decompression techniques would be employed.

E-mail packaging of e-publications

[20] Canada. *LEGAL DEPOSIT, National Library Act (excerpts with respect to legal deposit) and National Library Book Deposit Regulations*, 1995.

Concerning e-mail addresses and headers, it was first decided during the EPPP to retain and store this information with the e-mail e-journal issues. This practice can lead to an unattractive and somewhat confusing presentation of these e-publications.

In the case of e-journals delivered to the NLC via e-mail, the e-mail headers should be stripped from the text prior to storage.

Storage/preservation of e-publications

In the shorter-term, until the NLC can implement an e-publication architecture such as that described later in this paper, the NLC will store e-publications online on magnetic disk.

For the mid- to longer-term, preliminary observations indicate that a hierarchical storage management system (HSM) should be the preservation solution where current, currently requested, or frequently accessed materials are stored online on high speed retrieval media (currently, magnetic disk); where less frequently accessed material is stored on near-line media (such as, currently, CD-ROM); and low demand material is stored off-line (such as various magnetic tape media). If an e-publication has been migrated to off-line storage, the user accessing this material would be notified of its availability and there would be significant delay (minutes or hours, or next day) in access; then the requested off-line files are migrated online for viewing.

It might be ideal to find a long-term e-publication storage medium like permanent paper or microfilm which, under appropriate environmental conditions, has a

A current concern is that no electronic media are accepted as long-term archival media. There are further concerns about the costs and difficulties of having to transfer information from tape to tape and CD to CD every four to five years. Because of the ease of copying e-publications seamlessly and without damage to the original (unlike the print equivalent), preservation and access in an electronic environment will become more closely entwined. HSM systems can address most of these concerns:

- An HSM can refresh media automatically at appropriate periods.
- New media and storage devices are introduced into the HSM system as they become available and commercialized.
- An HSM can move an e-publication automatically between online, near line and off-line depending on the demand to access the e-publication or other criteria. Online e-publications are available immediately, response time is slower than online for access to e-publications stored on near line media, and, for e-publications stored off-line, the researcher would be informed of the existence of the e-publication, and when it would be migrated from off-line to online for viewing. Depending on the load on the e-publication server and other factors, such as the size of an e-publication being moved from off-line to online, the time to migrate an e-publication for viewing could vary from a few minutes to overnight.
- The HSM system automatically backs up on a specified period all the data under its control; all the data online, near line and off-line are automatically backed up. The HSM comes with utilities to restore or reconstruct files as necessary.

If an HSM is acquired by the NLC, it would be acquired through a competitive Request For Proposal in which the NLC requirements for the HSM would be specified and would include conformance to the NLC systems architecture, and having the capability of being integrated with the rest of the proposed NLC e-publication technical architecture. The NLC requirements for an HSM are further discussed later in this paper.

Decentralization of e-publication preservation

Decentralizing the responsibilities for preserving e-publications is a topic which has been raised numerous times in the EPPP. Arguments against the NLC attempting to archive "all Canadian e-publications" generally centre on the resources necessary to do this, coupled with the ease of sharing the responsibility for e-publications in an electronic environment. Regardless of decisions re: centralized/decentralized preservation storage, it is imperative that the NLC discuss cooperative activities related to the identification, acquisition, management, and preservation of e-publications.

The NLC must examine various decentralized preservation storage options. One scenario might be to have institutions across Canada responsible for specified sets of e-publications, migrating them to new storage media and keeping these media in environmentally sound conditions so that the information will remain available to future generations. Storage of one copy of a publication at a named site may be sufficient; or two or more copies may be stored at several locations. Bilateral decentralized arrangements are desirable, but will be time-consuming. In another scenario modeled on the NLC's responsibilities as repository for the microform masters of Canadian newspapers, the NLC might assume a role in storing electronic hard copies (e.g., diskettes, tapes, CD-ROMs, etc.) of publications which are stored on remote servers.

A third scenario might be to work with other federal players and, through national consensus, create a central repository for electronic information. This, in the end, might be the most cost-effective way to ensure the long-term availability of the published heritage. Regardless, the NLC must be prepared to act as the archive of last resort for Canadian e-publications and should take the lead in examining decentralization of e-publication preservation responsibilities.

There are related cooperative initiatives underway in Canada, such as Can-Linked[21]. Can-Linked focuses on digitization of collections. Note, however, that preservation of the digitized collections will involve the same preservation issues as networked e-publications.

The NLC and its partners should explore other Canadian cooperative library programs, such as the Canadian Cooperative Preservation Project and the Decentralized Program for Canadian Newspapers to determine the applicability of such models to the preservation of Canadian e-publications.

Identification of e-publications for long-term preservation

Preservation at any stage of a publication's life cycle is concerned with provision of access, both immediately and at some point in the future. Because of the ease of copying seamlessly and without damage to the original, preservation and access to publications in an electronic environment should become more closely entwined.

The NLC will acquire some e-publications and place them on NLC servers, and will point to e-publications which exist on remote servers. Some acquired e-

[21] *The CAN-LINKED Initiative: a proposal for the co-ordinated development of a distributed national digital library system in Canada*. Ottawa: a group of Academic and Research Libraries and the Canadian Association of Research Libraries, 1995

publications will be considered for long-term storage and preservation whereas some (e.g., dictionaries which are brought in to NL servers to facilitate access, but which do not fit the definition of Canadiana) will not.

"Identification" is the process of choosing items from the pool of e-publications for long-term preservation. There is some debate over whether there would be some sort of "first filter" applied at acquisition to determine whether a publication is to be "archived". If a publication is "acquired", in the sense of being mounted on a local server, then perhaps it should be identified for long-term preservation. However, there are cases where the NLC might temporarily mount an e-publication on a local server to expedite access, but would choose not to store it long-term (e.g., earlier issues of publications which cumulate).

The NLC must delineate the selection, collection management and preservation policies for e-publications, and determine the criteria by which publications will be identified for long-term preservation.

This report recommends that the NLC acquire and provide access to standard versions of e-publications in section IV.1.3. Electronic Publication Integrity. In order to provide access, such e-publications must be preserved.

E-publication access software/hardware issues

Access software may be proprietary, shareware, freeware, or in the public domain; and may include word processing software; WWW browsers (e.g., Mosaic, Netscape, Cello, etc.); viewers for audio, video, graphics, and other types of files; software for compression and decompression of files of all types; and could include software specific to various types of hardware such as Macintosh and PC computers.

Proprietary software is frequently upgraded and sometimes loses "backwards compatibility", i.e., the ability to handle electronic files created for use with an earlier version of the software. Monitoring the functionality of access software will be mandatory as new versions appear, and retention of older versions of access software may be needed to ensure continued access to publications. Vendors of the more popular packages usually present clear comparisons of new and older versions as selling points, so the monitoring of this software may not be difficult. If e-publications that require proprietary software are retained (this is contrary to the recommended practice of this report), then e-publications should always be converted to the newest version of the software. If the proprietary package disappears from use, then every effort should be made to convert the e-publication from a proprietary format to a standard format.

Shareware, freeware, and public-domain software are also subject to upgrade, but the mechanisms for announcing and distributing new versions are sometimes inadequate or lacking. The NLC should not attempt to retain and archive e-publications in shareware, etc. format.

There is also potential for obsolescence of either the EDP equipment or the software platform upon which access software is run. For example, if DOS were to become obsolete, it might be necessary to maintain a DOS-based PC in order to access publications using DOS-based access software. It is not reasonable to consider archiving hardware; if hardware obsolescence precludes access to publications in the future, publications would have to be converted from one format to another. Additionally, decisions on how to transmit knowledge of the "look and feel" of the original format would be needed.

IV.2.3 Public Services

Early in the EPPP, it was decided to use the World Wide Web as the primary means of access to the NLC collection of e-publications. Some of the journals were available on the Web only and could not be adequately viewed other than via a Web server and client. The Web appears to be the preferred means for formal e-publishing. The NLC has established a Web site for the EPPP at the URLs: http://www.nlc-bnc.ca/eppp/e3pe.htm (English) and http://www.nlc-bnc.ca/eppp/e3pf.htm (French). A bilingual pre-page with general information about the NLC and pointers to the EPPP was established in line with evolving federal government standards.

The NLC recognizes that Internet users have various means of access and that the Canadian public has access to a growing number of freenets. Gopher access (gopher.nlc-bnc.ca) was provided via the Canadian e-publications menu on the NLC Gopher and, resources permitting, file transfer protocol (FTP) access to the e-publications will also be provided. (The NLC has now migrated its Gopher services to the Web.)

The Web and Gopher access to the EPPP were opened to the public in March 1995. The EPPP became part of the overall NLC Web structure in June 1995. Initial access was alphabetic by title and by broad Dewey subject category. In August 1995, a title keyword search and full-text keyword search of the EPPP archive were added. A news release on the project was released in January 1995 and the Internet availability was announced on INFO_L (the NLC news dissemination list), on other discussion lists, and in *National Library News* in March 1995.

NLC's internetworking Acceptable Use Policy[22], NLC must be able to identify any NLC originator of an Internet transmission. Capabilities and security features of the public access workstations include:

- Read-and-search access to the Internet, including the NLC Web and Gopher servers, and the National Capital FreeNet is provided.
- Access to the NLC office systems environment is prevented.
- Access to the workstation hard drive and floppy drive is prevented.
- Patrons are uniquely identified when using the network, while maintaining their confidentiality.
- Security software is used to protect the Windows Program Manager so that it cannot be accessed.
- Some of the Windows Program Manager functionality is replaced by a user-friendly, menu-based, shell front-end. It, too, has security features, such as preventing access to DOS.
- NLC patrons are required to logon to the public access workstation using their NLC patron barcode.

NLC has also drafted an Acceptable Use Policy[23] for patrons who use the NLC's public-access Internet workstations. This document explains the policies and standards which NLC on-site patrons are expected to follow when using the NLC networked facilities, and what patrons can expect from the NLC staff assisting them.

It was decided that for the purposes of the EPPP, existing document-delivery policies and responsibilities will serve for the handling of e-publications. Delivery of print items is by loan of the item itself, photocopy of a portion of the publication, FAX, or electronic document scanning. For e-publications, a printout of a portion of the publication (if appropriate for the e-publication) may be provided, or the item (or part of an item) may be sent electronically by FAX or e-mail. Any form of delivery is dependent upon the remote client's capabilities for receiving the requested item.

Access policy

The reason for acquiring, cataloguing, preserving, and managing the collection of e-publications at the NLC is to provide long-term access to these e-publications, especially in support of Canadian researchers and educators.

---

[22] National Library of Canada. *Internetworking and Email Acceptable Use Policy (AUP).* 1995.
[23] National Library of Canada. Reference and Information Services Branch. *Public Internet Workstations : Acceptable Use Policy*. 1995.

NLC and will make the acquisition of e-publications easier.

1.2. Many of the e-publications acquisitions tasks and activities must be supported through automation. For example:
   1.2.1. The many variations in obtaining individual issues of serials should be automated as much as possible.
   1.2.2. Acquisition of HTML e-publications must be streamlined via automated mirroring, for example. Facilities should be developed so that acquisitions personnel set up and maintain the mirroring processes.
   1.2.3. When acquiring the current version of an HTML title, the NLC must develop automated means to help recognize when there have been retrospective changes or back issues have been deleted, and provide support for appropriate preservation action before the e-title is mirrored.

1.3. Copyright clearance and recording access rights should be supported by automation for greater efficiency. This support could include e-mail of forms to copyright holders to obtain access rights, and automatic recording of replies to trigger the setting of system parameters.

2. Serials Control

2.1. The serials-control functions must be automated and streamlined. Examples of required features include:
   2.1.1. A feature must be developed to check Web host servers automatically to determine whether or not to capture the new "issue" of an HTML e-publication.
   2.1.2. Automated mechanisms should be developed so that e-publication subscriptions can be verified as "active." This feature will ensure that the NLC is not inadvertently removed from a subscription list, in particular an e-mail e-journal subscription list.

3. Cataloguing

3.1. The AMICUS unique identifying number of an e-publication should be stored with the e-publication on the NLC e-publication server. This would be a link from the title on the e-publication server to the AMICUS bibliographic record of the title. Assuming that the e-publication was accessed by a WWW browser, the user could click on the AMICUS number to obtain the AMICUS bibliographic information.

3.2. The metadata to be stored and indexed with each e-publication must be specified (the e-publication cataloguing information is stored in AMICUS) including:
⇒ variants,
⇒ modifications,
⇒ evolution,
⇒ electronic signature.

3.3. As with other publications, different forms of the same e-publication must be identified and linked through the AMICUS bibliographic system; that is, there must be a means of distinguishing between forms so that requests are for the most appropriate format.

4. Union Cataloguing

4.1. The NLC needs to provide union catalogue services for e-publications by providing their locations (e.g. URLs).

4.2. The NLC union catalogue should also be a register of where and how a publication was digitized.

5. Public Services

5.1. Programming should create a dynamic link between the tag 856 information (e.g. URL) in the bibliographic record in AMICUS or Dynix and the title mounted on the NLC Web server. Clients should be able to click on the URL and be taken immediately to the title as it is listed on the Web server. For users of the AMICUS Graphical User Interface, clicking on the URL would invoke the Web browser that the AMICUS user had mounted at the workstation. For users of AMICUS character based interfaces, choosing a menu item or initiating a command to “read” the URL would invoke the user’s character-based Web browser. Data modeling will be required during development of this feature to determine where the URL type information should be stored in the AMICUS database.

5.2. Location and access of information in the NLC’s e-collection should be both via AMICUS (as outlined in these requirements), and via Internet access tools. The Internet access tools would use both e-publication metadata and e-publication text in the generation of indexes, etc.

5.3. The requirements for full text access to e-publications should be formulated to indicate which parts of documents should be full-text indexed.

5.4. FTP should be used for e-publication delivery where appropriate. The user could specify time of delivery, desired compression techniques, etc.

6. Preservation and Collections Management

6.1. The Manage E-publications application must be able to keep track automatically of which software versions are required to access every part of every e-publication in the NLC collection. This could be done by establishing a relationship between each file (i.e., object) in the NLC e-collection and the software package(s)/version(s) needed to access the file. When the package(s)/version(s) becomes obsolete, the new version of the package or a replacement package could be tested with an appropriate sample of the collection. If replacement software cannot be obtained, the objects in question should be converted to a form that would be readable by then current software. This recommendation includes any compression/decompression software required to access e-publications.

6.2. As part of the Manage E-publications suite of software, the NLC must acquire:

- ⇒ extensive conversion software to convert various formats to standard formats;
- ⇒ extensive standard compression/decompression software (NLC already has a significant library of this category of software);
- ⇒ standard encryption/decryption software.

7. Other

7.1. The links between AMICUS and Dynix should be improved so that all the data in the tag 856 can be downloaded from AMICUS to Dynix. Also, both AMICUS and Dynix must be able to handle all the data changes - fixed fields and variable fields - related to format integration.

7.2. The NLC should keep "copyright clearance" and rights management information for each e-publication in the NLC collection indicating the status of copyright clearance and whether or not the item is in the public domain. Free items would have

status “public domain” while chargeable items would have status “copyright protected”. The system should provide for overriding the default and setting the copyright information for an item.

7.3. The recommended technical platform architecture and configuration for support of Manage E-publications include a scaleable and flexible server computer system with the UNIX operating system, a commercial Web server with security and other features, an object management system or an e-document management system, a hierarchical storage management system, and HTML testing and validation tools. The figure below gives an overview of the technical architecture.

7.4. The pros and cons of using an object management system (e.g., the object management extensions to the Ingres relational database management system employed in AMICUS) versus using an e-document management system for storing and accessing e-publications should be analyzed. The system selected must interface with the hierarchical storage management system to be employed for preservation.

7.5. Requirements for the hierarchical storage management (HSM) system include:

7.5.1. The HSM system must have the capability of automatically identifying and migrating an e-publication, including all its objects, to/from near-line and/or off-line storage. This automatic migration should be based on parameters, such as the longest time period since last access.

7.5.2. The HSM system must refresh storage media automatically at appropriate periods.

7.5.3. The HSM must have a good track record of introducing new media and storage devices into the HSM system as they become available and commercialized.

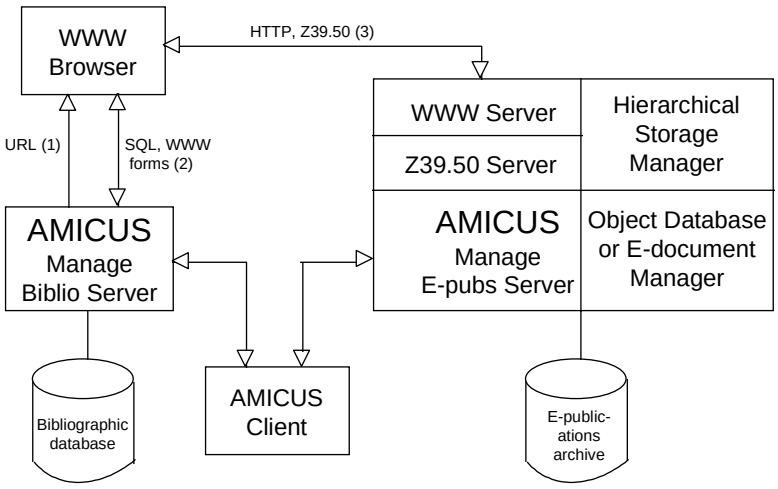
7.5.4. The HSM system must automatically back up on a specified period all the data under its control; that is, all the data online, near-line and off-line are automatically backed up. The HSM has utilities to restore or reconstruct files as necessary.

7.5.5. The HSM system must interface smoothly with the object management system or the e-document management system employed in Manage E-publications. In particular, an object or a group of objects comprising an e-publication must be HSM “addressable” entities capable of being migrated individually or as a group by the HSM system.

7.5.6. Application Program Interfaces (APIs) to the HSM must exist to enable the Manage E-pubs application to inform the user of the expected delay when accessing an e-pubs that is stored on near-line or off-line storage.

7.5.7. The HSM system must have a large enough installed customer base that, even should the HSM system supplier go out of business, there would be an excellent business case for another software company to maintain the HSM system.

Manage E-Publications Architecture Overview



0Notes concerning the Manage E-publications Architecture Overview:

The boxes in the overview diagram do not necessarily indicate physical computers. For example, the AMICUS client and the WWW browser can be on the same machine.

The WWW Browser is the client implementation of WWW. WWW browsers are commercially available.

The WWW Server is the server implementation of WWW. WWW servers are commercially available. WWW server packages are being continually enriched with such features as the ability to support secure financial transactions.

Hierarchical storage management (HSM) packages are available commercially.

Object database management systems (ODBMS) are available commercially. They include object implementations of relational database management systems. ODBMS are optimized toward storing large electronic objects (also known as “blobs”). The e-objects in e-publications, such as sound bites and movie clips, are very large. Electronic document management systems are available commercially and often include extensive facilities for format conversion.

Indexing and searching the e-collection metadata and text via Internet tools would be supported by the Manage E-publications server software, or by the ODBMS or e-document management system.

It is expected that the NLC would issue a Request For Proposal (RFP) for all the commercial software products in the Manage E-publications Architecture. It would be a mandatory requirement that all the components work smoothly together.

Following are explanations of the numbers in parenthesis in the above diagram.

- The URL stored in the holdings area of the AMICUS bibliographic database is a “hot key” to the e-publication that the AMICUS bibliographic record describes.

- SQL and WWW forms are employed to communicate between the AMICUS bibliographic server and client Web browsers.

- Hypertext Transfer Protocol (HTTP) is employed to communicate between the NLC e-publication server and WWW client browsers. This protocol also has gateways (called Common Gateway Interfaces, or CGIs) to other protocols, such as the Z39.50 search and retrieval protocol. Keyword searches of e-publication contents could be communicated by Z39.50.

V. Conclusion

Electronic publishing is exploding. This report has made recommendations to address issues in identifying, locating, acquiring, reporting, storing, preserving, and accessing e-publications to prevent major gaps in the Canadian heritage and research tools of the nation.

The NLC personnel who participated in the EPPP were very keen to learn about e-publications and perform EPPP assignments. Their high level of interest extended to all aspects of e-publishing. Perhaps because they were dealing with a new type of bibliographic entity, all participants felt the excitement of "breaking new ground". The cooperation among all areas of the NLC was excellent.

Appendix A: EPPP Project Organization

Project Review Committee: NLC Electronic Collections Committee with membership as follows: Francine Bédard; Rolande Blair (chair), Nancy Brodie, Alison Bullock, Marie Lanouette, Chantal Métivier, Alice McClymont, Elizabeth McKeen, Bill Newman, Jane Thacker.

Project Team

EPPP Project Leader: Bill Newman, Information Technology Services.
Technical Services (TS) Subproject Leader: Roselyn Lilleniit, Acquisitions and Bibliographic Services.
Public Services (PS) Subproject Leader: Donna Dinberg, Research and Information Services.
Informatics Services Subproject (IS) Leader: Yvon Lauriault, Information Technology Services.
EPPP Publicity Coordinator: Nancy Brodie, Information Resource Management.

Project team members and resource persons were assigned to a subproject as indicated below, where:

ABS is Acquisitions and Bibliographic Services,
IRM is Information Resource Management,
ITS is Information Technology Services,
MAP is Marketing and Publishing,
RIS is Research and Information Services.

advertisements or publication catalogues for the print publications or permit electronic communication with the publishers. Other sites contain the table of contents of the latest issue of a magazine, with perhaps the text of an editorial or one or two articles. Others contain a substantive portion of the print publication and may contain additional information, that is not in the print edition, or valuable links to related information. Karen Hebb has analyzed trends in Canadian newspaper publishing on the Internet.[26] The NLC is currently developing criteria for selection of these variant electronic editions.

[26] Hebb, Karen. The future of online publications: are they a replacement for or a supplement to newspapers? Halifax: University of King's College, 1994. (Honours thesis, School of Journalism) URL: http://www.ukings.ns.ca/Docs/onlineeditdocs/Studentpages/KarenHebbThesis.html

IDRC							

Title	Format	Distribution	Method of acquisition	ISSN?	Paper equiv-alent?	Genre	Peer review-ed?
Information, communication et le CRDI SEE Information, Communication and IDRC							
International Teletimes	html	WWW	mirror		No	Broad, international culture magazine	No
International Teletimes	ascii	ftp	ftp	ISSN 1198-3604	No	Broad, international culture magazine	No
National Library News	ascii	gopher	from NLC Gopher	ISSN 1195-2326	Yes	Newsletter; government	No
Network Notes	ascii	gopher	e-mail	ISSN 1201-4338	on demand	Series of information sheets; government	No
Nouvelles de la Bibliothèque nationale SEE National Library News							
Surfaces	ascii	ftp	ftp	ISSN: 1188-2492	paper, diskette	Scholarly culture journal	Yes
The Children's Reader	html	www	mirror	No	paper	Annual review of children's literature	No
The Disseminator	ascii	e-mail, gopher	gopher	No	No	Newsletter; government	No
The Disseminator	html	www	e-mail: zipped archive	No	No	Newsletter; government	No
The Proofreader	html	www	mirror	No	No	Book reviews	No
The Reader	html	www	mirror	No	paper	Book reviews and announcements	No

Titles in progress 30 June 1995 (Publishers have agreed to participate)

Title	Format	Distribution	Method of acquisition	ISSN?	Paper equiv-alent?	Genre	Peer review-ed?
@ Uplink			mirror				
Capital Current			mirror				
Chorus	html	WWW	mirror	No	No	Academic and educational magazine	No
Ctheory, WWW ed.	html	WWW	mirror	No	No	Scholarly book reviews	No
DSP's weekly checklist			mirror & gopher		Yes		
Electronic Journal of Sociology			mirror				
EMLS: Early Modern Literary Studies	html	WWW	mirror	ISSN 1201-2459	No	Scholarly journal	Yes
Euphony	html	WWW	mirror	No	No	Movie & music magazine; includes sound clips	No
Global Monitor	html	WWW	mirror	No	Yes	Computer news magazine	No
green Cart	html	WWW	mirror	No	No	Magazine	No
Ideas DIGest	html	WWW	mirror	No	Yes	News magazine for small business	No
InterFace	html	WWW	mirror	No	Yes	Technology magazine	No
Info tech			mirror				
International Journal of Continuing Education			e-mail				
Lytx			gopher				
Media Tribe			mirror				
Money Issues			mirror				
NWHQ	html	WWW	mirror	No	No	Art magazine	No

Title	Format	Distribution	Method of acquisition	ISSN?	Paper equiv-alent?	Genre	Peer review-ed?
Stage Works	html	WWW	mirror	No	Yes electronic version includes highlights only	Technical journal	No
TAC: Theory and Applications of Categories			mirror				
Ways Online	html	WWW	mirror	ISSN 1198-7324	Yes	Business news magazine	No

Table of Copyright Statements

Titles acquired to 30 June 1995

Title	Copyright statement
Ability Network Magazine	All articles are copyrighted and copying for other than personal reference use without express permission is prohibited.
Axe	(c) Axe & J. Przychodzen
Big Dreams	copyright (c) Dare to Dream Enterprises. Permission is hereby given to print out and distribute unlimited exact copies as long as no fee is charged.
Canadian Journal of Educational Administration and Policy	Copyright 1995, the Canadian Journal of Educational Administration and Policy Archives. Permission is hereby granted to copy any article provided that CANJEDADPOL ARCHIVES is credited and copies are not sold.
Cropduster	copyright 1992 by Steven Meece and Chris Woodill. Naturally permission is granted to distribute Cropduster in any way you would like, but please leave it as it is so that others can see our mistakes as well.
CTheory	Example: Copyright of this article is shared by the author and CTheory. Permission to reprint requires the written permission of the author and CTheory.
DiSCORDER	DISCORDER 1994 by the Student Radio Society of the University of British Columbia. All rights reserved.
FLASH INFORMATION	Copyright Centre for Information Technology Innovation, 1995
Flash réseau	
SEE Network Notes	
Government Information in Canada	may be freely copied, distributed on condition that the journal and its contents are not altered, journal and authors are credited, and there is no charge for the journal or its contents. Rights for reprinting, republication for profit reserved for authors
IDFORUM	No statement
infoCycle	All material copyrighted (except as noted).
Information, Communication and IDRC	No statement
Information, communication et le CRDI	
SEE Information, Communication and IDRC	

Title	Copyright statement
International Teletimes	International Teletimes is a publication of the Global Village Communication Society and is copyrighted ©1993 by the same. All articles are copyrighted by their respective authors. However International Teletimes retains the right to reprint all material unless otherwise requested by the author. This magazine is free to be copied and distributed UNCHANGED so long as it is not sold for profit.
National Library News	Articles published in National Library News may be reproduced without permission, but a credit note would be appreciated.
Network Notes	No statement
Nouvelles de la Bibliothèque nationale	
SEE National Library News	
Surfaces	Copyright for texts published in SURFACES remains the property of authors. However, any further publication should be accompanied by an acknowledgement of SURFACES as the place of initial publication.
The Children's Reader	No statement
The Disseminator	No statement
The Disseminator	No statement
The Proofreader	No statement
The Reader	No statement

Titles in progress 30 June 1995 (Publishers have agreed to participate)

Title	Copyright statement
@ Uplink,	
Capital Current	
Chorus	
Ctheory, WWW ed.	
DSP's weekly checklist	
Electronic Journal of Sociology	
EMLS: Early Modern Literary Studies	[on article]Copyright (c) 1995 by the author, all rights reserved. Volume 1.1 as a whole is copyright (c) 1995 by Early Modern Literary Studies, all rights reserved, and may be used and shared in accordance with the fair-use provisions of U.S. copyright law. Archiving and redistribution for profit, or republication of this text in any medium, requires the consent of the author and the Editor of EMLS.
Euphony	All design, layout, and original Euphony artwork contained within the internet web site Euphony Net-Magazine (http://euphony.com/euphony/) is Copyrighted 1994, 1995 by Brad Pitzel. All rights reserved. This copyright includes, but is not limited to: graphic artwork, html coding, xhtml coding, the classified ad software, and all other custom software written to enhanced the production or look of Euphony Net-Magazine. Original photography, and articles texts are copyrighted by the respective authors. No material contained within the web site Euphony Net-Magazine may be copied, cached, mirror, or duplicated by any electronic or physical means, except for private non-profit use.
Global Monitor	None
green Cart	
Ideas DIGest	All rights reserved. No part of this publication may be reproduced in any form without the express written consent of the Publisher or Editors.
InterFace	(c) 1994 InterFace Magazine IFM/94 Please be advised that these works of art are free for download and personal use. They are prohibited from unauthorized publication. Contact the artist or InterFace Magazine for such authorization.
Info tech	
International Journal of Continuing Education	
Lytx	

Title	Copyright statement
Media Tribe	
Money Issues	
NWHQ	Copyright © 1994 Knossopolis and the individual artists. All Rights Reserved. You may make verbatim copies of this document for non-commercial purposes by any means, provided that this notice appears on all such copies. Commercial use may be permitted under some circumstances. Contact the copyright holder for details.
Stage Works	None
TAC: Theory and Applications of Categories	
Ways Online	Copyright _ 1995, Issues for Canada's Future, Inc. All rights reserved. No portion of this publication may be reproduced in any form without written permission from Issues for Canada's Future, Inc.

- no charge for usage
e.g. publications on Wimsey server

6. Promotional model
- e-publication promotes print publication(s)
- electronic service includes publications catalogue and communication mechanism
- remote host - publisher
- no charge for usage
e.g. Halifax Daily News Worldwide

7. Dissemination model
- author/editor/publisher absorbs costs
- goal is maximum dissemination
- no host (e-mail distribution) to multiple hosts
- no charge for replication, access or usage

The Basic Electronic Journal (Model 3) has existed for over a decade on commercial services such as InfoGlobe, InfoMart and Dialog. The Dissemination Model predominated on the Internet and in the early titles identified for the EPPP. The Advertising Model is predominant in the popular titles included in the EPPP and is becoming common for technical titles. By its selection criteria, the EPPP has excluded titles or sites using the promotional model but these are growing daily among Canadian newspapers and trade publishers. The models described for scholarly commercial publishing have yet to appear outside of pilot projects such as Project Tulip, Red Sage and the ISI Electronic Library Project. The commercially secure servers and bandwidth necessary to support these models on the Internet are just emerging.

The attitude of publishers to enforce copyright on an e-publication is closely tied to the economic model in use for the e-publication.

All authors and publishers wish to protect the integrity of their works and often explicitly request that the work not be changed if reproduced and that the source be acknowledged. No author or publisher wants someone else to use his work for profit without permission.

Some publishers carry the Dissemination Model to its extreme, encouraging replication and redistribution without permission. The copyright statement used by Treasury Board in a draft version of *The Internet; A Guide to Internet Use in the Federal Government* takes this approach.[27]

27

Other publishers using the Dissemination or Advertising Models state restrictions on reprinting, republication or reproducing in their copyright statements. At the same time, these publishers are disseminating their publications without restriction on the Internet where, it is acknowledged, the user must reproduce text even to read or browse it. For the most part, the copyright statements still reflect the print world. Those most serious about controlling reproduction, publishers with unique and creative images or software, are most explicit in their copyright statements e.g., “This copyright includes, but is not limited to: graphic artwork, html coding, xhtml coding, the classified ad software, and all other custom software written to enhance the production or look of Euphony Net-Magazine.”

Publishers using the Subscription and Document Delivery Models will likely be very explicit and restrictive in their copyright statements and will back these statements up with licenses that further control use.

All these models and more will likely co-exist. Libraries must develop mechanisms to deal with all the models.

---

source work and the Treasury Board Secretariat be included.” The Internet; A Guide to Internet Use in the Federal Government. [Ottawa: Treasury Board Secretariat] April 18, 1995.

Documents concernant l'Internet
Laval library publications

University of Calgary
Electronic Library
    Electronic journals and newsletters
        alphabetic list of 50 titles, most held locally, some links to CICNet
    Electronic documents
        4 titles and links to Alex, Notre Dame (which links to Runeberg, ERIS, Gutenberg and OBI),
    Electronic Newstand
    Directory of Scholarly E-Conferences

University of Manitoba
Electronic materials 1 title and links to Alex, CICNet, OBI, Gutenberg, Electronic Newstand

University of Saskatchewan
EJournals -- most stored locally as the result of recommendations of a pilot project in 1993

Education	4 titles
Humanities and Social Sciences	5 titles
Library and Information Science	5 titles
Science and Technology	2 titles

    Link to CICNet
    Search by keyword (locally held titles)
    Bibliography and Project report
Electronic Books
    Business empty directory
    Humanities - a Department of English initiative with articles, books, journals and reports some stored locally
    Social Sciences 1 title
Government documents links to Statistics Canada, UN

University of Toronto. Centre for Computing in the Humanities
goal: prototype of an electronic research library
Academic resources by discipline
    45 subjects
    links to Georgetown, OBI, Oxford Text Archive

University of Toronto. Faculty of Information Studies
Electronic Resources Project led by Marte Misiek
Library services and electronic collections
    Electronic journals 19 titles held locally

Electronic texts 18 titles - some local, some links
Link to other electronic collections just citation guides

University of Toronto Libraries
Publications: Electronic books, journals, archives
6 titles, U of T Newsletters (2 titles), U of T computer Science Technical Reports archive
Links to Electronic Books at U. Minnesota, CICNet, Electronic Newstand, Global Electronic Library (LC), OBI, std.com
Canadian Government Publications links to NL, Industry Canada, Natural Resources, StatsCan
UN link

University of Waterloo
Finding information
Resources - general
Resources - by discipline

World Wide Web sites

University of British Columbia URL: http://unixg.ubc.ca:7001/
Electronic materials
2 serial titles, 1 monograph
links to Canadian electronic publications (NLC), CICNet, Electronic Newstand, Gutenberg, OBI, Alex, University Press catalogues (UChicago)

Technical services personnel register new publications in the EPPP registration database and subscribe to new publications.

The registration database contains information about each publication managed through the EPPP system.

Data elements stored in the EPPP registration database include:

- Display form of e-publication title,
- Physical name of directory used to store each issue
- Broad Dewey classification code,
- E-mail address of the e-publisher,
- E-mail message subject-line string used to match a message with an e-publication,
- Any other matching string that can be used to match an e-mail message with an e-publication.

One of the early decisions in the EPPP was to support viewing of e-publications from NLC Gopher and WWW servers.

Automated processes were developed to construct the Gopher menus and the WWW pages as e-publications are received. These menus and pages contain the links necessary to browse and view the e-publications stored in the EPPP archive area.

Gopher menu entries (links) and WWW pages (containing hypertext links) are constructed on the fly employing EPPP registration database information.

The EPPP registrar function is employed to add or modify the registration database content. The registrar interacts with the user in order to add new entries into the database.