

Collection Development in the E-content World: Challenges of Procurement, Access and Preservation

N.V. Sathyanarayana

Informatics (India) Ltd., Bangalore-560 004
sathya@informindia.co.in

ABSTRACT

Collection Development is one of the most demanding professional functions of a library that involves deeper understanding of the users' needs, organisational eco-system in which a library functions, and the supply-chain of information and publishing industry. The importance of this function to library's usage value continues to be more demanding in the e-content world. However, the multitude of choices offered and controls exercised by the content creators and providers (authors, publishers, aggregators, and distributors) and the laws that regulate the ownership, pricing, and usage rights of e-content makes the collection development process one of the most challenging functions in the libraries of digital world today. It is a very different process demanding the libraries to redefine their purpose and functions. This article is a broad sketch of emerging challenges of e-collection development and the changes it will bring in other key functions of the libraries. The scope of discussion is limited to journals and books in the digital format.

Keywords: Collection development, e-content, access, preservation

1. INTRODUCTION

In the print world of educational and research content, the collection development function revolved around the library's philosophy of ownership and preservation for perpetual use. When the libraries purchased the content from the publishers, the lending and perpetual usage rights came by default under the first-sale right doctrine of the copyright laws. Purchase for permanent retention was the only business model available for print collections. However, publishers continued to own the intellectual property (IP) rights of the content often without even keeping a single copy of books that have gone out-of-print and the back volumes of journals. It is well known that when publishers decided to come out with digital archives of their old journals from volume 1, many of them had to fall back on the preserved collection of library's bound volumes! Evidently, the publishers did not directly control the method of access and preservation of print material after it was sold to the libraries.

In the e-content world, publishers have taken direct and strategic control of content, its IP ownership, and its delivery and usage through the entire legal life-cycle of the content until the copyright period expires. Technology will facilitate them to maintain just one copy of the content, and a few mirror copies if need be, to serve the needs

of all the global users through internet. What the libraries buy will be right-of-access to this copy for limited or unlimited usage, for a limited or unlimited period, but by a limited number of users defined as authorised users of the library (members of the library). The process of obtaining and maintaining the right-of-access, which is known as licensing, is gradually replacing the traditional procurement model of content buying.

2. REDEFINING PURPOSE OF COLLECTION DEVELOPMENT

With the publishers taking direct control of managing the access and preservation, the selection of content that adequately serve the needs and purpose of library's user community at affordable costs today becomes more important than its needs in a distant future. The new mission of collection development may have to be—just-in-time access to the relevant and required content accessible by the user from anywhere. This imperative mission will force the following changes in the collection development practice:

- Collection development will become a short-term, need based plan. It will move from purchase for permanent retention to licensing for a definite or indefinite (perpetual) period of access and usage.

- Investment focus will shift from content ownership for preservation to the provision of technology infrastructure that will enhance the access, delivery and usage value of the licensed content.
- Libraries will have to engage the publishers to develop content licensing models that will be more affordable and serve the needs of the user community for content at a global level.
- A new kind of usage awareness program will have to be created for the user community to understand and appreciate the new content licensing models, associated copyright compliance requirements, and the price value of the content in addition to its usage value.

3. BUSINESS MODELS

Many business models are evolving in various permutations and combinations around the new licensing process for e-content procurement. These models may control the access by content volume, duration of access, size of the user population and the computing devices used for access. They generally fall into a variant of one or many of the following general models.

3.1 Perpetual Access Rights

This model is similar to libraries buying books and journals in the print-model. The library can license the content that grants access rights by its user community for an unlimited period. Technically this means that the library has the access rights for its (organisational) life. The questions that arise around this model are:

- Will the library have local archiving rights for the content? If so, how and when will the publisher deliver the content and at what cost?
- How does the publisher or the service provider guarantee the continuation of access in the event of closure of his organisation?
- Will this model provide the ownership privilege under the doctrine of first sale right that the libraries have under the copyright laws? Or, will it reduce ownership to the state of virtual reality forever?
- Is this right transferable in the event of the library or its parent organisation merging into another entity?
- How is this right shared in the event of a split? For example, if the University is split into two groups or campuses, or the library is split into two independent blocks in the same campus.

3.2 Limited Period Access Rights

This model is similar to renting or leasing the access rights for a limited period which could vary

from one month to one year or more. Consumer models are available, directed to end-users where a user can license access rights on a monthly payment/usage model which of course will not be appropriate for libraries. Normally, the libraries sign the license agreement for a period from one to three years. By general understanding, library will cease to have right of access to the content once the lease period expires. Following are some issues and questions that need to be addressed under this model:

- This model is popularly used by aggregators who license the content from the publishers. The publishers generally grant limited period of license to the aggregators are constrained by the business compulsions of limited options in dealing with licensing period and pricing terms. While delivering the content to libraries aggregators need to have a common licensing and pricing model for the content of all the publishers that their aggregation platform provides. For this reason, aggregators are not able to guarantee perpetual access rights to the content they deliver. There are a few exceptions in aggregation models which have tried to offer perpetual access rights. They are difficult to manage, both by the publishers and the libraries. JSTOR, EBSCO, PROQUEST, etc. are the examples of this model.
- Many publishers offer this model as an alternative. Some publishers offer this model only as conditional and refuse to make available content for perpetual access.
- Libraries develop collections for meeting the educational and research needs of their users. While licensing the content under this model may restrict the library's access by the period of lease, what happens to the usage rights of a user who has downloaded a portion of the content for his study or research that may be lost beyond the library's lease period? User has downloaded the content, like he takes a photocopy of an article from a journal under the fair-use doctrine which permits him to use the article for the entire duration of his academic or research needs. Does it amount to infringement if the end-user retains the downloaded copy even after the library's subscription has expired and not renewed?

3.3 Subscription Model

This model is perceived to be similar to libraries subscribing to print journals with many variants in the right of access. It is the most confusing model due to too many varied practices by publishers in setting the terms of licensing. It is a hybrid model of both perpetual access rights and limited access

rights. Following practices and issues concerned with this model need the attention of libraries for engaging with publishers to fine-tune this model for reducing the ambiguity:

- In the initial days of e-journals, publishers offered free access to e-version, limited to the current year of subscription as part of the print subscriptions. E-access was for limited content (current year's content only) and for a limited period (one year). This was a business strategy to promote e-journal usage. Most publishers have dropped this model but some still continue to offer.
- As a standard term of this model, when libraries subscribe to e-version of the journal either, independently or as part of the print subscriptions with additional payment for e-version access, libraries get access to e-version of the content for the current year (subscription year) with archival access to a few years of back-files depending upon the period for which the publisher maintains the back-files on his site. The current year's (subscription year) content will be made available for perpetual access, but the back-years' content will not be available for perpetual access. This means:
 - if the publisher has his collection available for 15-years;
 - if the library has subscribed to e-version for 5-years, with or without print subscription;
 - if the library stops renewing after 5-years;
 - the access will be perpetual for the 5-years of paid subscription content, and will be stopped for the rest of the 10-years content.
- Some publishers provide back file access as part of current subscription for rolling period of 3-5 years. Rolling period is one of the most confusing and least practical model.
- Of late publishers are optionally applying subscription model for e-books also. They are applying this model largely for selling their complete collection or subject collections which will include their back-lists. Besides finding new market for their books, they have brought in a fundamental change by digitizing their back-lists. The concept of out-of-print or out-of-stock will no longer be applicable for books. Subscription model for books has raised a serious issue of long-term affordability of this mode to libraries.

3.4 User Population

Within the business models discussed above, publishers exercise different levels of access control as a means of inventing new business models and to manage price-flexibility in product pricing.

In the print world the content had inherent access disadvantage. A copy could be available for one user at a time and the libraries had to buy multiple copies where demands were more. A fundamental change that the e-content has brought in is the feasibility of 24/7 access to all the content that a library licenses to all its users, and simultaneously. This will never happen in reality and in any situation only a small part of the library's content will be under use by a small section of the users. Taking advantage of this fundamental change, publishers have replaced the multiple copy purchase concept by concurrent access model. It is like saying three concurrent users are equivalent to licensing three copies. This model is particularly relevant for text books. The model has opened up two distinct pricing options:

- Pricing based on number of users in the campus or full-time-equivalent (FTE) model with no limit on the number of concurrent users.
- Pricing based on number of concurrent (simultaneous) users accessing the content irrespective of the user population of the library.

3.5 Pay-per-view Model

This model is equivalent to paid document delivery service. Publishers never offered this model in the print world. Digital content has made this model a new, though very limited, revenue opportunity for the publishers and a strategy to protect the stable revenue from the subscription model. This model has all the potential to bloom as a cost-affordable option to both libraries and end-users. But, neither the libraries due to practical problems nor the publishers due to revenue considerations, are willing to give a serious chance to this model. One of the reputed e-Journal aggregators admitted to this author in a recent conversation in 2011 that his system is recording around 100 million turned-away users. Even charging \$1 per article would have brought him a revenue three times more than his current subscription based revenue. Turned-away users refer to users landing for full-text in a publisher portal from discovery services like J-Gate or Google or any abstracting and indexing service that provides links to full-text.

3.6 Subscription-based Article Delivery Model

Libraries rarely buy all journals due to affordability constrains even from speciality publishers in their subject area. Users approach to published content is rarely by journal but more by their topics of interest and study needs. Every publisher knows that subscription model limits the access to users for articles from his other non-subscribed journals by the library, but has rarely ventured into develop article delivery model as a parallel and equally

remunerative business model comparable to subscription model. Of late, as a means of making the article delivery (Pay-per-view) model affordable, a few publishers have started bundling article delivery through subscription model. For example: a library serving the needs of users in the chemical research group subscribes to five journals from the Royal Society of Chemistry (RSC) which publishes about 30 journals. Its users have definite and regular need for articles from the rest of the RSC journals but the library can't afford to subscribe them all. RSC's subscription-based offer for a pre-defined bundle of articles in the range of 100-1000 from the non-subscribed journals can meet library's the needs in as affordable model.

4. RIGHTS & LICENSES

When a library buys a book or a journal, it acquires the right to make it available for unlimited and unrestricted use by its user community and retain the acquired copy forever. In effect, it acquires unlimited lending rights to its users. This right is called the first-sale-right which includes the right to protect the copy for ever or re-sell that copy to anybody. The implicit right in protecting the physical copy for ever will also extend the right to take a backup copy for use in the event of the book suffering mutilation. The only limitation imposed on the library by the copyright law is that the total restriction from copying for any purpose other than limited copying of a small portion (not substantial portion) of the book required by an authorised user of the library for his/her own individual educational or research and study purpose. This concession is given under the fair-use doctrine.

4.1 Managing the Rights and Licenses

When the library moves to e-content what it acquires is the right-of-access to the content wherever the publisher is hosting the digital copy. Since the library doesn't get a physical copy and the medium limits the right to access only, limited or perpetual, it opens up a new world of controlled or managed access. Due to the fear of the fact that the copyright laws do not adequately cover the situations and implications of content acquisition and access in the digital medium, publishers have started the practice of licensing the copyrighted content with additional terms to protect their IP rights and economic interest. With every purchase of e-content signing a license agreement has become a normal practice now. The digital collection of content that libraries will be developing in future and their ownership patterns will be represented by two elements:

(a) Rights acquired for access based on the business model of the publisher chosen by the library

(b) A Licensing document describing terms of contract under which the rights are acquired.

These two elements will control and manage the access all through the perpetual contract. Hence, all the systems, functions and services of the library such as the library automation system used by the library, circulation of e-content, access through discovery services like J-Gate or library's OPAC or a federated search systems, renewal tracking, local archiving, etc., will have to be redefined to work through controls of these rights and licenses.

5. PERPETUAL ACCESS VS. TRUE OWNERSHIP

The print world has created a ownership myth for the libraries which is a result of the social responsibility that the libraries assumed as social memories of the civilisation. It was a limited ownership of the physical copy that the library acquired through purchase. The real owner remained the IP rights owner (author represented by his publisher). The physical ownership guaranteed perpetual access to the content by virtue of the provisions of first sale right doctrine of the copyright law as mentioned earlier in this paper. With that came the perpetual maintenance responsibility to the library and the associated costs.

When the libraries purchases a digital copy of a book, the privilege of perpetual access should be applicable to this purchase also through this doctrine of first sale right. But, there is a serious technology bottleneck for exercising this right at the usage level, particularly in the function of lending the digital book. Copyrighted works donot comfortably fit within the constraints of the first sale doctrine¹. Copyright laws completely restrict substantial copying of the copyrighted material. When the library loans a printed book to its patrons, it involves no copying. But, when it lends a copy of the digital book, the lending process creates additional digital copies at the delivery and usage point. The library retains the source copy unless deleted/deactivated temporarily from its digital content repository through some special technology and restores only after the user returns the digital copy. There must be a control at the user's end to permanently delete the copy from the user's machine once the loan period is expired. While the digital rights management (DRM) technologies exist to manage this process, it is expensive and often proves to be an irritation to the user.

The perpetual access is highly desirable to libraries from the century-old mind set of retention for re-use. However, in the e-content world libraries have an opportunity to bring in a paradigm shift in the pricing models for digital content as service rentals. Some disruptive economic model like Google needs to emerge for the paid content to make this happen.

6. CONCLUSIONS

Nothing explains the implications of the changes the digital technology and the internet are driving, better than the recent uproar in the massive digital lending of e-books. When a library buys a print book, by the doctrine of first sale right it acquires unlimited lending privilege except that the nature of the medium puts a restriction on lending to one-user-at-a-time and only until the copy gets mutilated through circulation to a point of non-use. Normally, a book survives 20-50 circulations depending on the quality of paper and binding. This system of lending is undergoing fundamental change with e-books. Overdrive is a company that provides a platform for facilitating and managing lending of e-books as a technology service to the libraries. In just one year (2011), American libraries checked out loaned 350 lakhs digital titles, with 170 lakhs reservations pending²! Threatened by potential loss of new sales due to the massive upsurge in digital lending by libraries, several publishers have withdrawn selling e-books to libraries. In an open letter to librarians explaining its switch to limit the number of check-outs a library can offer on an e-book, Harper Collins said that its previous policy of "selling e-books to libraries in perpetuity, if left unchanged, would undermine the emerging e-book eco-system, hurt the growing e-book channel, place the additional pressure on physical bookstores, and in the end

lead to a decrease in book sales and royalties paid to authors"³. When a library now buys (licenses), an e-book with perpetual access rights, technically it can use one copy (one license) to lend the book to *any-number-of-its-members-at-a-time* with no fear of loss or mutilation. Extend this concept to the public library system of a country. The implications could be disastrous to the publishing system of the world. Obviously, the e-content access and delivery system needs to invent a new model for library buying.

Libraries will be increasingly buying the e-content through new form of licensing for a fee. But, what will not change is, that library's users will continue to access the content free unless the libraries also shift to commercial eco-system of charging for usage.

REFERENCES

1. First-sale-doctrine. Wikipedia. http://en.wikipedia.org/wiki/first-sale_doctrine.
2. e-book discovery and sampling skyrocketing at public libraries. <http://overdrive.com/news/e-book-discovery-and-sampling-skyrocketing-at-public-libraries>.
3. Open letter to librarians. Harper Collins, library love fest blog. 3 March 2011. http://harperlibrary.typepad.com/my_weblog/2011/03/open-letter-to-librarians.html.