E-Commerce and Digital Libraries

Suliman Al-Hawamdeh(*) & Schubert Foo(**)
Division of Information Studies
School of Information and Communication, Nanyang Technological University
Nanyang Link, Singapore 637718
Email: assuliman@ntu.edu.sg(*) assfoo@ntu.edu.sg(**)

Abstract

Until recently, digital libraries have provided free access to either limited resources owned by the organization or information available in the public domain. For digital libraries to provide access to copyrighted material, an access control and charging mechanism need to be put in place. Electronic commerce provides digital libraries with the mechanism to provide access to copyrighted material in a way that will protect the interest of both the copyright owner and the digital library. In fact, many organizations such as Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE) have already started to make their collections available online. The subscription model seems to be the favourable option at this point of time. However, for many ad hoc users, the subscription model can be expensive and not an option. In order to cater to a wider range of users, digital libraries need to go beyond the subscription models and explore other possibilities, such as the use of micro payments that appear to be an alternative logical solution. But, before even that can happen, digital libraries will need to foremost address a number of outstanding issues among which include those of access control, content management, information organization, and so on. This chapter discusses these issues and challenges confronting digital libraries in their adoption of e-commerce including e-commerce charging models.

Introduction

Digital Library Research Initiatives in the United States and the increased interested in digital libraries by computer science researchers had provided the impetus for the growing proliferation of digital libraries around the world. Most existing digital libraries have mainly focused on digitizing individual collections and making them available on the Web for users to search, access and use. They are providing a new means of fast and effective access to information in different forms and formats. Nonetheless, the development of digital libraries also translates into significant financial requirements, which in the past, has been borne largely by government funding agencies, academic institutions and other non-profit organizations.
By the virtue of the basic principles of economics and business, digital libraries are looking for alternative forms of revenue generation in order to meet the ever-increasing needs of users through the provision of new value-added services and products. In this respect, e-commerce can provide digital libraries with the means to support their operation and provide them with a sustainable source of funding. This is natural evolution in the use of digital libraries as content management and electronic publishing are gaining momentum and popularity.

However, before digital libraries can engage in e-commerce activities, many issues need to be addressed. Some of these issues include intellectual property, access control, backup and archiving, and micro payments. In this chapter, we will look at these issues and highlight problems and opportunities related to digital libraries as a viable e-commerce business model.

**Characteristics of Digital Libraries**

The “digital library” is a term that implies the use of digital technologies by libraries and information resource centers to acquire, store, conserve and provide access to information. But with the increased interest in other areas such as electronic commerce and knowledge management, the concept of digital library has gone beyond the digitization of library collection. It has been expanded to encompass the whole impact of digital and networking technologies on libraries and the wider information field. Researchers from many fields including computer science, engineering, library and information science are investigating not only the digitization of catalogues and collections or the effective use of networked resources but also the meaning of these developments for both information providers and users alike. Beside the technical issues that engineers are dealing with, there are a number of issues such as acquisition, content management, charging and intellectual property that require the help of the business and the legal experts to deal with.

As digital libraries are being embraced by many communities, the definitions and characteristics of digital libraries varies from one community to another. To the engineering and computer science community, digital library is a metaphor for the new kinds of distributed database services that manages unstructured multimedia. It is a digital working environment that integrates various resources and makes them available to the users. From the business community perspective, digital library presents a new opportunity and a new marketplace for the world’s information resources and services. From the library and information science perspective, it has been seen as “the logical extensions and augmentations of physical libraries in the electronic information society. Extensions amplify existing resources and services and augmentations enable new kinds of human solving and expression” Marchionini (1999).

According to the Digital Library Federation (DLF), digital libraries are “organizations that provide the resources, including the specialized staff, to select, structure, offer intellectual access to, interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.” (Digital Library Federation, 2001). From the above, it is clear that the stakeholders of digital libraries are many and wide-ranging. They include publishers, individual authors and creators, librarians, commercial information providers, federal, state and local governments, schools, colleges, universities and research centers, corporate technology providers, major information user organizations in both the public and private sectors. With this, it is unsurprising to find a myriad of different definitions and interpretations of a digital library. It could be a service, an architecture, information resources, databases, text, numbers, graphics, sound, video and a set of tools and
capabilities to locate, retrieve and utilize the information resources available. It is a coordinated collection of services, which is based on collections of materials, some of which may not be directly under the control of the organization providing a service in which they play a role. However, this should not be confused with virtual libraries or resource gateways that merely provides a link to the external resources without any extra effort to manage those resources. As those resources are normally not under the control of the organization, maintaining content and keeping the links up to date is extremely difficult.

But while the definition of the digital library is still evolving, it might be easier to look at the characteristic and functionality provided by the digital library. Garrett (1993) outlined some of these characteristics that are worth noting:

- **Ubiquity.** At least some set of services must be accessible at any time from any physical location.

- **Transparency.** The internal functioning of infrastructure components and interactions must be invisible to users. Users must be able to access services using their user interface of choice.

- **Robustness and scalability.** The infrastructure must be powerful enough to withstand a wide range of potential risks and continue to function without disruption to users and service providers.

- **Security and confidentiality.** The infrastructure must include mechanisms which ensure that parties to any transaction can reliably identified to each other, that confidentiality of the parties and the transaction can be assured where appropriate, and that the system cannot be easily compromised.

- **Billing, payment, and contracting.** The infrastructure must support both financial transactions in payment for goods and services, and the delivery and utilization of electronically generated and managed tokens (e.g., digital cash).

- **Searching and Discovery.** The infrastructure must provide for a wide range of resource identification strategies, from highly specific searches to generic browsing.

Clearly, the above characteristics involve access to information, content management, search and retrieval of information, payments, security and confidentiality, technology and infrastructure. While some of these issues sound manageable, other issues such payments and intellectual property still pose significant challenges and are still candidates for further research and development. The following sections address some of these issues confronting digital library development, and in particular, those affecting the electronic commerce aspect of the digital library.

**Issues Confronting Digital Libraries**

**Content Management**

Content management is an important and critical activity in digital libraries. It involves the creation, storage and subsequent retrieval and dissemination of information or metadata. In this respect, content management can be closely linked to online search services. While most of the collections in digital libraries are still text-based, this is expected to change in future as more and more material will be made available in multimedia format. As the content is
expected to come from various sources, it will also come in different formats such as word processor files, spreadsheet files, PDF files, CAD/CAM files and so on. However, Rowley (1998) pointed out that despite the growing importance of multimedia approaches, most of the collections are still text based. The volume of text base information is increasing at an alarming rate and its diversity of form, from the relatively unstructured memos, letters or journal articles, to the more formally structured reports, directories or books, is continually broadening. The management of content will also involves capturing and validating information. Nonetheless, issues related to ownership and intellectual property will continue to hamper the development of digital libraries. Most of the digital libraries that exist today are either own the content or just provide a link to the information resource. Access control and intellectual property are therefore fundamental issues in the operation of large digital libraries.

**Issues facing the content organization in digital format**

Information organization is an area that still evolving and will continue to do so for sometime. Statistical based information storage retrieval models have failed to provide an effective approach to the organization of large amount of digital information. On the other hand, more effective tools, which have been used manually by the librarians to organize information in the traditional libraries, are considered slow, tedious and very expensive. Given the vast amount of information available today, it is important to organize it in a way that allows for modification in the retrieval system. This is highlighted by Arms et. al. (1997) where flexible organization of information is one of the key design challenges in any digital library. The purpose of the information architecture is to represent the riches and variety of library information, using the building blocks of the digital library system. With the different types of material in a digital library, information can be organized using a hybrid approach that combines the statistical based techniques with manual organization tools. Many companies are developing tools that will enable libraries to create taxonomies and organize information in a more meaningful and useful way.

The growth in size and heterogeneity represents one set of challenges for designers of search and retrieval tools. The ability of these tools to cope with the exponential increase of information will impact directly on the content management of the digital systems. Another challenge pertains to that of the searcher’s behaviour. Recent studies have shown that users have difficulty in finding the resources they are seeking. Using log file analysis, Catledge and Pitkow (1995) found that users typically did not know the location of the documents they sought and used various heuristic techniques to navigate the Internet, with the use of hyperlinks being the most popular method. They also found that users rarely cross more than two layers in a hypertext structure before returning to their entry point. This shows the importance of information organization and content management in digital libraries.

The organization of information is still an issue in content management that needs to be addressed. Some outstanding issues include the following:

- The nature of digital materials and the relationship between different components. A digitized document may consist of pages, folders, index, graphics or illustration in the form of multimedia information. A computer program, for example, is assembled from many files, both source and binary, with complex rules of inclusion. Materials belong to collections can be a collections in the traditional, custodial sense or may be a compound document with components maintained and physically located in
different places. Although it appears to the user as one entity but in reality it can be put together as a collection of links or an executable components.

- Digital collections can be stored in several formats that required different tools to interpret and display. Sometimes, these formats are standard and it is possible to convert from one format to another. At other times, the different formats contain proprietary information that requires special tools for display and conversion, thereby creating content management and maintenance problems.

- Since digital information is easy to manipulate, different versions can be created at any time. Versions can differ by one single bit resulting in duplicate information. Also digital information can exist in different levels of resolution. For example, a scanned photograph may have a high-resolution archival version, a medium quality version, and a thumbnail. In many cases, this is required if we want to address the retrieval and display issues on one hand, and printing quality issues on the other hand.

- Each element of digital information may have different access rights associated with it. This is essential if digital libraries are used in an environment were information needed to be filtered according to confidentiality or to be sold at different price.

- The manner in which the user wishes to access material may depend upon the characteristics of computer systems and networks, and the size of the material. For example, a user connected to the digital library over a high-speed network may have a different pattern of work from the same user when using a dial-up line. Thus, taking into account the response time and the speed by which information can be delivered to the users becomes another factor of consideration.

It is clear from the above that the organization of information should take into consideration many issues. Borgman (1997) noted that the issues of interoperability, portability, and data exchange related to multi-lingual character sets have received little attention except in Europe. Supporting searching and display in multiple languages is an increasingly important issue for all digital libraries accessible on the Internet. Even if a digital library contains materials in only one language, the content needs to be searchable and displayable on computers in countries speaking other languages. Data needs to be exchanged between digital libraries, whether in a single language or in multiple languages. Data exchanges may be large batch updates or interactive hyperlinks. In any of these cases, character sets must be represented in a consistent manner if exchanges are to succeed.

Information retrieval in a multimedia environment is normally more complex. Most of the information systems available today including digital libraries still rely on keywords and database attributes for the retrieval of images and sound. No matter how good is the image descriptions used for indexing, a lot of information in the image will still not be accessible. Croft (1995) noted that general solutions to multi-media indexing are very difficult, and those that do exist tend to be of limited utility. The most progress is being made in well-defined applications in a single medium, such as searching for music or for photographs of faces.

**Copyright and Intellectual Property**

Digital libraries as any other Web applications are still not protected from copying, downloading and reuse. Digital technology makes reproduction of electronic documents easy and inexpensive. A copy of an original electronic document is also original, making it difficult to preserve the original document or treated different from the other copies. In a
central depository system where the original document is normally stored, the digital library system will have to make copies of this document for viewing or editing purposes whenever users access the document. In the Web environment, a copy is normally downloaded to the users machines and sometimes cached into the temporary directory for subsequent access.

The ease in which copies can be made and distributed prompted many to predict that the electronic publishing will not prevail, as there might not be many people willing to put their works on the Web due to lack of protection. As legislators grapple with the issues of copyright, electronic document delivery is already taking place both within and outside the restrictions of copyright. The sentiments expressed by Oppenheim (1992) reflect those of many with regard to copyright in that “the information world is essentially a global one ... and the legal framework in which the industry operates is in places very confused, and in some cases, such as data protection, it is unwittingly swept up by legislation not aimed at it all. In other areas such as liability and confidentiality of searches, it will face increasing pressures from its consumers in the coming years”.

Although the copyright issues in many areas have not been fully addressed, attempts have been made recently to introduce greater restrictions upon copyright and intellectual property. One such notable effort is by the Clinton Administration’s Intellectual Property Working Group, which issued its Copyright Amendment recommendation code named “Green Paper”. The Green Paper recommends amending the copyright law to guard against unauthorized digital transmission of copyrighted materials ((Mohideen, 1996).. The four main principal implications of the law include

- Copyright should proscribe the authorized copying of these works
- Copyright should in no way inhibit the rightful use of these works
- Copyright should not block the development of dissemination of these works
- Copyright should not grant anyone more economic power than is necessary to achieve the incentives to create

Based on these principles, the US Copyright Commission concluded that making some changes to the Copyright Act of 1976 could develop protection of computer programs. Congress has accepted the recommendations.

The question of Intellectual Property versus the Freedom of Information has been widely debated. There are two opposing views to this issue. One is that creators of information should be amply rewarded for their works. On the other hand, there is the notion that nobody really owns information and society would be better off if knowledge is available for all. In the old system, copyrights always protected the physical entities by prohibiting the reproduction of the work without permission from the author. This also includes photocopying with the exception of fair use for educational purpose. In the Internet environment, downloading and printing is not much different from photocopying, although controlling this activity is extremely difficult.

In the past, copyright and patent laws were developed to compensate the Inventors for their creations. The systems of both law and practice were based on physical expression. In the absence of successful new models for non-physical transaction, how can we create reliable payment for mental works? In cyberspace, with no clear national and local boundaries to
contain the scene of a crime and determine the method of prosecution, there are no clear
cultural agreements on what a crime might be (Barlow, 1995).

**Intellectual Property Management**

For digital libraries to succeed, an intellectual property system needs to be developed to
manage copyrighted material and ensure that the rights of authors and creators are protected.
Garett (1993) proposed having an Intellectual Property Management System to manage
Intellectual Property in a distributed networked environment. This system should assure
copyright owners that users would not be allowed to create derivative works without
permission or to disseminate the information beyond what is permitted. Besides controlling
the copying of information, owners and users also would like to ensure that information has
not been intercepted or altered in anyway. To be able to achieve this, Garett suggested that
the Intellectual Property Management System must be capable of the following:

- Provide for confidential, automated rights and royalty exchange;
- Ensure owners and users that information is protected from unauthorized, accidental
  or intentional misattribution, alteration, or misuse;
- Ensure rapid, seamless, efficient linking of requests to authorizations;
- Include efficient and secure billing and accounting mechanisms.

Another method of protecting Intellectual Property and Copyright as proposed by
Marchionini (1999) is through using technical solutions. The solutions are in the form of
encryption algorithms and digital watermarking. So far techniques have been developed
whereby visible or hidden watermarks on digital objects have been incorporated into
commercial products. According to Marchionini, these techniques insure the veracity of an
object and may discourage the copying and distribution in the open market place. Examples
of such systems currently being tested include Cybercash, DigiCash & Netbill. Cybercash use
a third party intermediary to effect transfer of property and payment while DigiCash issues
money in the form of bit stream tokens that are exchanged for Intellectual Property. Netbill
use prefunded accounts to enable intellectual property transfer.

**Cataloguing and Indexing**

The exponential growth of the Web has made available vast amount of information on a huge
range of topics. But the technology and the methods of accessing this information have not
advanced sufficiently to deal with the influx of information. There is a growing awareness
and consensus that the information on the Web is very poorly organized, and of variable
quality and stability so that it is difficult to conceptualize, browse, search, filter, or reference
(Levy, 1995). Traditionally, librarians have made use of established information organization
tools such as the Anglo American Cataloging Rules (AACR2) to organize, index and catalog
library resources. This works fine with the printed material by providing access to the
bibliographic information only. When it comes to content indexing on the Web, these tools
are inadequate and expensive to use due to the large amount of information available on the
Web. The other major problem with the traditional approach is the fact that it is a largely
intellectual manual process and that the costs can be prohibitive in the Web environment.
This is further exacerbated that information on the Web is prone to sudden and instant
updates and changes. An automated indexing process is therefore more useful and suitable.
The success of automatic indexing should therefore lead to fast access and lower costs. The
other major difference between traditional libraries and digital libraries is the content and format of the information stored. Digital libraries contain multimedia information, images, graphics and other objects where traditional cataloging rules do not deal with.

Currently, indexing and retrieval of images is carried out using textual description or database attributes assigned to the image at the time of indexing. Indexing and retrieval based on image content is still very much in the research stage. In the Web environment, metadata is used to provide description of an object for indexing purposes. Metadata is a data about data, which is highly structured like its MARC (MAchine Readable Catalogue) counterpart in order for retrieval software to understand exactly how to treat each descriptive element in order to limit a search to a particular field.

Some of the digital libraries such as the State Library of Victoria Multimedia Catalogue attempted to use the MARC format to catalog digital objects only to find that it did not work adequately. In some cases, it becomes very complex requiring highly trained staff and specialized input systems. Digital librarians have identified three categories of metadata information about digital resources: descriptive (or intellectual), structural and administrative. Of these categories, MARC only works well with intellectual metadata. Descriptive metadata includes the creator of the resource, its title, and appropriate subject headings. Structural metadata describes how the item is structured. In a book, pages follow one another, but as a digital object, if each page is scanned as an image, metadata must “bind” hundreds of separate image files together into a logical whole and provide ways to navigate the digital document. Administrative metadata could include information on how the digital file was produced and its ownership. Unlike MARC, which is a standard specified by AACR2, metadata standards are still evolving and still there is no consensus on a particular standard to follow (Tennant, 1997).

The other main concern with cataloging and indexing is the hefty costs involved. Basically, the cost to assign values to index attributes depends on the amount of work that is needed to determine what information to post. If the index is prepared before scanning, such as filling out a form, then adding index records to the database is strictly a data entry effort. However, if the information is derived from a reading or the document or an analysis of photographs, it will be very costly indeed. According to a report prepared for the Washington State Library Council (1999), a fifteen-element index record with 500 characters of entry may take between 30 seconds and a few minutes to complete. For thousands or hundred thousands of items, this translates into very high costs.

Access Control

Access to most digital libraries was initially made free to promote the site and attract users. Materials available on these sites are limited due to the lack of an appropriate and good access control system. When digital libraries deal with copyrighted material or private information, they are faced with the necessary task of developing access control facilities. A good example is the course reserve system developed by many universities to manage courseware. Most course reserve systems provide different levels of access control depending on the type of material and the enrollment of the students. Another reason for having a flexible and good access control system is the need for cross-organizational access management for Web-based resources. This is another area of great interest to information consuming institutions and information resource providers. These organizations would like to enable access to a particular networked resource or to a particular member of an institutional consumer community. While access to users should be easy and flexible, it should also
protect the privacy of the user and should not depend entirely on the user's location or network address but rather on the user's membership in appropriate communities. It should also provide the necessary management and demographic information to institutional consumer administrators and to resource providers.

A flexible and good access management system should do more than providing the technical infrastructure. It should also address a number of other difficult issues such as access policies and deployment of technology. Two important technical infrastructure components are required for an institutional access management system. First is the ability of a user to obtain an identity on the network, known as authentication, and the second is the ability to correlate a user's identity with rights and permissions to use various services, called authorization.

Given the problem surrounding the development of a good access control in digital libraries, there are a number of issues need to be taken in consideration when developing and deploying an access control infrastructure:

- The system must address real-world situation. It should take into consideration the technology being used to verify users as well as the level of user expertise. In the Internet and e-commerce environment, verification of users is rather difficult and a Public Key Infrastructure (PKI) might be needed to address the security and trust problems.
- The system should protect users privacy and protect users information from illegal or inappropriate use.
- It should provide different level of access to information depending on the type and nature of that information. Some services might be made accessible to the public while other can be restricted to paid users, managers or head of divisions.
- Access to information should not be hampered by technology and made difficult as a result of security or access right measures. It should remain efficient and simple.
- It should be easy to control and manage. Web based user registration, verification reduces the time, and cost involve in administering the system. It should be as painless to manage and to scale as current technology permits.

For libraries to engage in e-commerce activities, they need to deploy an access control system not only to protect information resources but to also enable them to charge and collect money. Thus, access control in digital libraries will need to be integrated with payment and intellectual property management.

**E-Commerce in Libraries**

Libraries have so far been very slow in embracing electronic commerce. This is largely due to that fact that most libraries are originally institutionalized as non-profit organizations. Furthermore, the cost of setting up an e-commerce infrastructure was a barrier as libraries are generally not cash-rich organizations. However, electronic commerce and Internet have played a significant role in the way libraries operate and the way library services have developed. Many libraries have made their presence felt on the Web by making their collections searchable and their services accessible. The web sites of the New York Public Library (NYPL), the British Library, and Singapore National Library Board (NLB) are good
examples of libraries using current technology to enhance and extend their services to current and future clientele.

Whether in digital or traditional environment, libraries were set to provide various mechanisms for knowledge archival, preservation, and maintenance of culture, knowledge sharing, information retrieval, education and social interaction. Barker (1994) states that as an educational knowledge transfer system, a library fulfils a number of important requirements, these being:

- The library is a meeting place – a place where people can interact and exchange ideas.
- The library provides a range of resources that access to them otherwise is difficult.
- The library provides an effective mechanism for information acquisition and dissemination.
- The library provides access to expert in different field and help users to locate relevant information.
- The library is an educational institution and plays an important educational role for the fulfillment of life-long learning.

In keeping up with the changes and advances in technology, and the need to creating self-sustaining entities, some libraries are changing their practice and adapting to the new environment by starting to charge their users for certain classes of value-added services such as document delivery, reference services and information research. The Canadian Institute for Scientific and Technical Information (CISTI) is an example of such a library or resource center that charges the public for value-added services (Song, 1999). In Singapore, the Library 2000 report recommended that basic library services remain free, however value-added services such as translating, analyzing, and repackaging information will be chargeable (Fong, 1997). Currently, the National Library Board (NLB) of Singapore has adopted and implemented cashless payments through the use of the cash-cards. The use of cash-cards at NLB branches for all transactions was introduced in 1998 in an effort to automate payment processing. Although the introduction of cash-card systems at NLB branches initially drew some negative responses, the majority of library users soon grew accustomed to this mode of payment.

The cash-card system developed by NETS (Network for Electronic Transfers (S) Pte Ltd) and KRDL (Kent Ridge Digital Laboratories) of Singapore enabled the cash-card to be conveniently used at NLB branches. C-ONE, Singapore’s first attempt at developing an electronic commerce system to enable cash-card payments over the Internet was introduced at some NLB libraries in 1999. The cash-card, which is basically a stored-value card, is useful for micro-payments. The value of the card can be topped-up at machines through the use of bankcards. However, the main drawback of the cash card and NETS is that they are only usable in Singapore.

As another example, the Library of Virginia introduced electronic commerce by enabling its patrons to adopt a book or shop on-line from its gift shop via its Web site that is credit card enabled (Harris, 2000). In more noticeable emerging trends, some libraries have begun to develop partnerships with vendors such as booksellers. The Tacoma Public Library is one such library where it allows its patrons to order books from the online booksellers,
Amazon.com, via its OPAC (online public access catalogue) system. For each transaction, it earns 15% commission of the sale (Fialkoff, 1998).

Digital libraries are being developed for the preservation and access of heritage material through digitization efforts. At the same time, the digitized documents are potential revenue generators for these digital libraries. In addition, the digital library is an avenue through which electronic publications and value-added services can be accessed. With the presence of NetLibrary, many options are available to libraries (physical and digital) to offer electronic books for access to their members. NetLibrary goes through the process of acquiring the distribution rights to books from publishers and has made approximately 14,000 books available for access. Some of these books can be accessed for free while others require payment (Breeding, 2000). Electronic commerce and digital libraries are complementary in that “a digital library may require the transactional aspects of EC to manage the purchasing and distribution of its content, while a digital library can be used as a resource in electronic commerce to manage products, services and consumers” (Adam & Yesha, 1996).

The platform for libraries to innovate within its designated roles is reaching new heights with the aid of technology and electronic commerce. Traditional methods of doing things can be performed more effectively through an electronic exposure. The World Wide Web has created new avenues of delivering traditional services and created an environment of creative business development within the realms of the library world.

**Charging Models for Digital Libraries**

Since the definition of a digital library is still evolving, there is no prevailing e-commerce model for digital libraries. However, most of the goods sold on digital libraries are content such as electronic journals and databases. But there is no reason to exclude the fact that digital libraries cannot still sell physical goods such as postcards, books, T-shirts, mugs and other forms of goods. Given that, digital libraries might have to adopt different charging models. These charging models need to be integrated into a seamless and convenient interface. Some of the charging models that can be used for digital libraries include the pre-paid subscription model, pay later subscription model, and pay now or as you use model.

**Pre-Paid Subscription Model**

In this model, the buyer maintains a virtual account with the seller that has been debited with a certain amount that is normally the annual or monthly subscription value. Depending on the terms and conditions of the subscription arrangement, the account can then be used for subsequent payments during payment transaction. This provides a very convenient form of payment where user need not submit payment details each time to effect a transaction.

**Pay later Subscription Model**

This is similar to the pre-paid subscription model with the exception of “use first and pay later”. This model also requires users (buyers) to register with the service provided. A virtual account is created and maintained. At the end of the subscription period, a bill is sent to the user for payment. Payment methods using credit cards fall in this category. In the case of E-commerce, credit card is a very common and convenient payment type and will be expected to be available on the digital library by many users. Most of the e-commerce websites operating under this model require the user credit card number upon registration. The credit card will be charged at the end of the subscription period.
Pay Now or As You Use Model

This model requires the user to pay immediately after completing the transaction. For example, if the payment method selected by the buyer is the cash card, then the amount of the transaction will be immediately deducted from the cash card and deposited into the electronic wallet of the seller. In the context of digital libraries, this type of payment mode is appropriate where payment must be collected before a user can access a service such as downloading of documents, articles, software, etc. The problem with this model is that some of the payment techniques used such as cash cards and NETS are not standard and users cannot use them from abroad. Since credit card payments are not suitable for micro-payments and small amounts, there is a need for a standardized charging mechanism such as cybercash or digital cash. Currently none of these charging techniques are accepted as a standard.

In fact, the issue of micro-payment is not restricted to libraries but to many sprouting new businesses on the Web.

According to the micro-payment model of Amazon.com, a micro-payment constitutes an amount of US$2 or less. It was cited that these payments are difficult for many sites to collect because the cost of processing such small transactions are higher than the fee itself (Regan, 2001). This issue greatly impacts the libraries, as they are generally non-profit making, so that they are effectively incurring more costs in providing such a service to allow users to make micro-payments online.

Given that, how can consumers pay for such micro-payments at their convenience and yet does not incur a high overhead for the online merchant? In respect of this, many various commercial suppliers such as CyberCoin of CyberCash.com, MicroPayment by IBM and Millicent by Compaq, have developed products in response to addressing the need for micro-payment transaction and processing.

These suppliers attempt to connect buyers and sellers on the Internet into a worldwide microcommerce environment. For the example of Millicent (Compaq Computer Corporation, 2001), buyers will first need to open an account with them or through authorised brokers, and fund it using one of three ways, namely, through an online credit card or debit card transaction, by direct billing their monthly ISP (Internet Service Provider) statement or telephone bill, or through pre-paid cards purchased anonymously through convenience stores. Funds are held in the account in any currency until needed and then spent at vendor websites with the simple click of a mouse. The system takes care of the actual payment, handles currency conversion if required, resolves content delivery problems, and automatically processes refund requests. On the part of the seller, they first open vendor accounts by first selecting a licensed payment hosting provider from the MilliCent website. This might be their ISP, a preferred Commerce Hosting Provider (CSP), or MilliCent itself. Once the account is opened, the vendor is live as part of the network. Using their browser, vendors assign prices to individual objects to be sold or to groups of objects. As a final step, vendors generally update their website HTML pages to visually reflect pricing to website visitors. Alternatively, vendors are allowed to integrate the software directly into their websites. By directly controlling their own payment processing, advanced vendors can increase their website responsiveness for end-users, better integrate MilliCent into their day-to-day operations, and eliminate any fees charged by MilliCent payment hosting providers.
Micro-payment is also an issue facing digital libraries as a direct result of intellectual property management. Digital libraries will not be able to provide services for copyrighted material unless a charging mechanism that takes into account intellectual property is put in place. Monash University implemented a course reserve system as well as a formula to calculate charges for copyright materials in an electronic course reserve system (Hawamdeh, 1999). The formula takes into account the number of pages scanned, viewed or printed, the retail price of the printed material or monographs and the number of users enrolled in the course. This could form a basis for licensing and use of copyrighted material in universities and large corporations. Implementing such a formula can help the library to cater for copyright clearance and manage charges when necessary.

XML Might Hold the Key

Digital libraries have started recently to take advantage of XML to better organize and manage digital resources. XML or Extensible Markup Language came about as a result of combining SGML (Standard Generalized Markup Language) and the Web. Due to the limitations inherited in HTML (HyperText Markup Language), there is a need to extend HTML capabilities to better display and manipulate Web pages. SGML, on the other hand, is powerful but too complicated. XML achieves much of the power of SGML but without the complexity associated with the use and implementation of SGML.

XML promises to solve many of the problems associated with diverse data types by allowing for user-defined markup rather than browser-defined markup. It provides markup that describes the content similar to that of SGML and goes beyond merely describing the format. This description of content has implications for extracting and reusing the content in ways that allows micro-payments and charge-per-use mechanisms. For example, a book marked in XML can then be displayed and manipulated in various ways. It can be displayed chapter-by-chapter, sections or even paragraphs. XML-encoded material can be viewed both on the Web and personal devices such as e-book readers and personal digital assistants. The Open eBook Forum is working on a standard method of encoding e-books in XML specifically to provide an easy method for interchanging books across reading devices (www.openebook.org).

The efforts to provide an open XML-based infrastructure enabling the global use of electronic business information in an interoperable and secure environment are underway. ebXML, sponsored by UN/CEFACT and OASIS, is a modular suite of specifications that enables enterprises of any size and in any geographical location to conduct business over the Internet (http://www.ebxml.org/geninfo.htm). By taking advantage of these efforts, digital libraries will be in better position to implement electronic commerce and micro-payment systems.

Conclusion

In this chapter we addressed some of the issues and challenges facing digital libraries in their adoption of e-commerce. Some of these issues can be resolved using technology such as access control, cataloging and indexing. Others, such as content management, quality of information, copyright and intellectual property goes beyond technical issues and extends to policies and regulation. To make e-commerce a viable option for digital libraries, the issue relating to micro-payments and charging must first be resolved. Such an issue is not restricted to digital libraries but also to many businesses on the Web that deals with content such as music and gaming. It is obviously infeasible to collect small amounts when the cost of processing such small transactions is higher than the fee to be collected itself. As
technology advances, digital libraries will continue to look for new ways in which content and other services can be delivered and a fee correspondingly collected. This could be achieved through using new charging mechanisms such as cybercash and across-country NETS mechanism.

References


Authors Biography

**Dr. Suliman Hawamdeh** is an Associate Professor and Programme Director of the Master of Science in Knowledge Management programme, School of Communication and Information at Nanyang Technological University, Singapore. He has more than 20 years of teaching and industrial experience in areas such as knowledge management, electronic commerce, document imaging, information retrieval, Internet and digital library. He holds a Master’s degree from University of Michigan, Ann Arbor in the US and a Ph.D from University of Sheffield in UK. He is the founder and president of Information and Knowledge Management Society (*iKMS*). He is also the author of a book on “Information and Knowledge Society” published by McGraw-Hill.

**Dr Schubert Foo** is the Head of Division of Information Studies and Programme Director of the Master of Science in Information Studies programme, School of Communication and Information at Nanyang Technological University (NTU), Singapore. He received his B.Sc. (Hons) in Mechanical Engineering, Ph.D. in Materials Engineering, and M.B.A. from the University of Strathclyde in 1982, 1985 and 1989 respectively. He joined NTU in 1990 and over the years, lectured in the Divisions of Computer Technology, Software Systems and Information Studies. His research interests include Internet and multimedia technologies, information retrieval and digital libraries. He has published over 100 international journals and conference papers to date in these areas.