

# Enhancing Services in a Digital Age – 10 Years of Experience from the Systems Librarians' Perspective

Edward F. Spodick and Ki-Tat Lam

The Hong Kong University of Science and Technology Library,  
Clear Water Bay, Kowloon, Hong Kong  
{lbspodic, lblkt}@ust.hk

**Abstract.** This paper is an attempt by the authors to share their experiences in equipping a young academic library with the information technologies needed to enhance services in a digital environment. After discussing the advent of *digital libraries*, the paper explores a progression of projects which make use of advancing technologies, from Web interfaces to XML metadata, and their effectiveness in a non-English (CJK) environment. These digital initiatives have become a core component of the Hong Kong University of Science and Technology (HKUST) Library's service infrastructure, in addition to enhancing its traditional roles. The past ten years' accelerating pace of technological change has had a tremendous impact on the provision of library services. Through this paper, the authors have provided one institution's experiences both in benefiting from and contributing to these changes.

## Introduction

The Hong Kong University of Science and Technology is a very young institution, having its first student intake in 1991. The early years of the University coincided with several streams of radical change and optimism – from an incredibly accelerated expansion of information technology to the economic and political progression of Asia, including a peaceful and mostly welcomed change of sovereignty for Hong Kong. Locally, the University was established and funded to serve as a leader in research and development, and was positioned as the leading edge of an expansion in tertiary education for Hong Kong. In a few short years the University has become ranked in the top tier in Hong Kong, and in many fields it holds such a place globally.

All of this was a major factor in the Library's ability to innovate and to embrace new techniques and technologies during what has been perhaps the most tumultuous upheaval in librarianship and the provision of library services in the history of our profession. From the Library's founding, it embraced new technologies in an effort to enhance information access and services. In 1991, the Library rolled out one of the first fully Chinese-capable multilingual integrated library systems in the world and the first large-scale campus-wide CD-ROM network in Asian academic libraries; 1992 saw the installation of the largest networked full text database system in Asia, followed in 1993 by an early Course Reserve Image System. And 1995 saw the first academic library Web server in Hong Kong. In subsequent years the Library estab-

lished a mirror site for some subscription databases, a wide-scale implementation of native XML-based database development projects, and a host of other innovations [1].

Upheaval, change and challenge have been our constant companions, and we believe they have been a significant factor driving the Library and University's achievements.

## From Traditional to Digital

The past few decades have seen a remarkable alteration in the technology available for information processing. Coupled with the ability to obtain ever more powerful equipment and more capable software for an ever-decreasing investment, this has allowed more smaller or less well funded libraries to participate in the exploration and integration of technology for the improvement of operations, collections and services.

Initially the advances promised since the development of a standard for MACHine Readable Catalog records were available only to the largest institutions. Even the smallest school library has the option to do far more than merely replace their card catalog with an electronic version. And thus the concept of a "Digital Library" began to be more than just a theoretical exercise – it became a topic of increasingly common discussion and experimentation. This was especially relevant to a small library in a brand new university – one that had the option of "going digital" from its creation.

While there may continue to be discussion of the differences between definitions used for "digital", "electronic" and "virtual" libraries, the authors agree with Roy Tennant when he states that while the term "electronic library" may be more inclusive it is "digital library" which has become the accepted terminology [2].

There are three basic ways of defining what is meant by the term "digital library":

1. Libraries providing a) access to digital information using a variety of networks, including the Internet, and b) services in an automated environment
2. The result of projects to digitize library materials for network access
3. A discipline which refers to research on the theories and technologies for the building of digital libraries

The first definition most fits the ongoing efforts of the authors and their colleagues. As part of these efforts, the second definition is representative of attempts to provide access and improve services in relation to materials that exist in the print collection.

The third definition comes into play as part of the ongoing efforts of the authors and others to learn how to incorporate new technologies and methodologies into the services and collections, as well as in their work to enhance collaborative efforts and to contribute new ideas to the profession as a whole.

One of the most important points the authors wish to stress is that the term "digital library" simply reflects a fundamental shift in **how** libraries do what they do, rather than in **what** they do. Since their initial establishment, libraries have provided access to information and services to assist their patrons in obtaining this access and utilizing the information acquired through it. None of this is new. Improvements in information technology have allowed a much greater degree of flexibility in all stages of these processes. But the basic mission of a library to support its user community through enhanced access and services remains the same.

## **Digital Libraries in Practice**

The HKUST Library began testing the Lynx and NCSA Mosaic Web browsers, for connection to Gopher, WAIS and Web sites, in 1994. In view of its potential capability for becoming the library's gateway to information, serving not just plain-text but multi-media documents, approval was obtained to set up a Library Web Server. Officially released in May 1995, it was the first such server in Hong Kong and the region.

Once the web infrastructure and technology were in place, a progression of digital library projects began to take place. The following sections discuss a selected number of issues and lessons learned during the implementation of these projects.

### **Content Digitization and Web-Based Document Management System**

In 1998, after more than four years of experience in digitizing course reserve materials in TIFF format for access from dedicated workstations, the authors began to base future digitization projects on Adobe's PDF document format with a Web interface. Although PDF is now a popular format for content delivery, it was not an easy decision seven years ago. PDF was selected because it was more portable and smaller than TIFF, with a free viewer which seamlessly launched as a web browser plug-in. The documents were digitized into PDF and hyperlinked to the bibliographic metadata. Text was extracted using Acrobat Capture and some Chinese OCR software.

Three digital library databases were created during that time, serving up HKUST Theses, the Digital University Archives and HKUST-related News Clippings.

It was also not an easy task at that time to identify a web-based document management system that could handle full-text searching and display of text in both English and Chinese. With much localization, the databases were created on a commercial document management system called BASIS, which was later replaced by native XML database systems, although the basic design concepts remained unchanged.

### **Interchange of XML Metadata with the Library Catalog**

The HKUST Library uses INNOPAC, an integrated library system, to host the Library Catalog and to provide automated library functions such as circulation, cataloging and acquisitions. The Library Catalog contains bibliographic metadata for all print, multimedia and electronic resources.

When building digital library databases, the authors frequently must export metadata from the Library Catalog to the database for indexing and displaying purposes. A pioneering attempt at HKUST to deploy XML technologies to extract the metadata from the Catalog in real-time began in 2000, using the INNOPAC's built-in *xrecord* command and XSL style sheets. These technologies were then adopted for most of HKUST's digital library projects, including an Electronic Journals database system rolled out in 2001.

XML is a very flexible format – being able to extract data in XML is critical to a Digital Library project's success. If your catalog or database does not have this capability yet, push the vendor hard or switch to something else.

### **CJK and Unicode**

The HKUST Library has been working very actively with the INNOPAC vendor to enhance support for CJK (Chinese, Japanese, Korean) characters in the Library Catalog, which ensured that CJK bibliographic metadata can be seamlessly stored, indexed, searched, sorted and displayed.

With advances in computer hardware and operating system software, the past decade saw many critical CJK issues resolved [3]. However, Web technology, together with its increasing adoption of Unicode, has brought another layer of problems.

The CJK-using library communities have traditionally used a legacy character set known as EACC (East Asian Character Code). It is used in MARC 21, an international bibliographic metadata content description standard, for cataloging CJK materials. Both MARC 21 and EACC are maintained by the Library of Congress (LC) and are widely adopted by bibliographic metadata suppliers and systems vendors. The cross-walk problem of mapping between EACC and Unicode began to emerge when suppliers and vendors began to develop Web- and Unicode-based applications. A failure in mapping usually results in various types of display and searching problems.

In early 2003, Library staff conducted an analysis of the EACC-Unicode mapping of CJK characters in the HKUST Catalog and concluded that Hong Kong users of the INNOPAC system should work more closely with the vendor to clean up the mappings. A Hong Kong Innovative Users Group (HKIUG) working group made up of members from the HKUST, City University of Hong Kong, Chinese University of Hong Kong and University of Hong Kong libraries was then established to develop an HKIUG version of the EACC-Unicode mapping table. The table was subsequently adopted by the vendor for implementation at its customer sites, with the first implementation being at HKUST in late 2003.

### **Global Name Access Control**

In 2002, the HKUST Library began the development of an XML Name Access Control Repository [4] to address the problems experienced by catalogers and catalog users in identifying personal authors whose names are in non-Latin scripts [5].

In cataloging, the use of names in bibliographic records is controlled by authority records, and the process involved is called Authority Control. The main purpose of name authority control is to authorize a form of a name (known as the established form) to be used in bibliographic records as an access point. The concept of Access Control as proposed by Barnhart [6] expands on the role of an authority record by enriching into a central place for looking up various forms of a name, allowing the linking of variant forms without declaring any of them to be the authorized form.

It was concluded from this research project that distributed repositories of metadata for describing various forms of an author's name, with encoding in original scripts and their Latin forms, can be built globally by various bibliographic utilities, national libraries and library consortia to form a virtual name access control platform. To achieve this, there must be a mutually agreed-upon metadata schema for data interchange between repositories. The XML Name Access Control Repository is an initiative to promote the establishment of such a standard.

## Open Source Software and Open Access

Thanks to the open source software movement the authors have been extensively using such tools for library applications. Two of our projects were based on the open source software *DSpace*. *EPrints* and *DSpace* were then the only options, and *EPrints* did not support Unicode, making it useless when CJK characters are required.

The HKUST Institutional Repository was established in early 2003 [7] to create a permanent record of the University's research and scholarly output. It was the first such initiative in Asia. By allowing the content of the Repository to be globally and openly accessible, we are contributing to the international open access effort that campaigns for free and unrestricted online access to research literature.

The Digital University Archives on DSpace was released in late 2004. In contrast to open access, the Archives requires authorization. The DSpace source code was modified to search and display only authorized documents based on LDAP authentication, so different user groups would access different document sets. This kind of localization would not have been feasible if DSpace were not open source.

## Conclusion

These are exciting times to be a Librarian. The expansions in information technology discussed in this paper have opened up many new vistas of information and service provision. The possibilities continue to be enormous, limited more by the inability to find time to explore them than by fiscal constraints. The authors urge all Librarians to set aside some time on a periodic basis to simply **think** about new options and technologies they see in their readings and explorations, and **ponder** how to relate them to solving identified local needs for the traditional library roles of providing access to needed information and services both to support that access and to improve users' ability to utilize the information obtained.

## References

1. Library's First Ten Years. (<http://library.ust.hk/info/exhibit/sep2001/>)
2. Digital v. Electronic v. Virtual Libraries. (<http://sunsite.berkeley.edu/mydefinitions.html>)
3. K.T. Lam. Chinese Information Access and Retrieval: Issues Facing Libraries. Presentation at the Seminar on Chinese Information Processing in Libraries, January 1998, The HKUST Library. (<http://hdl.handle.net/1783.1/1924>)
4. XML Name Access Control Repository. (<http://library.ust.hk/info/nac/>)
5. K.T. Lam. XML and global name access control, OCLC Systems & Services, vol. 18, no. 2, 2002, p.88-96. (<http://hdl.handle.net/1783.1/443>)
6. Barnhart, L. Access control records: prospects and challenges. Authority Control in the 21st Century: An Invitational Conference, March 31-April 1, 1996. (<http://www.oclc.org/oclc/man/authconf/>)
7. K.T. Lam. DSpace in action: implementing the HKUST Institutional Repository system. International Conference on Developing Digital Institutional Repositories : Experiences and Challenges, December 9-10, 2004, Hong Kong. (<http://hdl.handle.net/1783.1/2023>)