大學圖書館
10 卷 1 期（民國 95 年 3 月） 頁 102-122

From Idea to Reality: Building the HKUST Institutional Repository

Catherine S. Y. Kwok
Reference Librarian, Hong Kong University of Science and Technology Library
E-mail：lbckwok@ust.hk

Diana L. H. Chan
Head, Reference Department, Hong Kong University of Science and Technology Library
E-mail：lbdianac@ust.hk

Gabrielle K. W.
Reference Librarian, Hong Kong University of Science and Technology Library
E-mail：lbgabi@ust.hk

【Abstract】

The Hong Kong University of Science and Technology Library builds the HKUST Institutional Repository, aiming to preserve and disseminate its institutional scholarly output. This paper describes the background which led to its birth, the growth and development of the project, the challenges involved and strategies adopted to achieve sustainability.

Keywords：Institutional repositories; Content recruitment; Self-archiving rights; Open access; Hong Kong

臺灣大學學術期刊資料庫

The Hong Kong University of Science and Technology Institutional Repository (HKUST IR) was created by the HKUST Library as a response to the evolving scholarly communication landscape. This paper describes the background that led to the birth of the HKUST IR and charts its development since 2003. It also discusses the challenges involved, and the strategies adopted to achieve sustainable growth.

I. Changing Landscape of Scholarly Communication

Institutional repositories emerge out o f a changing scholarly communication l andscape which has been largely reshaped in recent years by a number of convergent factors: technological breakthroughs in information exchange, social and cultural trends, the open access movement, and the policies of research grant funding bodies.

a. Technological trends

The birth of the World Wide Web in the 1990's revolutionized ways for information exchange. Researchers increasingly posted their research results on the Web for immediate and wider dissemination. Subject-specific, web-based open archives became a popular medium for researchers to share preprints and working papers. In 1991, arXiv, the first preprint server on Physics and related disciplines, was developed. Other collections began to sprout: CiteSeer, the Cognitive Sciences Eprint Archive, SSRN, IDEAS, the NASA Technical Reports Server, and the National Academy Press. The success of these collections brought about the Open Archives Initiative (OAI), which promotes author's self-archiving and interoperable standards for file sharing. A major outcome of OAI is the Open Archives Initiative Protocol for MetaData Harvesting (OAI-PMH). This protocol allows the harvesting of metadata from different archives so that an indexing or a distribution service can be provided.

b. Social and cultural trends

While the academia was experimenting with new modes of sharing research results, academic libraries had to endure the serials crisis which had plagued them since the 1980's. According to a study done by the U.S. Association of Research Libraries, the unit cost of journals had shot up 188% in the past two decades while library budgets continued to shrink. The inflation rate for materials acquisitions was an alarming 73% in 2004. (Take Figure 1: Monograph and Serial Expenditures in ARL Libraries, 1986-2004) An inevitable consequence was that academic libraries could no longer afford to

臺灣大學學術期刊資料庫

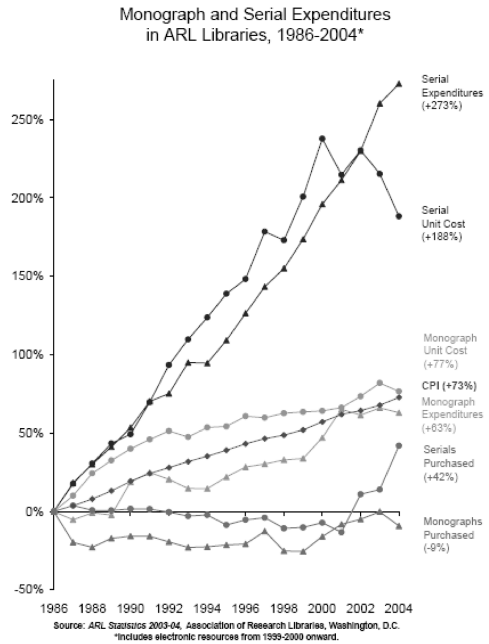


Figure 1

subscribe to journals their researchers and students need.

The success of arXiv and other open archives and the permeating Web generated a whole new culture of sharing scholarly output. Users had a growing expectation of finding quality information online, full-text and free of charge. These trends sparked off campaigns for free and unrestricted online access to research literature worldwide. One such campaign was the Budapest Open Access Initiative which arose from a small meeting convened by the Open Society Institute (OSI) in Budapest, December 2001. The purpose of the meeting was to accelerate international efforts to make research articles in all disciplines freely available on the Internet. Such campaigns developed into the open access movement.

c. Open access movement

The beginning of the open access movement can be traced back to the founding of the Scholarly Publishing and Academic Resources Coalition (SPARC) in 1998. SPARC was created to address the market dysfunctions in scholarly publishing. It offers a partial solution to the serials crisis by expanding competition and advocating open access to research publications. Sponsored by the Association of Research Libraries, its various partners have published 200 free or low cost journals in the past seven years.

Since then, the open access movement has borne many fruits. Numerous open access journals are published and made freely available on the Web. Many of them are refereed or peer-reviewed journals with high impact factors, such as *Emerging Infectious Diseases* and *Journal of Machine Learning Research.* The Directory of Open Access Journals (DOAJ) now covers 2,000 journals in 17 subject categories, providing access to a total of 85,000 articles. These journals can be read online at no charge.

臺灣大學學術期刊資料庫

Institutional repositories (IRs) are another result of the open access movement. They are built to capture, preserve and disseminate digitally the scholarly output of an institution on a free and interoperable basis. The development of IRs gained momentum with the release of two open source systems: Eprints by the University of Southampton in late 2000, and MIT's DSpace in 2002. Both systems now have numerous installations.

Developed at the University of Michigan, OAISter is another fruit which provides a single searchable interface for 600 open archives holding millions of documents. They include articles from open access journals, working papers, discussion papers, conference papers, theses and dissertations, and contents of institutional repositories around the world.

d. Funding bodies

Funding bodies also play a crucial role in propelling open access and the creation of institutional repositories. The U. K. Parliament Science and Technology Committee issued a report in July 2004, recommending that publicly-funded institutions should "establish institutional repositories on which their published output can be read free-of-charge online" (Albanese, 2004). The Wellcome Trust, the biggest non-governmental funding body of biomedical research in the U.K., embraces open access by announcing that "from October 2005, all papers of new research projects must be deposited in PubMed Central within six months of publication" (Wellcome Trust, 2005). Similarly, the U.S. National Institutes of Health proposed that NIH-funded research should be made freely accessible online six to twelve months after publication. Australia's National Scholarly Communications Forum also shares a similar vision (Douglas, 2004).

II. HKUST Institutional Repository: Some Facts and Figures

HKUST is a young university with just 14 years of history. With English as the medium of instruction, the university community consists of 440 full-time faculty members, 5,600 undergraduates, and 3,200 postgraduates. Despite its short history, in 2005, the university was ranked 43 among 200 top universities in the world (The Times Higher Education Supplement, 2005). Our researchers publish widely in international peer-reviewed journals.

The HKUST Library showed its early support to open access by joining SPARC in 2001. It has also been at the forefront of library digitization. The idea of establishing an institutional repository at HKUST grew from a library staff de-

臺灣大學學術期刊資料庫

velopment workshop in late 2002. Kimberly Douglas, the University Librarian of the California Institute of Technology, spoke at this workshop on "E-prints, OAI and Institutional Repository". A task force was then formed to investigate the issue. Like other universities, HKUST Library started with the basic premise that the scholarly output of HKUST researchers is an institutional intellectual asset, one that should be carefully guarded and preserved for posterity. Researchers, therefore, should be encouraged to exercise their self-archiving rights and make their scholarly work openly accessible so long as publishers' requirements are complied with.

In February 2003, the HKUST Institutional Repository (HKUST IR) <http://library.ust.hk/repository/> was launched with 105 computer science technical reports. After three years of development, it grows to over 2,200 documents, holding preprints, technical reports, working papers, conference papers, journal articles, presentations, book chapters, patents, and PhD theses. They are mainly PDF files with some Powerpoint and program files. (Take Figure 2: Total No. of Documents as of October 5, 2005) The documents are grouped into a hierarchy of communities and collections. Communities refer to the university's teaching departments and research centers while collections stand for the various document types. (Take Figure 3: Communities and Collections) Not surprisingly, disciplines which have an established tradition of sharing preprints and working papers constitute the bulk of the content. Almost 50% of our documents are authored by research staff in our Computer Science and the various engineering departments. (Take Figure 4: Contributors by Departments as of October 5, 2005)

Collection	Size	Percentage
Conference Papers	579	26%
Working Papers, Technical Reports, Research Reports, Pre-prints	534	25%
Journal Articles	493	23%
Doctoral Theses	394	18%
Patents	58	3%
Presentations	56	2%
Book Chapters	37	2%
Miscellaneous	8	1%
Total	2,159 (incl. 100 duplicates)	100%

Figure 2: Total Number of Documents as of October 5, 2005

臺灣大學學術期刊資料庫

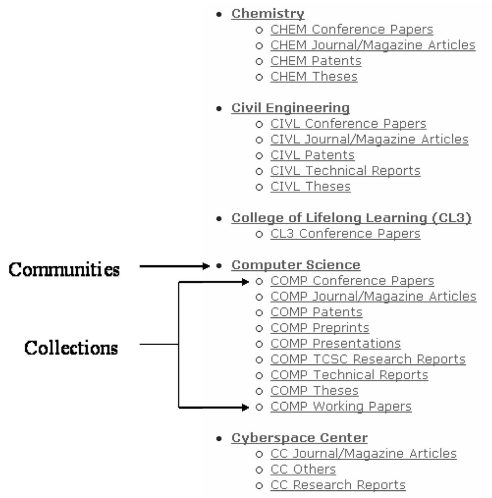


Figure 3: Communities and Collections

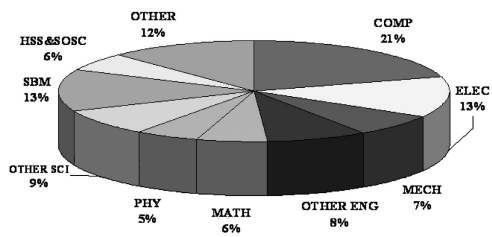


Figure 4: Contributors by Departments as of October 5, 2005

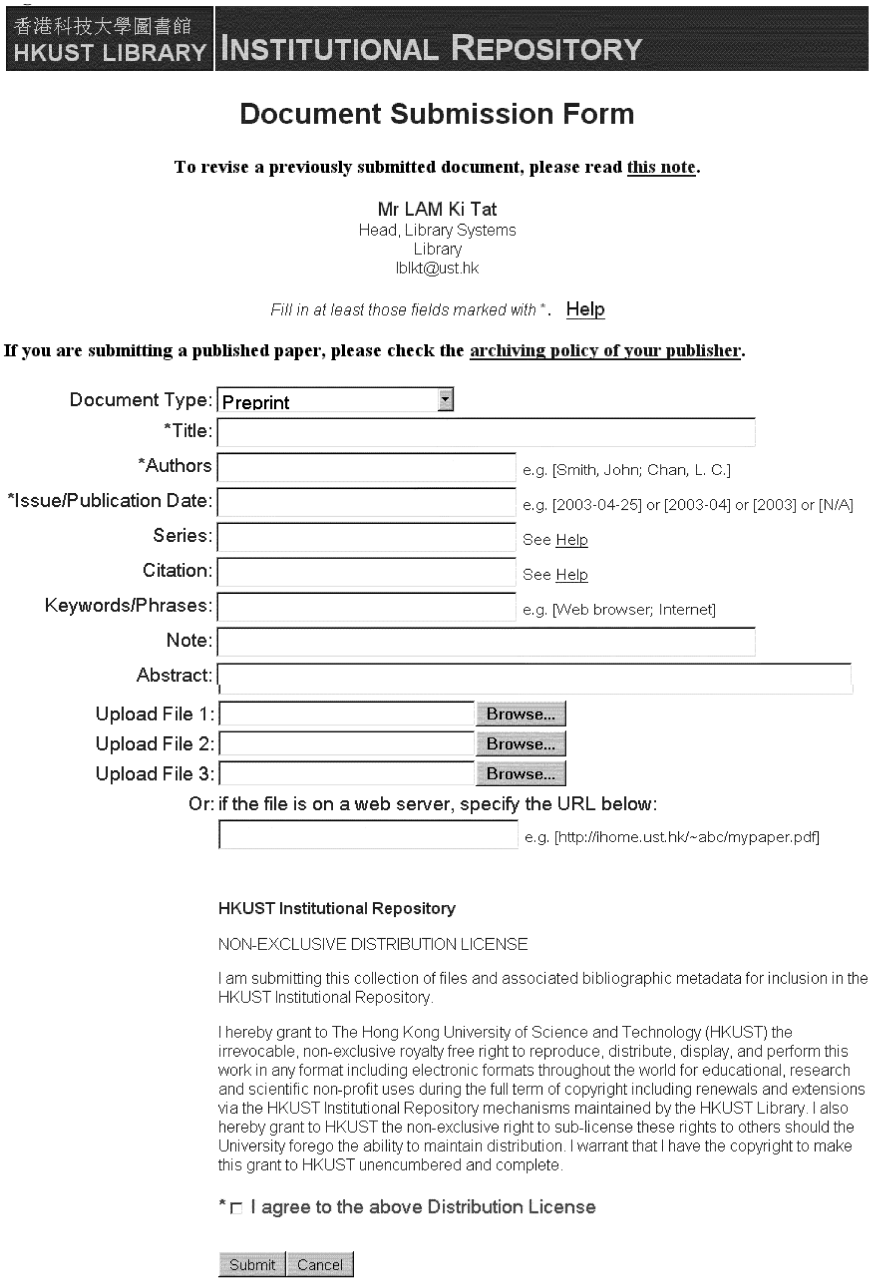
a. Systems issues

At the outset, it was decided that the HKUST IR should run on an open source software with full OAI-PMH compliance. Eprints and DSpace were two popular IR software programs in early 2003. Our Systems colleagues evaluated the two systems and finally chose DSpace mainly because EPrints did not fully support unicode and was not Java- and servlet-based at that time. That DSpace had the advantage of designing its software based on the Eprints experience was also a consideration.

Several customizations and enhancements were performed on DSpace. The most pressing problem was that though DSpace supported unicode, its lucene search engine was unable to search by CJK characters. This was solved by replacing DSpace's Tokenizer with a CJK Tokenizer. It was also discovered that the URLs of queries containing CJK characters were not properly encoded. Setting the Tomcat web servlet to URIEncoding="UTF-8" offered a practical solution. Display problems concerning subscript and superscript and number of items retrieved were also overcome.

Document submission is an important function in any IR and should be made simple and effortless. The original DSpace submission function was cumbersome, so our Systems colleagues replaced it with an intuitive submission program written in Perl. All a depositor needs to do is to fill out the bibliographic details of a document, attach the associated electronic file and click the submit button. (Take Figure 5: Document Submission Form)

臺灣大學學術期刊資料庫



香港科技大學圖書館 HKUST LIBRARY | INSTITUTIONAL REPOSITORY

Document Submission Form

To revise a previously submitted document, please read this note.

Mr LAM Ki Tat
Head, Library Systems
Library
lblkt@ust.hk

*Fill in at least those fields marked with *.* Help

If you are submitting a published paper, please check the archiving policy of your publisher.

Document Type: Preprint
*Title:
*Authors e.g. [Smith, John; Chan, L. C.]
*Issue/Publication Date: e.g. [2003-04-25] or [2003-04] or [2003] or [N/A]
Series: See Help
Citation: See Help
Keywords/Phrases: e.g. [Web browser; Internet]
Note:
Abstract:
Upload File 1: Browse...
Upload File 2: Browse...
Upload File 3: Browse...
Or: if the file is on a web server, specify the URL below:
e.g. [http://ihome.ust.hk/~abc/mypaper.pdf]

HKUST Institutional Repository

NON-EXCLUSIVE DISTRIBUTION LICENSE

I am submitting this collection of files and associated bibliographic metadata for inclusion in the HKUST Institutional Repository.

I hereby grant to The Hong Kong University of Science and Technology (HKUST) the irrevocable, non-exclusive royalty free right to reproduce, distribute, display, and perform this work in any format including electronic formats throughout the world for educational, research and scientific non-profit uses during the full term of copyright including renewals and extensions via the HKUST Institutional Repository mechanisms maintained by the HKUST Library. I also hereby grant to HKUST the non-exclusive right to sub-license these rights to others should the University forego the ability to maintain distribution. I warrant that I have the copyright to make this grant to HKUST unencumbered and complete.

* ☐ I agree to the above Distribution License

Submit Cancel

Figure 5: Document Submission Form

臺灣大學學術期刊資料庫

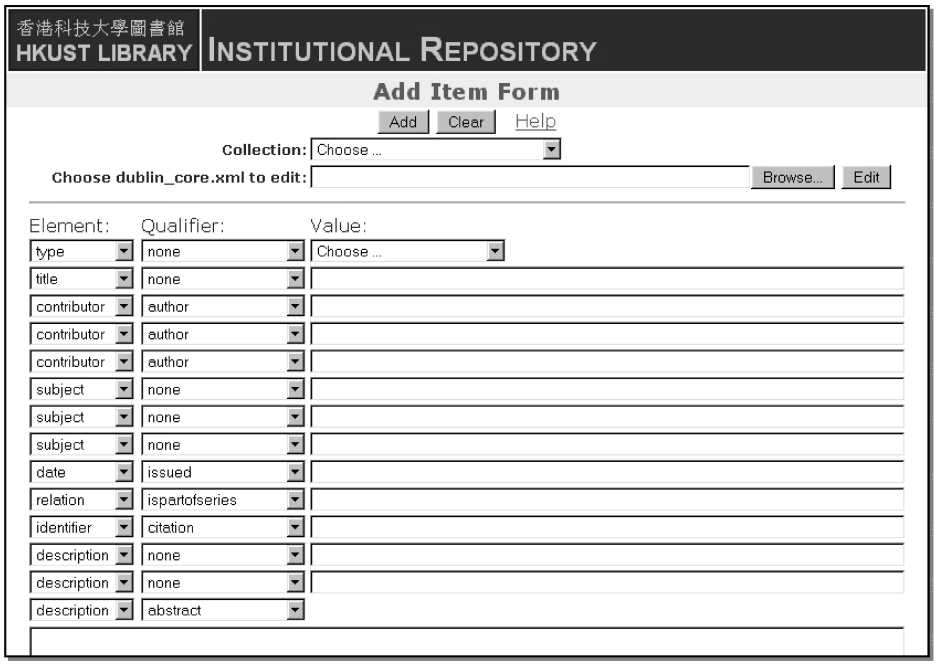


Figure 6: Add Item Form

An Add Item form using a JSP application was also locally developed to allow library staff to add records to the IR with ease. Library staff can either create a new item from scratch or load the metadata previously submitted by faculty onto the Add Item form and edit it before final submission. (Take Figure 6: Add Item Form)

The HKUST IR is an OAI data provider. Our documents are indexed and easily discovered in OAIster and Google Scholar. Our IR will soon be indexed as a preferred web source in Elsevier's Scirus, a search engine which specializes in scientific information. These OAI harvesters collect the metadata in Dublin Core format from various IRs around the world including HKUST's and provide value-added indexing and searching services.

b. Staffing issues

Building the HKUST IR has truly been a library-wide project. Colleagues from Systems, Collection Development, Archives, Acquisitions, Administration, Cataloging and Reference took part in it. The Library formed a team of eight reference and subject librarians and five data

臺灣大學學術期刊資料庫

	Copyrighted Journal articles, book chapters, conference proceedings, theses, presentations		**Non-copyrighted** Working papers, technical reports
Archivable	University Owned	Author's Permission	Author's permission Department's Permission
	Author Owned	Author's Permission	
	Publisher Owned	Publisher's and Author's Permission	
Non-archivable	Selected items to ask for author's & publisher's permission		

Figure 7: Copyright vs. Self-Archiving Rights

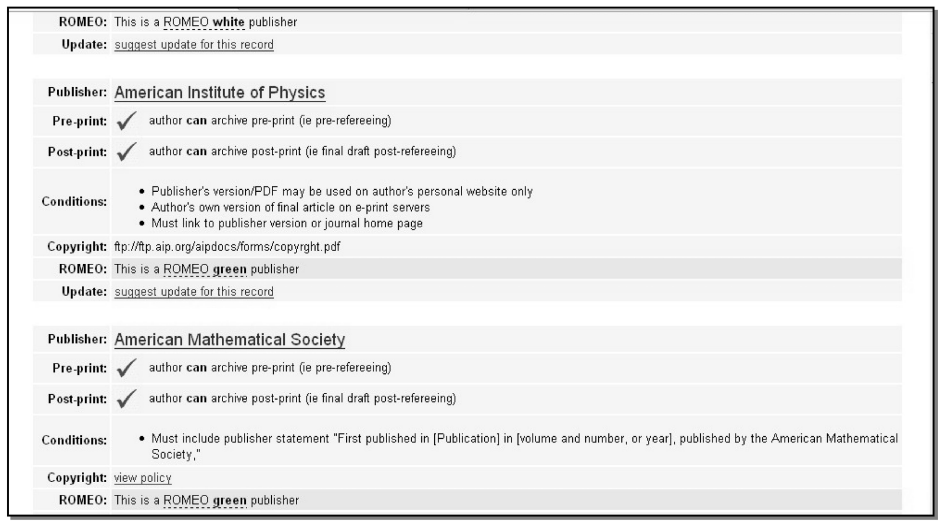
entry staff to take charge of content recruitment and the day-to-day operation of the IR. Policies, procedures, and liaison issues fell within the purview of the committee of reference librarians while the committee of department heads gave strategic guidance and ruled on critical issues.

III.Building HKUST IR

With the systems and staffing issues resolved, the work team was ready to act. The issues of open access, author's self-archiving and institutional repositories were unfamiliar concepts to most of our faculty members. Before the Library could blow the horn and campaign for support, we had to seed the IR with a critical mass of papers. Susan Gibbons (2004) comments that the "recruitment of content, not technology, is the greatest barrier to success". Ware (2004) surveyed 45 IRs and found out that the average number of documents per repository was only 1,256. Both point to a fact experienced by all libraries: content building is a painfully slow and uphill process.

The difficulties of adding content to the IR were manifold. To start with, the Library had to actively seek out documents that were potentially archivable. Voluntary submissions by faculty members were scarce and could not be relied on at this stage. Next, the Library had to ascertain the copyright ownership of a document in order to determine whether it could be legitimately archived. (Take Figure 7: Copyright vs. Self-Archiving Rights) Certain document types were

臺灣大學學術期刊資料庫



ROMEO:	This is a ROMEO **white** publisher
Update:	suggest update for this record
Publisher:	American Institute of Physics
Pre-print:	✓ author **can** archive pre-print (ie pre-refereeing)
Post-print:	✓ author **can** archive post-print (ie final draft post-refereeing)
Conditions:	• Publisher's version/PDF may be used on author's personal website only • Author's own version of final article on e-print servers • Must link to publisher version or journal home page
Copyright:	ftp://ftp.aip.org/aipdocs/forms/copyrght.pdf
ROMEO:	This is a ROMEO **green** publisher
Update:	suggest update for this record
Publisher:	American Mathematical Society
Pre-print:	✓ author **can** archive pre-print (ie pre-refereeing)
Post-print:	✓ author **can** archive post-print (ie final draft post-refereeing)
Conditions:	• Must include publisher statement "First published in [Publication] in [volume and number, or year], published by the American Mathematical Society,"
Copyright:	view policy
ROMEO:	This is a ROMEO **green** publisher

Figure 8: SHERPA's Summary of Publishers' Policies

easier to handle than others. Working papers, technical reports and presentations, for example, could be archived after obtaining the authors' consent. U.S. patents held by our faculty members could also be harvested as they are considered as public domain information.

Document types such as journal articles and book chapters required tactful handling. These are largely copyrighted materials and the publishers may have variant policies towards author's self-archiving. A convenient starting point was the SHERPA/RoMEO publishers' copyright listing <http://www.sherpa.ac.uk/romeo.php> which summarizes the policies of over 100 publishers. (Take Figure 8: SHERPA's Summary of Publishers' Policies) However, not everything was clear cut. While some publishers such as Elsevier and Springer have distinct policies on self-archiving, other publishers have unclear policies, and many others have no policy at all.

The different versions of a paper further complicate the issue. A journal article has a pre-refereed version (preprint), a post-refereed version (postprint) and a publisher's version. A publisher may allow the archiving of a particular version of a paper which differs from the one on hand. We had to observe this constraint and tried to obtain the permitted version from the author. For those publishers that were not analyzed on the SHERPA listing, we searched the Web to identify their poli-

臺灣大學學術期刊資料庫

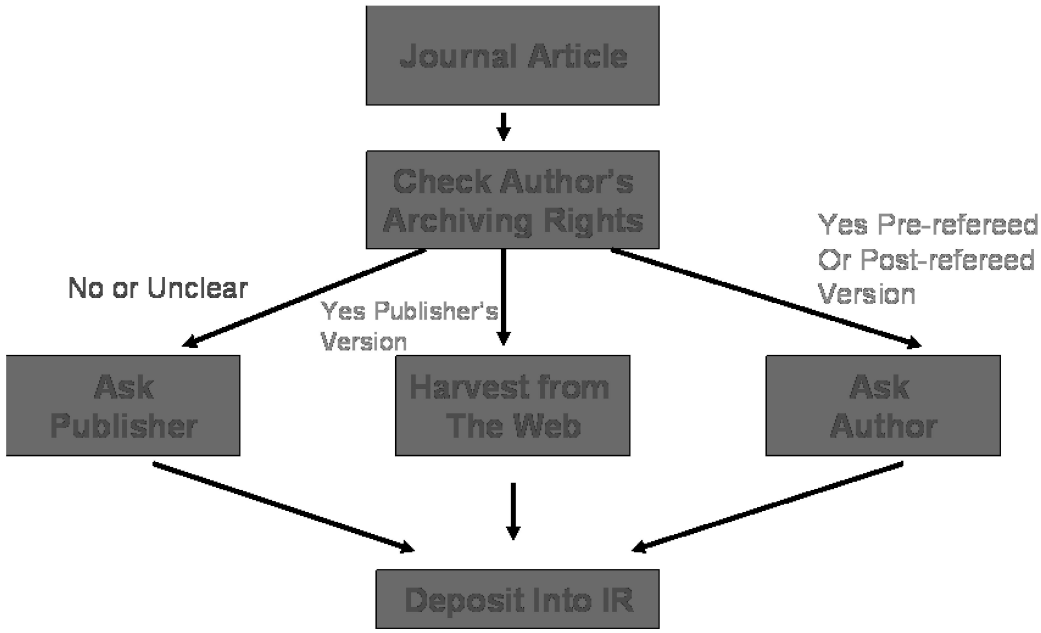


Figure 9: Procedures Ensuring Archivability of Journal Articles

cies and requirements. If we could not trace the policies of certain publishers, we wrote to them to ask for permission.

The Library undertakes the arduous task of deciding what to include and what to exclude. While we would like our researchers to assert their self-archiving rights, we must ensure that all research literature that crosses the IR threshold is cleared of copyright concerns. The process requires identification and interpretation of publisher's policies and frequent liaisons with individual faculty members and publishers. (Take Figure 9: Procedures Ensuring Archivability of Journal Articles)

Strategies to recruit content

Our goal was to build a critical mass of papers within a short time frame. With this in mind, the Library set out to explore different channels to acquire scholarly content for the IR. We emailed all faculty members in late 2003 to invite direct submissions to the IR. The response was pathetic; only two papers were received. Academics have busy schedules, and may consider self-archiving extra administrative work, however much they

臺灣大學學術期刊資料庫

understand and support the idea of open access and institutional repositories. The Library had no choice but to adopt more proactive strategies to acquire content.

a. Scanning websites

The reference librarians checked the departmental and personal homepages of individual faculty members, and learned that a good number of them posted their full-text research publications on the Web. The University Librarian emailed them one by one and obtained permission to post about 150 documents in the IR. At a later stage, a similar scanning exercise was done on all the research centers and institutes on campus. This strategy was effective, but labor-intensive. Enormous amount of time had been spent on obtaining the correct versions of papers and sifting through the archivable and the non-archivable.

b. Working papers and technical reports

We then hoped to capture pre-existing collections of grey literature, i.e., working papers and technical reports. We surveyed all departments and found a wealth of these was tucked away either in printed or electronic format. The University Librarian wrote to these departments and offered to preserve them digitally in the IR. This proved to be an effective means. We managed to add close to 400 papers to our repository.

c. PhD theses

The Library maintains an Electronic Theses Database which provides the metadata, abstracts, and full-text of all the MPhil and PhD theses produced at HKUST. It was decided to include only PhD theses in the HKUST IR. Our colleagues in Acquisitions contacted PhD alumni to secure their consent. To simplify workflow, we created a metadata record for each thesis and provided a link to the full-text file residing in the Electronic Theses Database. A mechanism has been established to capture new PhD theses on a regular basis. To date, we have added 400 PhD theses to our IR.

d. Conference papers

Over the years, HKUST has hosted a number of conferences and symposia. We searched the library catalog and singled out 50 conference proceedings to work on. For those proceedings published in-house, copyright belongs to the university. We targeted those papers authored by HKUST members and emailed them for permission. The Collection Develop-

臺灣大學學術期刊資料庫

ment manager also helped to secure copyright clearance from professional societies and commercial publishers. A total of 142 out of 193 papers identified were archived, which represented a decent success rate of 74%.

e. University archives

With the help of colleagues in the Archives Department, we sifted through a number of publications affiliated to HKUST and added 80 more documents to our IR. Every little harvest counts, however small it is. A bigger and more exciting discovery is that the Department has been receiving boxes and boxes of pre-published research papers from faculty members whose projects are funded by the Hong Kong Research Grants Council. This goldmine of scholarly papers is kept in 160 boxes. They will be inspected one by one for archivable materials.

f. Open access sources

Open access sources were another channel we explored, which turned out to be quite fruitful. HKUST researchers publish in open access venues such as DOAJ and PubMed Central. These sources were scanned for suitable papers for inclusions. Some publishers, to our pleasant surprise, have liberal policies towards author's self-archiving, and allow the posting of the publisher's version in institutional repositories. Many of our researchers are U.S. patent holders. After securing their approval, we collected the patents by using U.S. Patent Fetcher which has the ability to concatenate a multi-page patent into a single PDF file. The size of the HKUST IR jumped by another 200 as a result of all these efforts.

g. Publishers' approval

Our Collection Development manager negotiated with 40 publishers and successfully obtained approval from half of them for 120 journal and conference papers. The Society of Information Display alone accounted for 32 papers. Our experience suggests that the small publishers and professional societies are more inclined to accede to our request. Understandably, many publishers either were wary of our request or simply ignored us. This was not a very effective means, but worth trying nonetheless.

h. Liaisons with faculty members and departments

In addition to soliciting readily available collections of working papers and technical reports from departments, the reference librarians contacted faculty

臺灣大學學術期刊資料庫

members, asking for their complete publication lists, if the ones on the Web were only representative. We asked for the correct versions of papers. We encouraged them to dig up their files from their computers and send us the soft copies. We also reached out to retired and emeritus professors to recruit content. Our efforts sometimes yielded surprisingly good results. A Computer Science professor, for example, agreed to deposit all his papers in the IR.

Naturally, faculty members will be more responsive if they are aware of the self-archiving policies for major publishers in their subject area. We update our faculty members on this via email. For instance, the academic staff from the Computer Science and Electrical and Electron-ic Engineering Departments would appreciate knowing that IEEE accepts the posting of the post-refereed version. We also advise on the embargo period imposed by publishers, and how to sign copyright transfer agreements so as to retain self-archiving rights.

On every instance of contact, we seized the chance to explain what open access is, the trends of open access publishing, increasing governmental support for institutional repositories, and nitty-gritty issues like how to distinguish the pre-refereed version from the post-refereed version of a paper. We introduced new concepts, clarified any misunderstandings, and responded to their worries and concerns.

It was indeed a novel experience for the reference librarians to recruit content for a database "from scratch". The entire process of content building has been an ongoing experiment. It requires a great deal of time, patience, and meticulousness. Tact and coordination efforts are prerequisites. We have to be flexible and persistent. We explore and exhaust all likely channels, both within and outside the library.

IV.Librarians as Advocates

Promotion is an essential element of any IR initiative. It follows after a critical mass of papers has been gathered. We have to reach out to university administrators, department heads, center directors, researchers, and graduate students in whatever appropriate ways: on a one-to-one basis through informal conversations, small group discussions, departmental visits, and campus-wide promotion.

In this connection, we published articles on the development of the HKUST IR in library and university newsletters, and even in local and international newspapers. To raise the visibility of the IR, we held an anniversary celebration in 2004 when our IR topped 1,000 documents. There, we gave prizes to our top

臺灣大學學術期刊資料庫

ten contributors. The event gave a social platform for us to interact with faculty members, and for faculty members to discuss the issues among themselves. We organized informal promotion sessions during which we demonstrated to our attendees the impact and benefits of putting their papers on an open access platform. The events were well received and generated document submissions.

Departmental visits were made to the four schools at HKUST in 2005. A half-hour presentation was followed by a question and answer session. All the workshop attendees were faculty members and PhD students who were interested in the issues of open access and institutional repositories. These outreach occasions helped to dispel most of the mist and worries, and increased their receptivity towards these latest trends in scholarly publishing.

We have been striving for a cultural shift among researchers so that they will develop the habit of archiving their scholarly output in the IR. After months of email announcements and promotional efforts, we have witnessed a shift of attitude among some of our faculty members from apathy to support.

V. Sustainable Growth

Archiving all the university's research output is the ultimate goal of building the HKUST IR. We realize we have made only a very humble start. Since our establishment in 1991, HKUST researchers have published over 8,000 journal articles indexed in the Web of Science. Thus far, we have only succeeded in capturing a tiny fraction of this journal output into our repository: a total of 500.

At the initial stage of content building, the Library focused largely on the archiving of "retrospective materials" or materials that already existed. To sustain the growth of the IR, it is imperative for the Library to establish a channel through which fresh scholarly content can steadily flow in.

At HKUST, faculty members are requested to report the progress of their government-funded research projects to the university administration. They need to submit a report as well as a list of associated publications to our university on an annual basis. In March 2005, the Library worked with the Office of the Vice-President for Research and Development to install an archiving mechanism for the HKUST IR in the university's research output reporting system. As a result, a checkbox was added to the research report form. Those faculty members who wish their publications to be openly accessible and permanently preserved in the IR will select the checkbox. An email notice will then be sent automatically to a library account for follow-up actions. They can

臺灣大學學術期刊資料庫

also choose to deposit their publications right away using the IR Document Submission form. The response so far has been satisfactory. Cumulatively, over 600 positive replies have been received.

The Library continues to explore other means of nurturing the culture of IR archiving among the research staff. More departmental visits will be necessary. One probable strategy is to suggest to the individual departments that they should appoint a designated IR co-coordinator among themselves to oversee the deposit of documents into the IR and to provide bridging support to the Library. How this proposal works out remains to be seen.

VI. Usage and Reference Assistance

The HKUST IR has been in place for three years. Is there any indication to show that it is indeed serving its purpose?

One good benchmark is the reference queries we receive on the IR. Once, we received an email from someone in the U.K. who wanted to contact the author of a PhD thesis. It turned out that the requestor was the father of a son suffering from a type of cancer called Ewing Sarcoma. He discovered the thesis on the IR via the Web. We acted as the intermediary and passed his enquiry to the author concerned. On another occasion, a user as far away from Hong Kong as Argentina downloaded a working paper from our IR, and he sent us a question on the bibliography of the paper. Queries and feedback like these are good indicators showing that information seekers worldwide are benefiting from open access, and are able to tap seamlessly into our institutional scholarly output. Locally, reference librarians find the IR a convenient research tool for tracking down grey literature produced by our faculty members.

In October 2004, our Systems colleagues introduced two new features on the HKUST IR homepage: the total number of document downloads per month and the Top 20, i.e., the 20 most accessed documents of the month. These figures provide good estimates of the number of real web users by excluding access by most search engine robots. A generally rising trend of document download activity from October 2004 to November 2005 can be observed from Figure 10.

A snapshot of the usage pattern of the IR was taken for November 2005. There were about 2,000 items in the database in that month. 5,877 robot-free downloads were recorded. The document with the highest access was a 2004 journal article authored by a faculty member in Civil Engineering. This paper recorded 147 accesses in the month. The number of access among the top 20 documents ranged from 27 to 147. They accounted for 725

臺灣大學學術期刊資料庫

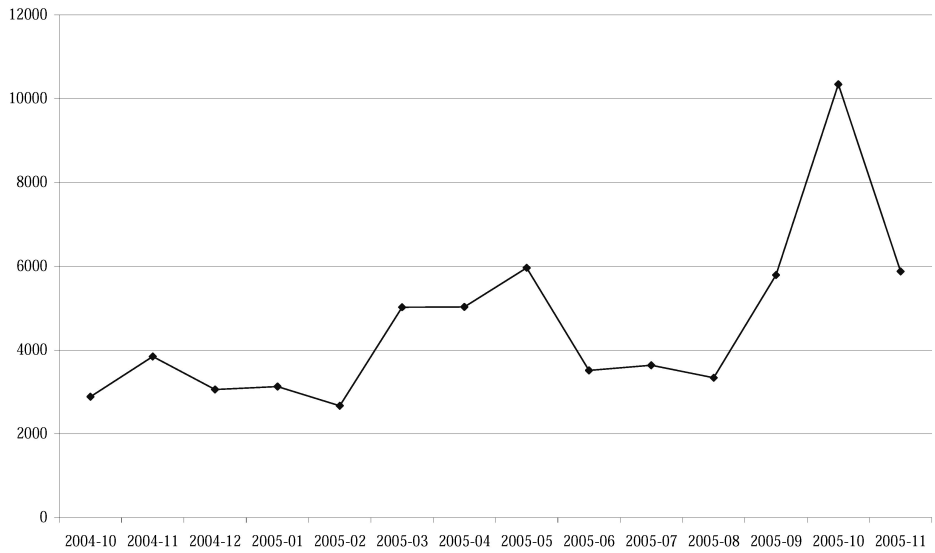


Figure 10: Document Download (Excluding Most Robot Access)

times of access, about 18% of the total access of the database in November.

In the thirteen months from October 2004 to November 2005, a total of 106 documents made into the Top 20 lists. Analyzing this sample of documents gives a good picture of users' interest and usage pattern. Apparently, these documents come from all disciplines: science and engineering, business, social sciences and humanities. Quite a number of documents authored by the HKUST Language Center received high usage. In terms of document types, journal articles and conference papers are the most popular. The cumulated download figures are represented by Figure 11. It is interesting to note that two documents from the Language Center persistently remained on the Top 20 lists. A Finance working paper stayed for 12 months, and two journal articles stayed for 11 months.

VII.Barriers and Challenges

Summing up our three years' experience with developing the HKUST IR, we have met with both successes and failures. Our challenges are multi-faceted, emanating from different directions: our faculty members, the university, the publishers and the Library.

臺灣大學學術期刊資料庫

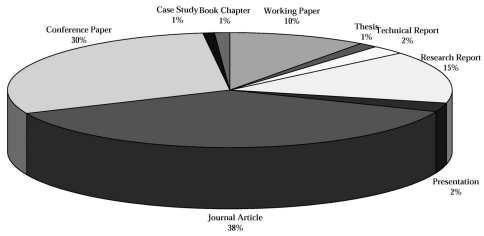


Figure 11: Cumulated Access of 106 documents from Top 20, Oct 04 to Nov 05

a. Faculty members

Generally, faculty members have low awareness of open access, IRs, and self-archiving rights. They normally will not negotiate with publishers for self-archiving rights or non-exclusive copyright transfers. We have to do the ground work for them and advise them accordingly. Many of them have reservations about posting the pre-published version of their papers on the Web without the packaging and endorsement of traditional publishers. Some fear that research papers in IRs are likely to attract plagiarists. It is always difficult to ask for the correct versions of papers from the faculty members. They either do not respond or tell us they do not keep the requested version. Their apathy and resistance undermine the success of the project. Hopefully, continuous advocacy will dissolve these initial sentiments.

b. Institution

Institutional support is vital for the sustainability of any IR. Swan and Brown (2005) conducted a survey on 1,296 journal authors in late 2004. It reveals that 81% of them would willingly self-archive all their publications in their IRs if they were mandated to do so by their employers or funding bodies. We hope to see more and more institutions mandating open access self-archiving. The Eprints Institutional Self-Archiving Policy Registry <http://www.eprints.org/signup/fulllist.php> maintains a list of institutions that adopt such a forceful policy.

At this stage, HKUST may not be ready for such a drastic step. In our view, it would make a tremendous difference if our university administrators could strongly encourage our faculty members to support the cause. The university could also develop coordinating policies such as funding the library and providing financial assistance to faculty members to submit papers to open access journals.

c. Publishers

SHERPA identified that 75% of 127 publishers allowed some sort of self-archiving as of January 2006. The percentage seems high, but more often than

臺灣大學學術期刊資料庫

not, only a handful of articles from a long bibliography can be archived. The main stumbling block is that authors do not usually retain the correct version of their papers as required by the publishers. Another source of frustration is that many publishers have no or unclear policy towards author's self-archiving. To rectify this somewhat chaotic situation, a probable solution is to incorporate an appropriate statement into the license agreements with publishers.

Commercial publishers have been understandably slow in responding to the changing landscape. In May 2004, however, Elsevier took the lead by announcing that it would allow its authors to self-archive their papers in IRs (Peek, 2004). Since January 2005, authors of the Nature Publishing Group can also deposit their manuscripts into their own IRs six months after the original publication. These are good signs pointing towards the development of a complementary platform of scholarly communication.

d. Library

The HKUST IR has placed heavy demands on the Library. The current mode of operation is that the Library absorbs all the extra duties and costs involved. The reference librarians, in particular, shoulder the extra roles in educating, promoting and liaising. We have to grapple with new skills and roles while not neglecting our current duties. As the project has become a long-term commitment, the Library will need additional funding and resources from the government or other bodies. The injection of funding would allow us to purchase additional equipment or software for digital content management, and release precious staff time for IR work.

Looking ahead, we welcome the chance to collaborate with other institutions such as the public libraries, public records office, or non-profit organizations. We can share our knowledge and experience with those institutions which have the financial resources but not the expertise. At this juncture, the Library will need to continue its role to advocate and showcase until the self-archiving culture is ingrained into the work cycles of our research staff.

VIII.Conclusions

A new scholarly communication landscape is in the making. The trends for open access, author's self-archiving and institutional repositories are irreversible. Libraries often stand in the front line to engender change and lobby for support. The process may be slow and haphazard. Despite the many challenges involved, libraries should take stewardship seriously. It certainly pays in the long run to in-

臺灣大學學術期刊資料庫

stall an infrastructure which can manage and guard an institution's intellectual output.

The HKUST Library embarked on the project following the example of Caltech. In December 2004, the two parties co-hosted the International Conference on Developing Digital Institutional Repositories: Experiences and Challenges. We were encouraged by the high level of enthusiasm in the subject. Recently, other conferences and workshops on institutional repositories were held in China, Taiwan and Thailand. These events, sponsored by IGroup, again made a stir among the attendees and many expressed an interest in starting their own IRs to manage their organizational intellectual output. We are positive that more and more IRs will be set up in the Asia-Pacific region in the near future. It is estimated that within ten years, all major universities are likely to be running an IR (Yeates, 2003). We hope this prediction will one day be realized.

References

Albanese, A. (2004). UK report calls for publicly available STM research. Library Journal, 129 (3), 16-17.

Chan, D, Kwok, C. and Yip, S. (2005). Changing roles of reference librarians: the case of HKUST Institutional Repository. Reference Services Review, 33 (3), 268-282.

Douglas, K. (2004). Institutional repositories: preparing for the future. International conference on developing digital institutional repositories : experiences and challenges. Hong Kong: California Institute of Technology Libraries and the Hong Kong University of Science and Technology Library. Retrieved January 23, 2006, from http://library.ust.hk/conference2004/papers/douglas-keynote.doc

Gibbons, S. (2004). Benefits of an institutional repository. Library Technology Reports, 40 (4), 11-16.

Peek, R. (2004). The cascading effect. Information Today, 21 (8), 17-18.

Swan, A. and Brown, S.N. (2005). Open access self-archiving: an author study. Retrieved January 23, 2006, from http://www.jisc.ac.uk/uploaded_documents/Open%20Access%20Self%20Archiving-an%20author%20study.pdf

World university rankings. (2005, October 28). The Times Higher Education Supplement, p.3.

Ware, M. (2004). Universities' own electronic repositories yet to impact on open access. Nature.com's Web Focus: Access to the Literature. Retriveved January 23, 2006, from

臺灣大學學術期刊資料庫

http://www.nature.com/nature/focus/accessdebate/4.html.
Wellcome Trust. (2005), Wellcome Trust announces open access plans. May 19. Retrieved January 23, 2006, from http://www.wellcome.ac.uk/doc_WTX025191.html
Yeates, R. (2003). Institutional repositories. VINE, 33 (2), 96-99.

臺灣大學學術期刊資料庫