
Cataloguing for the world: motivation, method and madness

Peter Sidorko, Connie Lam
University of Hong Kong Libraries

OCLC Asia Pacific Regional Council Membership Conference
8 October 2013, Bangkok, Thailand

Outline

- HKUL
- Cataloging in HKUL
- Chinese e-resources cataloguing
- Challenges and solutions
- Cataloguing outputs

HKUL



- Established in 1912
- Comprises 1 main library + 6 branches
- Collection

	Print	Electronic
Books	2.96 million	3.58 million
Journals	51,513	81,886

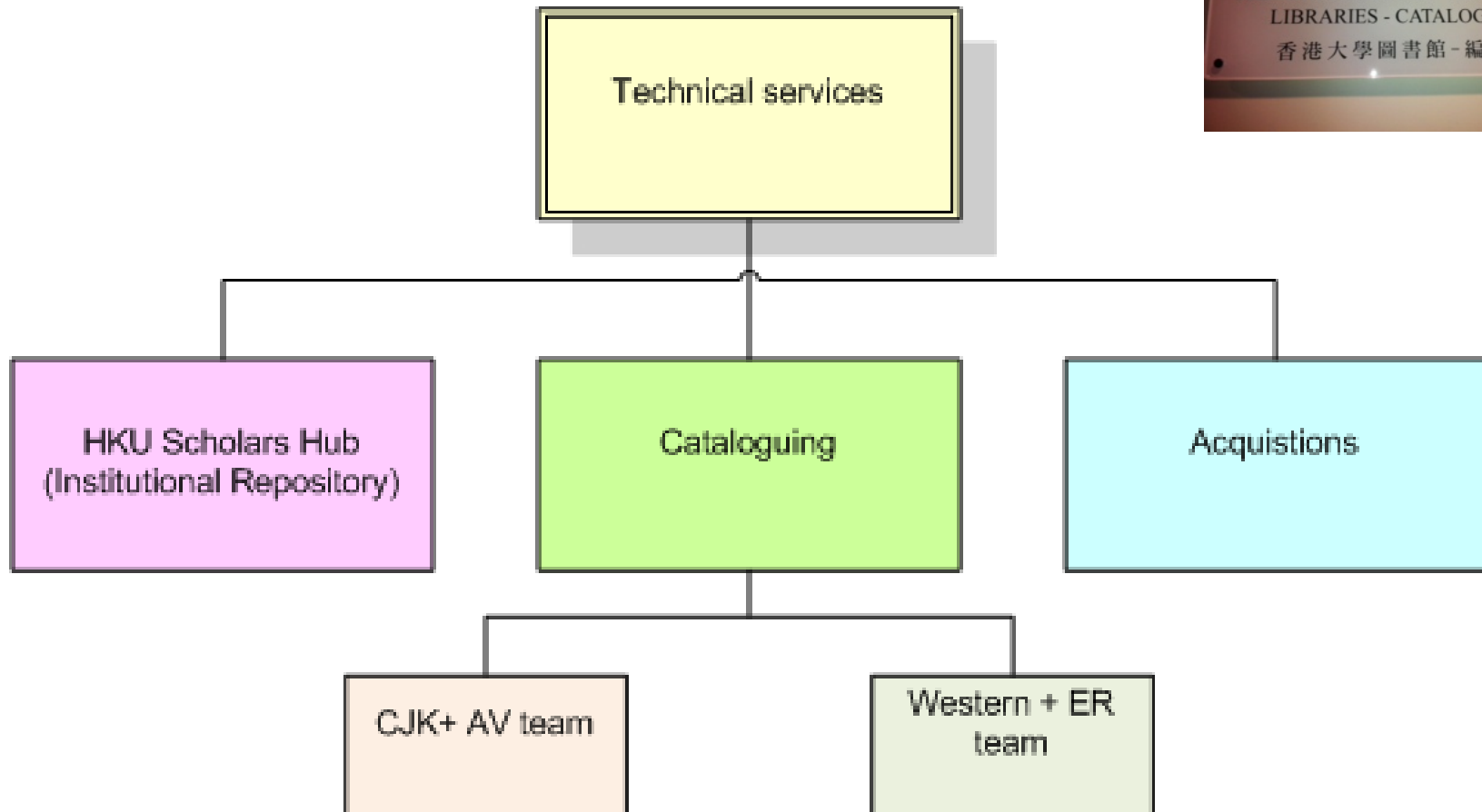
- Staff

Professionals: 32

Support staff: 197

figures as of June 2013

Cataloging in HKUL



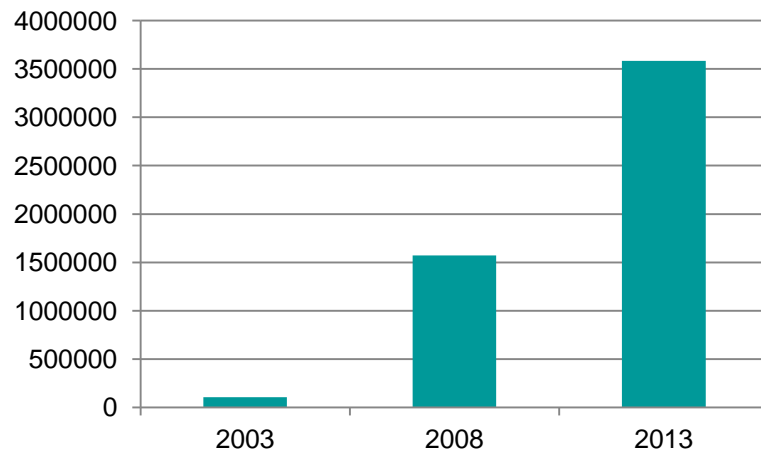
Objectives of HKUL Cataloguing

- Provide **our users** quality bibliographic records in timely manner.
 - Catalog individual titles in databases to easy the search for **our users**.
 - Employ all kinds of technology to speed up the workflow to allow **our users** have promptly access.
- We focus on **our users** → Cataloguing for **our users**.

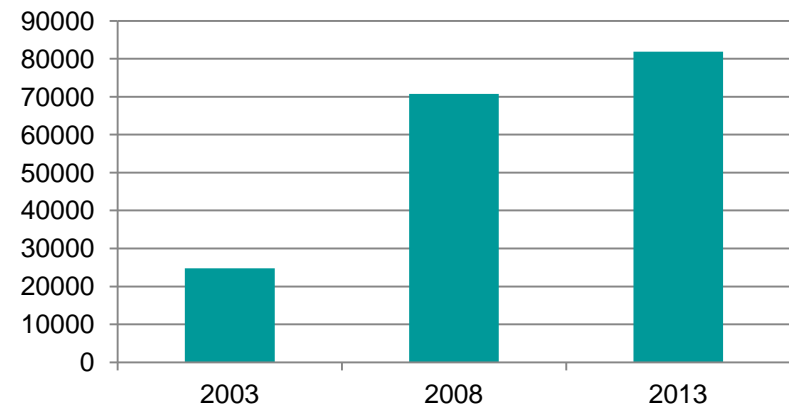
Challenges we encountered

- Rapid increases of e-resources

ebooks

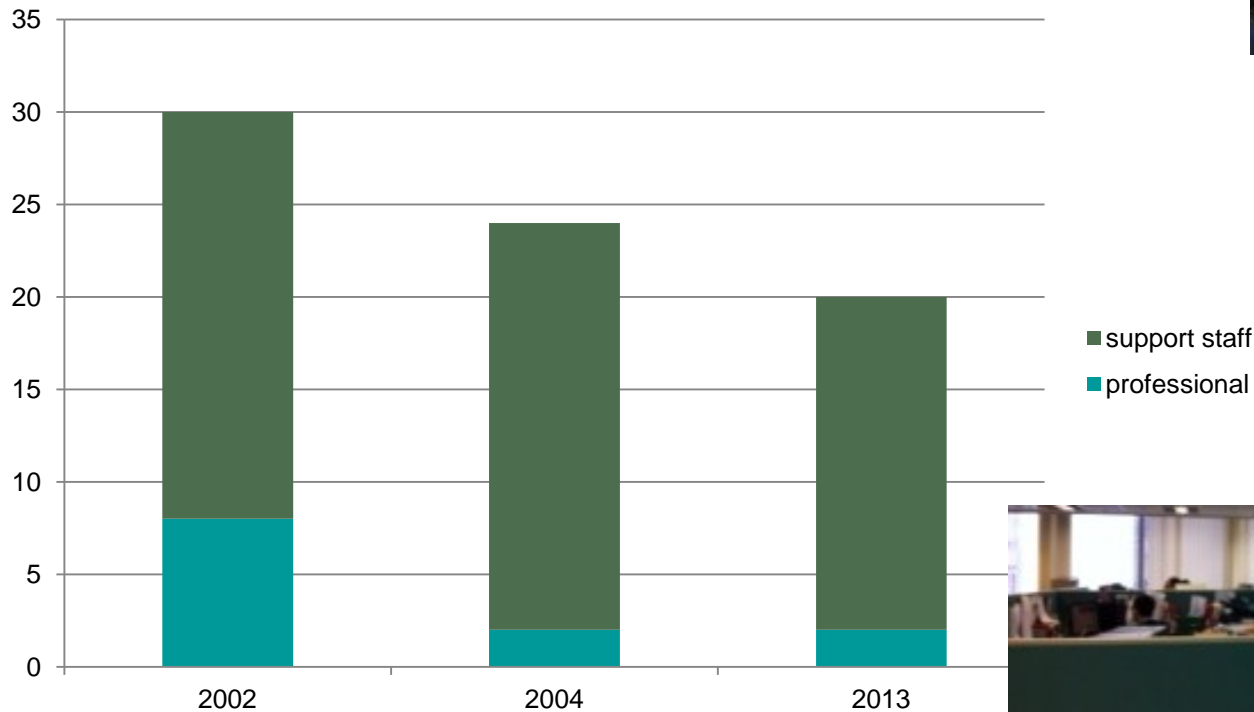


ejournals



Challenges we encountered

■ Decreases of cataloguing staff



What can we do?

- Collection ↑

- Staffing ↓



- Vendor supplied records

- For Western e-resources, especially the large package, we rely on vendor records. The quality of the records normally are acceptable.
- But for Chinese e-resources

The e-Chinese Collection in HKUL

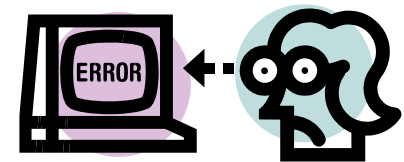
- Major acquisitions (to 2005):
 - 1998: Sikuquanshu (四庫全書)
 - 1999: China Journal Net (China Academic Journal) **5,000+ titles**
 - 2000: China Core Newspapers Database
China Doctors & Masters Dissertations
 - 2002: Sibucongan (四部叢刊)
 - 2003: Apabi e-book **around 40,000 titles**
 - 2005: SuperStar Digital Library **around 400,000 titles**

Our journey in cataloguing Chinese ER

- Chinese Journal Net (CAJ)
 - With the limited information provided by vendor, we started our cataloguing project in 2000.
 - Only a title list in print format. No ISSN, no information on title relationship, no coverage details.
 - We requested vendor to provide more detail information in excel file in later stage, which helped to speed up the cataloging work.
 - The project completed in 2004. We spent 4 years for around 5000 titles.

Our journey in cataloguing Chinese ER

- With the large amount of Chinese ebooks purchased, we asked for Marc records.
- Vendor provided Marc records
 - ❑ Problems with the Chinese scripts.
 - ❑ Not follow AACR2.
 - ❑ Improper tag used, incorrect indicators used.
 - ❑ Not follow LC Pinyin guidelines.
 - ❑ Multi-records for multi-volume set
 - ❑ Multi-records for serial titles.



Our journey in cataloguing Chinese ER

- Communicated with vendor for desired standard.
- Provided advice to enhance the records
- Provided example records to vendor.
- Spent months in communication → no good news → Dead end???
- Why vendor unable to fulfill our requirements
 - Records were generated by metadata in database
 - Problems cannot be tackled by program

In the meantime ...



港大圖書館編目 全球第二

★香港大學圖書館在今年度的原始編目數量，成為全球書目數據庫 (World Cat) 第二位，其提交的資料達到45,758份，超越不少大型圖書館，牛津大學圖書館亦以不足4萬份資料，排名次於港大，而美國研究圖書館中心則以63,549份資料居首。全球書目數據庫有1.2萬成員館，所謂原始編目，就是各成員館根據其所收藏的書籍，將與其他成員館不同的書籍資料提交到數據庫，以港大而言，即有逾4.5萬本書籍，期刊等出版資料是其他成員館所沒有的。

2004年全球書目數據庫排名

機構	原始編目數量
美國研究圖書館中心	63,549
香港大學	45,758
牛津大學	39,322
俄亥俄州州立大學	25,048
威斯康辛州立大學	24,073

資料來源：香港大學

Message from the Librarian

HKU Libraries is Second in WorldCat's Top Ten

Anthony Ferguson

We were pleasantly surprised to learn last month that HKU Libraries is the 2nd largest contributor of original cataloguing records for the year to the WorldCat database of the Online Computer Library Consortium, a membership organisation made up of 60,540 libraries in 84 countries. Our 45,758 records submitted to this database provided a large revenue to the Libraries, and also placed us ahead of such famous and large libraries as the Library of Congress, the British Library, etc. The University of Oxford Library came third with 39,322 records, HKU Libraries second, and the Center for Research Libraries (USA) placed first with 63,549 records. For the second year in a row, our contribution has exceeded 40,000. This is a reflection of the massive amount of work that our Technical Services staff members are able to do year after year.

With severe budget cuts, this high contribution of unique records is one way that HKU Libraries is able to claw back some of the costs, and in some cases, pay completely for some of the many services we subscribe to and offer our users.

With over 12,000 libraries contributing records to WorldCat, and with over 53 million records, this database is the de facto worldwide union catalogue. When searching WorldCat, the record is usually already entered. With so many contributors and so many libraries buying and cataloguing the same materials, it is very rare for libraries to enter so many unique records in the way that HKU Libraries has done and continues to do. We have been able to do this because the nature of our contributed records is relatively rare compared to other would-be contributors to WorldCat. We are the largest catalogue of Chinese materials in WorldCat. We are the largest catalogue of Chinese e-journals and ebooks in WorldCat. On the other hand, for most of our English language materials, instead of contributing records, we download records from WorldCat.

As the largest and premier cataloguing database of the world, WorldCat allows contributing members such as HKU Libraries to share resources worldwide with other libraries and universities. In this manner we can increase HKU's reputation and further fulfill its mission of disseminating knowledge. End-users use WorldCat for information discovery. Upon finding that it resides in HKU Libraries,

Contents

Message from the Librarian	1
Exhibitions	2-3
Book Talks	4
SuperStar in Libraries	5
HKU e-Theses: 1,000 More	5
WebWise	6
Spotlight on Treasures	7-8
Feature Collection	9
Selected Notable Acquisitions	10
Library Updates	11
Upcoming Events	11
Library Visitors	12-14



they can visit the Libraries, request inter-library loan, and now increasingly, simply click to connect to the resource at HKU Libraries. WorldCat is now searchable through Yahoo and Google, which means that HKU Library records will begin to appear in these two giant search engines.

On top of this 45,758 unique records added to WorldCat, HKU Libraries sent 350,000 records last year to OCLC so that our retrospective collection of Chinese materials also could be shown in WorldCat. Records for our western language materials had been sent the year previous.

These numbers are very surprising in light of the events our Technical Services staff members have experienced in the last two years. We lost several key staff due to the University's voluntary retirement scheme. Consequently, we restructured Technical Services and rationalised its workflow. These numbers show that these changes have worked very well indeed! 📖

Top 10 original cataloguers

Organisation	Original Records Added Online
Center for Research Libraries	63,549
University of Hong Kong	45,758
University of Oxford	39,322
The Ohio State University	25,048
University of Wisconsin	24,073
Godfrey Memorial Library	23,647
University of California, Berkeley	23,223
Singapore Integrated Library Automation Services	19,165
Union Theological Seminary	15,729
University of Texas at Austin	15,309
Fiscal 2004	



Our journey in cataloguing Chinese ER

With severe budget cuts, this high contribution of unique records is one way that HKU Libraries is able to claw back some of the costs, and in some cases, pay completely for some of the many services we subscribe to and offer our users.

+

contributors to WorldCat. We are the largest cataloguer of Chinese materials in WorldCat. We are the largest cataloguer of Chinese ejournals and ebooks in WorldCat. On the other hand, for most of our English language materials, instead of contributing records, we download records from WorldCat.

= Way out

Our journey in cataloguing Chinese ER

- We decided:
 - Request vendors to provide title lists with the following data in excel file
 - Author
 - Title
 - Publisher
 - ISBN
 - Direct url
 - Control number
 - Create records ourselves, upload to OCLC to earn more credits.

Our strategy

- Handle titles we have print records first.
- We estimated that 20-30% of the titles have print records in HKUL catalog.
- Based on the title list, we developed scripts
 - To search ISBN or title of print records in HKUL catalog.
 - To generate records based on the print records.
 - To add url and control number to records.
 - To use “Global update function” to convert records from print to electronic.

Try and error process

- Search titles by program → Fail
 - Simply rely on title is not reliable.
 - A lot of titles in the file have errors.
 - Same title sometimes has duplicate entries due to
 - Different punctuation used
 - Simplified/Traditional characters
 - Different space used
 - We tried one batch and found that too much effort required for the clean up.

Try and error process

- Search ISBN by program → Fail
 - The matching rate is much lower than our estimation.
 - Some publishers use same ISBN.
 - For different titles.
 - For same title, but different editions.
 - We used ISBN program to process 3 batches and have to re-check every title in these batches to ensure the records are correct.

Our solution

- The title list provided has a lot of incorrect data and duplicated entries.
- Manually search for matching records is the only solution.
- We expand the search to other libraries via Z39.5 and also OCLC.
- We set up a small project team for cataloguing of Chinese e-books.

Our tactics

- Outsourcing cataloguing of Western monographs to free some manpower.
- Load vendor supplied records to our catalog for patron driven acquisition package without editing.
 - Not to waste time to handle titles we don't buy
 - Upgrade those titles we purchased
- Minimize local practice
 - Give up all local practices, except Hong Kong subject headings for Hong Kong materials.

Our tactics

- Plan more before action.
 - Analyze title list to locate serial titles, collectanea and have them excluded from searching list.
- Members of the newly set up team are mainly responsible for searching.
 - Minimal training
 - High productivity
- Gradually develop their cataloging skills by assigning different tasks based on performance.

Our tactics

- Task breakdown
 - ❑ Searching
 - ❑ Descriptive cataloging
 - ❑ Authority checking
 - ❑ Call number and subject assigning
 - ❑ Review and quality control
- Matching work with staff skills
 - ❑ Tasks are allocated to staff according to experience and knowledge
 - ❑ Staff only handle tasks they are able to handle



Our outcomes

- Titles in collectanea which are not searchable in the databases, can be located by searching HKUL catalog.
- More precise title information in HKUL catalog than the databases which help users to locate the resources easily.
- By uploading records to OCLC and with Open WorldCat, our users can identify the e-books by searching on the web. Our collection becomes more visible.

Start Over Return to Browse Limit this Search Search As Word Another Search (Search History)

TITLE 艾蕪密藏 View Entire Collection Search

Limit results to available items

Limited to: Material Format "EBOOK"

Save Marked Records Save All On Page

TITLES (1-2 of 2)

<input type="checkbox"/>	1	EBOOK	艾蕪密藏 [electronic resource] ; 一卷 ; 河洛私見 : 一卷 ; 太極圖說 : 一卷 / 蔡清撰. Cai, Qing, 1453-1509. Jinan : Qi Lu shu she, 1995. Click to access	1995
<input type="checkbox"/>	2	EBOOK	艾蕪密藏 [electronic resource] ; 河洛私見 ; 太極圖說 / 蔡清撰. Cai, Qing, 1453-1509. Shanghai : Shanghai gu ji chu ban she, [1995] Click to access	1995


Chinamaxx

Title 艾蕪密藏 Search
[Help](#)
[Advanced Search](#)

Search Within the Results

No records

AUTHOR [Cal. Qing, 1453-1509.](#)
[蔡清, 1453-1509.](#)

TITLE **Ai an mi zhen [electronic resource] : yi juan ; He lu si jian : yi juan ; Tai ji tu shuo : yi juan / Cao Qing zhuan.**
艾菴密箴 [electronic resource] : 一卷 ; 河洛私見 : 一卷 ; 太極圖說 : 一卷 / 蔡清撰.

IMPRINT Jinan : Qi Lu shu she, 1995.
濟南 : 齊魯書社, 1995.

EDITION Di 1 ban.
第1版.

Permanent URL for this record=> <http://library.hku.hk/record=b3716244>

[Click to view the book via SuperStar digital library](#)

OCLC # ocn124085811

DESCRIPT. p. 662-695 : ill. ; 27 cm.

SERIES Si ku quan shu cun mu cong shu. Zi bu ; di 6 ce.
四庫全書存目叢書. 子部 ; 第 6 冊.
Si ku quan shu cun mu cong shu. Zi bu Ru jia lei.
四庫全書存目叢書. 子部儒家類.
[Si ku quan shu cun mu cong shu. Zi bu ; di 6 ce.](#)
[四庫全書存目叢書. 子部 ; 第 6 冊.](#)
[Si ku quan shu cun mu cong shu. Zi bu Ru jia lei.](#)
[四庫全書存目叢書. 子部儒家類.](#)

NOTE Yu "Zhongguo ke xue yuan tu shu guan cang Qing Qianlong qi nian (1742) Cai Tingkui ke ben" ying yin.
據"中國科學院圖書館藏清乾隆七年(1742)蔡廷魁刻本"影印.
Fu "Si ku quan shu zong mu, Xu zhai san shu wu juan shu" ti yao.
附《四庫全書總目·虛齋三書無卷數》提要.
"Xu zhai san shu ... yi Kan Tai ji tu shuo gai ming Tai ji tu shu, yi Kan He tu Luo shu shuo gai ming He Luo si jian er zeng yi Ai an mi zhen wu shi tiao gu yue san shu".
"虛齋三書... 以看太極圖說改名太極圖說, 以看河圖洛書說改名河洛私見而增以艾菴密箴五十條故曰三書".
Reprint. Each page represents 2 leaves of the original.
Electronic reproduction. Beijing : SuperStar, 2005. System requirements: SuperStar Reader. Available via World Wide Web.

LC SUBJECT [Yi jing.](#)
[易經.](#)
[Ethics -- China.](#)
[Show similar items](#)
[Divination.](#)
[Divination -- China.](#)
[Show similar items](#)

ALT AUTHOR [Cal. Qing, 1453-1509, He lu si jian, 1995.](#)
[Show similar items](#)
[蔡清, 1453-1509, 河洛私見, 1995.](#)
[Show similar items](#)

The screenshot shows the SuperStar search interface. At the top, there is a search bar with the text "Title" and a dropdown menu showing "四庫全書存目叢書. 子部 第6冊.". To the right of the search bar is a "Search" button and a "Help Advanced Search" link. Below the search bar, there is a checkbox labeled "Search within the Results".

The search results section shows a list of books. The first result is:

Books on [四庫全書存目叢書. 子部 第6冊](#) --1 titles. Elapsed time: 0.156 seconds, the current page 1/1 Sort by: Title | Author | Publication Year

1  [《四庫全書存目叢書 子部 第6冊 子部·儒家類》](#)
Author:四庫全書存目叢書編纂委員會 Pages:924 Book concern:齊魯書社 Publication date:1995年09月第1版 ISBN:7-5333-0478-0
Title of the series:四庫全書存目叢書
Brief:大陸新領中國大陸發行
Key words:四庫全書(学科:內容提要) 叢書(地点:中國 年代:清)
[收藏](#) [打印](#)

[Start Over](#)
[Return to Browse](#)
[Limit this Search](#)
[Search As Word](#)
[Another Search](#)

Limit results to available items

Limited to: Material Format "EBOOK"
 Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) ... [51 Next](#)

[WebBridge](#)
[HKALL](#)

TITLES (1-50 of 2502)

<input type="checkbox"/>	1	EBOOK	家語正義 electronic resource] : 十卷 / 姜兆錫撰. Jiang, Zhaoxi, 1666-1745. Jinan: Qi Lu shu she, 1995. Click to access	1995
<input type="checkbox"/>	2	EBOOK	孔叢子正義 electronic resource] : 五卷 / 姜兆錫撰. Jiang, Zhaoxi, 1666-1745. Jinan : Qi Lu shu she, 1995. Click to access	1995
<input type="checkbox"/>	3	EBOOK	孔子家語 electronic resource] : 八卷 / 何孟春註. Jinan : Qi Lu shu she, 1995. Click to access	1995
<input type="checkbox"/>	4	EBOOK	千秋金鑑錄 electronic resource] : 五卷 / 張九齡撰. Zhang, Jiuling, 678-740. Jinan : Qi Lu shu she, 1995. Click to access	1995
<input type="checkbox"/>	5	EBOOK	太極圖說論 electronic resource] : 十四卷 / 王嗣槐撰. Wang, Sihuai, b. 1620. Jinan : Qi Lu shu she, 1995. Click to access	
<input type="checkbox"/>	6	EBOOK	太極圖說論 electronic resource] : 二卷 ; 通書釋義 : 一卷 / 舒芬撰. Shu, Fen, 1484-1527. Jinan : Qi Lu shu she, 1995. Click to access	
<input type="checkbox"/>	7	EBOOK	正統釋 electronic resource] : 四卷 / [張載撰] ; 高學能集註 ; 徐必遠發明.	

<input type="button" value="Save Marked Records"/> <input type="button" value="Save All On Page"/> <input type="button" value="Save to My Lists"/>			
TITLES (1-12 of 243)			
1	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995
2	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995
3	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995
4	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995
5	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995
6	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995
7	<input type="checkbox"/>	四庫全書存目叢書. 子部 [electronic resource] / 四庫全書存目叢書編纂委員會編纂. 濟南: 齊魯書社, 1995.	1995

<input type="checkbox"/>	1	EBOOK	20世紀中國散文英華. 京華卷 [electronic resource] / 哈九增, 李兵, 任麗青編選. Shanghai : Fu dan da xue chu ban she, 1999. Click to access More...
<input type="checkbox"/>	2	EBOOK	20世紀中國散文英華. 巴蜀·荆楚卷 [electronic resource] / 潘頌德編選. Shanghai : Fu dan da xue chu ban she, 1998. Click to access More...
<input type="checkbox"/>	3	EBOOK	20世紀中國散文英華. 江南·嶺南卷 [electronic resource] / 黃樂琴編選 Shanghai : Fu dan da xue chu ban she, 1998. Click to access More...

檢索結果共9條

自定義排序: [按相關度排序](#) [出版時間↑](#) [出版時間↓](#) - 只顯示可整本閱讀的書



20世紀中國散文英華

作者: 吳歡章 沙似鵬 / 出版社: 復旦大學出版社 / 出版時間: 1998

歌林霖東方情懷

辭著邁向新的高度——《20世紀中國散文英華》跋

[線上閱讀](#)



20世紀中國散文英華

作者: 吳歡章 沙似鵬 / 出版社: 復旦大學出版社 / 出版時間: 1999

吳歡章總序: 百年文海難夕

葛乃福前言: “畢竟這是一個散文的世紀”

[線上閱讀](#)



20世紀中國散文英華

作者: 吳歡章 沙似鵬 / 出版社: 復旦大學出版社 / 出版時間: 1997

吳歡章總序: 百年文海難夕

哈九增前言: 浦江浩浩東流去

[線上閱讀](#)



20世紀中國散文英華

作者: 吳歡章 沙似鵬 / 出版社: 復旦大學出版社 / 出版時間: 1997

吳歡章總序: 百年文海難夕

潘頌德前言: 獨具魅力的藝術長廊

[線上閱讀](#)



20世紀中國散文英華

作者: 吳歡章 沙似鵬 / 出版社: 復旦大學出版社 / 出版時間: 1998

吳歡章總序: 百年文海難夕

<input type="checkbox"/>	4	EBOOK	20世紀中國散文英華. 海上卷 [electronic resource] / 哈九增, 陳偉娟編選. Shanghai : Fu dan da xue chu ban she, 1997. Click to access More...
<input type="checkbox"/>	5	EBOOK	20世紀中國散文英華. 海外游子卷 [electronic resource] / 嚴蓓編選. Shanghai : Fu dan da xue chu ban she, 1997. Click to access More...
<input type="checkbox"/>	6	EBOOK	20世紀中國散文英華. 香港澳卷 [electronic resource] / 葛乃福編選. Shanghai : Fu dan da xue chu ban she, 1999.
<input type="checkbox"/>	2		20世纪中国散文英华 [电子资源] / 吴欢章, 沙似鹏. [中国]: 复旦大学出版社, 1998.
<input type="checkbox"/>	3		20世纪中国散文英华 [电子资源] / 吴欢章, 沙似鹏. [中国]: 复旦大学出版社, 1998.
<input type="checkbox"/>	4		20世纪中国散文英华 [电子资源] / 吴欢章, 沙似鹏. [中国]: 复旦大学出版社, 1999.
<input type="checkbox"/>	5		20世纪中国散文英华 [电子资源] / 吴欢章, 沙似鹏. [中国]: 复旦大学出版社, 1997.
<input type="checkbox"/>	6		20世纪中国散文英华 [电子资源] / 吴欢章, 沙似鹏. [中国]: 复旦大学出版社, 1997.
<input type="checkbox"/>	7		20世纪中国散文英华 [电子资源] / 吴欢章, 沙似鹏. [中国]: 复旦大学出版社, 1997.
<input type="checkbox"/>	8		20世纪中国散文英华 [电子资源] / 潘颂德. [中国]: 复旦大学出版社, 1998.
<input type="checkbox"/>	9		20世纪中国散文英华 江南·岭南卷 [electronic resource] / 主编吴欢章, 沙似鹏; 黄乐琴编选. 上海: 复旦大学出版社, 1998.
<input type="checkbox"/>	10		20世纪中国散文英华 [electronic resource] / 吴欢章, 沙似鹏主编. 海上卷 / 哈九增, 陈伟娟编选. [electronic resource] 上海: 复旦大学出版社, 1997.

Our outcomes

- We developed our skills and knowledge in
 - Planning the work
 - Monitoring the workflow
 - Employing different automated tools

The e-Chinese Collection in HKUL

- Major acquisitions (from 2005)
 - 2008:
 - China yearbook full-text database
 - China reference works online
 - 2009:
 - Database of Chinese classic ancient books (中國基本古籍庫)
 - Taiwan electronic periodical services
 - Apabi e-references
 - 2010: Chinamaxx digital library
 - 2011: China online journals. [Medical journals]
 - 2012: Wanfang video. Medicine

Chinese e-resources we cataloged

- Chinese e-journals : 18,273 titles
- Chinese e-books: 447,051 titles
- Chinese e-news: 947 titles
- We manage to have records for all our e-resources, except:
 - Chinamaxx digital library (50% cataloged)
 - China Doctors & Masters Dissertations (no title list provided)

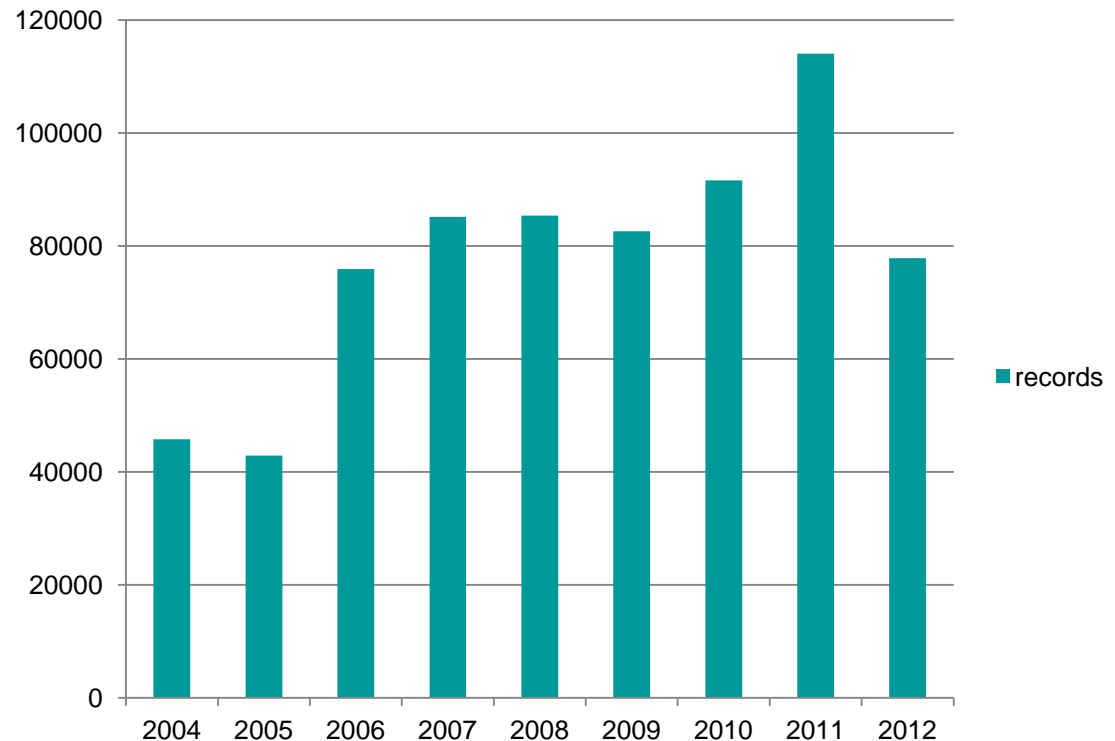
Collection sets in OCLC

- Colleagues in US learnt that we have all China academic journals (CAJ) cataloged and urged OCLC to approach us for creating a collection set.
- We created our first set for CAJ in 2006
- Added another set for China Yearbook Full-Text Database in 2009
- We maintain the updates of these two sets

Cataloging outputs

- Top 3 original catalogers in OCLC since 2004

Year	Rank
2004	2
2005	3
2006	1
2007	2
2008	2
2009	1
2010	1
2011	2
2012	1



Are we mad?

- While others are accepting vendor records even they are poor in quality, we decide to create records by our own.
- While others are waiting for records in OCLC for downloading, we contribute huge amount of records to OCLC.
- NO!
 - It is because we understand both ways are unable to provide our users quality records in timely manner.

Conclusion

- Cataloging for our users is our objective.
- **Cataloging for the world** is the side effect.

Thank You!

Peter Sidorko
peters@hku.hk
Connie Lam
csllam@hku.hk