# CALIS Digital Library Project—— Repository of Special Collections in Chinese Higher Education

Liu, Xia
Wuhan University Library
October, 2008

## **Outline**

- Principles, methods & contents
- Project organization & management
- Project implementation
- Project achievement
- Future work

- Repository of Special Collection (RSC) in Chinese higher education is one of the important sub-projects of China Academic Library & Information System (CALIS) during the "tenth five-years plan" of China
- RSC project Approved by CALIS in 2003

#### General principles & objects

- Creating distributed databases in clusters by subprojects/different libraries.
- Adopting uniform standards or criteria in the database construction
- Forming a unified RSC system based on integrated metadata sets in the central portal and distributed digital objects in the local system
- Providing open access to the metadata and document delivery service to CALIS members upon the contract agreements

#### Building methods

- Adopting appropriate standards and protocols
  - DL standards and protocols, such as OAI Protocol, OpenURL and METS
  - unified Metadata Standards to facilitate interoperability and establish RSC central portal which based on integrated metadata set
- Establishing a series of criteria for local system, central portal and digitalization process
- Unifying functions of interface and services for subprojects
- Using centralized management and defining quality control methods

- The contents digitized mainly focus on
  - Special educational collections on key disciplines
  - Unique resources of research topics
  - Interdisciplinary studies
  - Regional and historical humanistic resources
  - Special collections owned by only one or a few libraries, or scattered information resources which is hard to use

## **Outline**

- Principles, methods & contents
- Project organization & management
- Project implementation
- Project achievement
- Future work

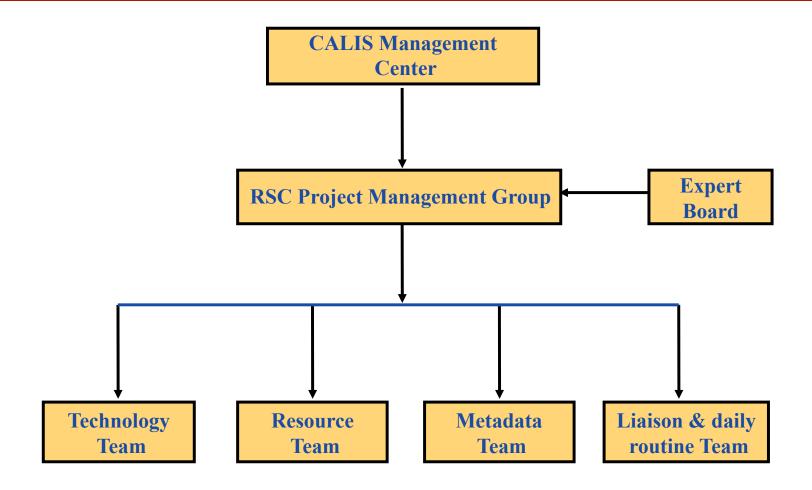
#### Organization and management

- RSC project is organized by the Project Management Group which run under Wuhan University Library
- Project Management Group consists of 7 members from the following academic libraries
  - Peking University Library
  - Tsinghua University Library
  - Wuhan University Library
  - China Agricultural University Library
  - Shanghai Jiao Tong University Libraries
  - Xi'an Jiao Tong University Libraries
  - Jilin University Library

#### Organization and management

- The Management Group set up four project teams dealing with the whole project operation and progress:
  - Technology team
  - Resource team
  - Metadata team
  - Liaison and daily routine team

#### Structure of management organization



# **Project Funding & Budget**

Initial budget 1 million yuan, additional investment 300,000 yuan, total in 1,300,000 yuan

Item		Planned outlay (yuan)	Actual outlay (yuan)	
Project management		50,000	50,000	
Fund	Key-funded	450,000	450,000	
Fund	Normal-funded	450,000	360,000	
Awarded		0	340,000	
Platform		0	100,000	
Others		50,000	0	
Total		1,000,000	1,300,000	

## **Outline**

- Principles, methods & contents
- Project organization & management
- Project implementation
- Project achievement
- Future work

- There are 7 steps in the whole project constructional flow
  - Publicity of the project initiating, including planning, contents and requirements
  - Candidates of sub-projects' application
  - Review and approval by Project Expert Board
  - Project Implementation
  - Mid-term examination
  - System trial running and probative period
  - Final examination and evaluation by Project Expert Board

- 2003.12-2004.3
  - Project planning and preparation
  - Peer review of the sub-projects' applications by the Experts Board
- **2004.4-12** 
  - Formulation & development of technology criteria and standards
  - testing & validation of local system
  - User training of local system
  - Constructing of sub-projects

- **2005.1-12:** 
  - Design and development of RSC central portal
  - Key phases for sub-projects construction
- 2006.1-3 : Mid-term examination of the subprojects
- 2006.4-6: Trial running of Central portal and the entire system
- 2006.6 : RSC Project passed final examination and evaluation by the Expert Board of CALIS

- **2006.9-2007.3** 
  - Evaluation of the sub-projects
- **2007.4-6** 
  - Peer to peer review of the sub-projects by the Expert Board and Management Group of RSC
- **2007.6-7** 
  - Announcement of the sub-projects evaluation results, the RSC project achieved goals of the "tenth five-years plan" of CALIS

## **Outline**

- Principles, methods & contents
- Project management & organization
- Project implementation
- Project achievement
- Future work

- 91 applications from 65 academic libraries applied to join the project
- Through the peer review of Project Expert Board, 75 sub-projects were granted and divided into three categories:
  - 9 key-funded projects, each in 50,000 yuan, total in 450,000 yuan
  - 37 average-funded projects, each in 10,000 yuan, total in 370,000 yuan
  - 27 non-funded projects, supported mainly with technical standards, database building criteria and expert consultation services

- 63 sub-projects are finished and 58 passed reviewed evaluation by the Expert Board
- some statistics of 55 sub-projects by the end of May 2006:
  - The total records are 4.45 million
  - The full-text documents have reached 1.76 million items which account for the 40% of the whole repository

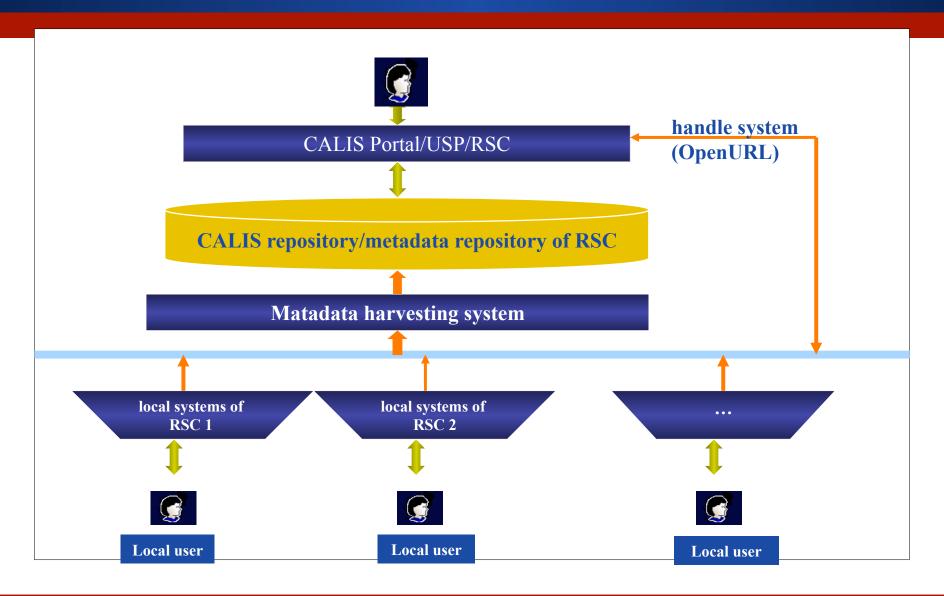
The whole project achieved the anticipated goals

- 1.RSC developed a set of standards and criteria, and gained useful experience on practicing them in large scopes
  - Adopted metadata descriptive standards from Chinese Digital Library Standards developed by the Ministry of Science and Technology project
  - Developed a series of technical criteria for local system, central portal and digitalization process

- 2. The whole system architecture is based on three technical protocols
  - OAI Protocol for Metadata Harvesting (OAI)
  - Metadata Encoding & Transmission Standard (METS)
  - Object Identifiers developed by CALIS (CALIS-OID)

- 3.Established metadata harvesting and object accessing mechanism between central server and local systems
  - The whole system architecture is constituted of two levels, the central service system of CALIS-SCR and the local systems for the partners, which form the structure of centralized metadata sets and distributed object data.
  - Central service provider harvest metadata from data provider of local system through OAI or by FTP.
  - Central service provider access object data from local system through OpenURL protocol.
  - Central service system is part of CALIS Central Union Search Platform, which consists of integrated digital library system with other CALIS resources

#### Data exchange and user service structure

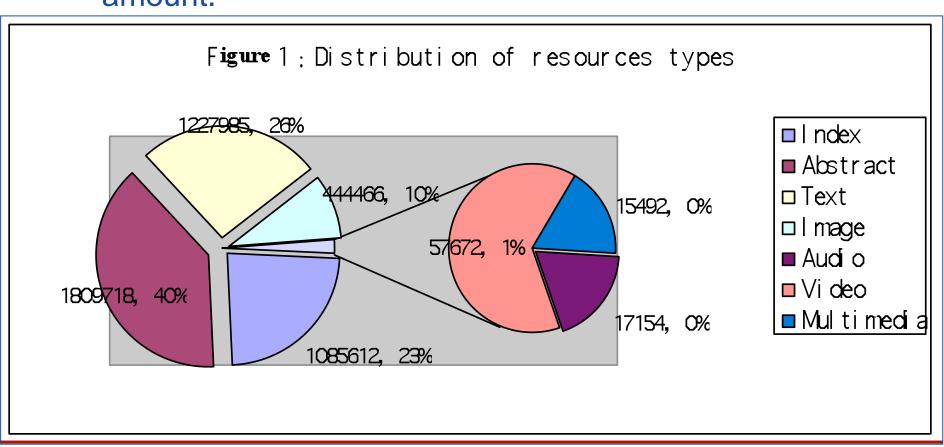


#### 4. Plentiful subjects on special resources

The subjects of 55 special collections by the end of May 2006

Subject	Data amount	Subject	Data amount	Subject	Data amount
Local document	460,294	Agriculture	265,811	Light engineering	154,009
Historical document	407,044	Economics	15,576	Constructional engineering	89,676
politics	77,627	Education	15,703	Hydraulic engineering	371,418
Medicine	179,436	Materials science	165,863	Transportation	83,105
Literature and art	45,250	mining petroleum	86,271	Aerospace engineering	63,394
Personality	116,600	Machinery industry	409,100	Environmental science	83,357
Regional culture	67,466	Communication technology	197,710	<u>Others</u>	7,953

5. The types of digital resources are plentiful, and the proportion of full texts has exceeded 40% of the total amount.



#### 6.Key-funded projects run well

Sub-Project	Number of records	Resource type	Proportion of full text
Digital Library Database on the History of Science & Technology in China	127,335	Text, image, audio, video	50.4%
Yangtze River Resource Database	199,951	Text, image, video, audio	33.83%
Robot Information System	150,823	Text, image	27.14%
Northeast Asia Research Database	91,375	Text, image, video	21%
Ancient Chinese Rare Book and Documentation in Academic Libraries	201,055	Text, image	54%
Database on Southeast Asia & Fujian- Taiwan Studies	95,271	Text, image, video,audio	51%
Large Zhu River Delta Area Research Database	68,856	Text, image	20%
Ancient Education System	47,000	Text, image, multimedia	96%
Special Collection on Bashu Culture	107,521	Text, image	20.5%

# 7. Functions of local systems performed smoothly by sub-projects

- There are 8 Local systems selected by 60 academic libraries. All the systems were validated under the unified identifying and testing by CALIS and RSC Management Group, in order to ensure the consistence and integrity of the running platform.
- Open platform, unified constructional style, unified standards, all of these, trying to insure the quality of databases and perfect resource sharing in a DL environment with large scopes

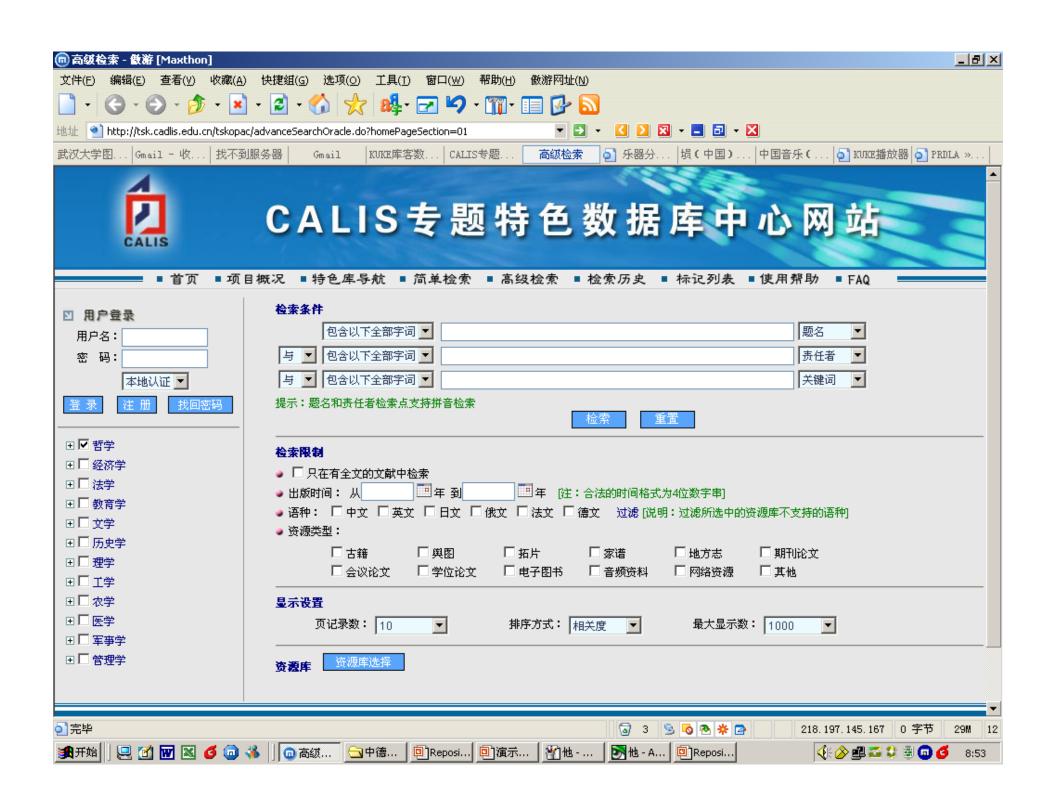
#### 8. The integrity of central portal is successful

- All sub-projects issue on the platform of CALIS-RSC central portal, which provides the hyperlinks (detail records/full-texts/objects) to the local systems
- Central service portal harvests metadata from local system and stores in a central database
- Full-text documents and digital objects are provided directly by the local systems
- Central portal provides to the public free access to the secondary documents

**CALIS-RSC** central portal:

http://tsk.cadlis.edu.cn/tskopac/





#### Personalized service web page of central portal





■ 首页 ■項目概况 ■特色库导航 ■ 简单检索 ■ 高级检索 ■ 检索历史 ■ 标记历史 ■使用帮助 ■ FAQ

欢迎: 4 [本地用户]

建出臺東

- 个人信息
- ○修改密码
- 我的电子书架
- ○我的检索式
- ○检索式定制

#### 我的电子书架

记录数:21 总共页数:3 当前页码:1

首页 上一页 下一页 末页 图 景縣 图 打印 图 电邮 图 下數

全选	序号	題名	责任者	资源库	类型
	1	運動物估自成都至關外巴衷一帶路工清冊並圍脱	李鴻治	高校古文獻特色库	20
	2	清理二十七保十圓並南十二圓载塚楽由	陸文蔵	高校古文獻特色库	2
	3	清上晚晴楼游稿:4卷,晚晴丙子游2卷	初始	高校古文獻特色库	R
	4	清内関庫貯舊福報刊:6線,並首編,於緑	方胜生	高校古文獻特色库	10
	5	明洁史料甲、乙、丙擒	國立中央研究院歷史語言研究 所	高校古文獻特色库	R
	6	Low Power Data Dependent Transform Video and Still Image Coding	Xanthopoulos, Thucydides	CALIS NDLTD	
	7	被征服者抵杀的文明美洲考古大发现	张家梅	NLC_TOC網試用仓储	•
	8	被征服者抵杀的文明美洲等古大发现	张家梅	NLC_TOC網試用仓储	•
	9	被征服者抵杀的文明美洲考古大发现	张家梅	NLC_TOC網域用金額	•
	10	周易略例:1卷	主36	高校古文獻特色库	10

首页 上一页 下一页 末页 図 無除 □ 打印 □ 电邮 □ 下數

## **Outline**

- Principles, objectives & contents
- Project management & organization
- Project implementation
- Project achievement
- Future work

#### **Future work**

- ❖ We still have some problems. The biggest trouble is funding. The total investment on the project is limited. Funds granted to the sub-project are so low that participating libraries should spend much more money on local system, equipments etc. The prompting to the participants are very limited
- The second problem is the coexisting of 8 different local system software. Their development and different features, such as AP Interface, plug-in... bring certain difficulties to the progress of the whole project

#### **Future work**

- Keep the financial support to the RSC, ask for more fund both from sponsors and the participating libraries, to maintain the continuity and integrality of the whole project
- Choose some excellent sub-projects, give them concentrated support and build demonstration databases with excellent content and advanced technology, as the guiding design of digital library

#### **Future work**

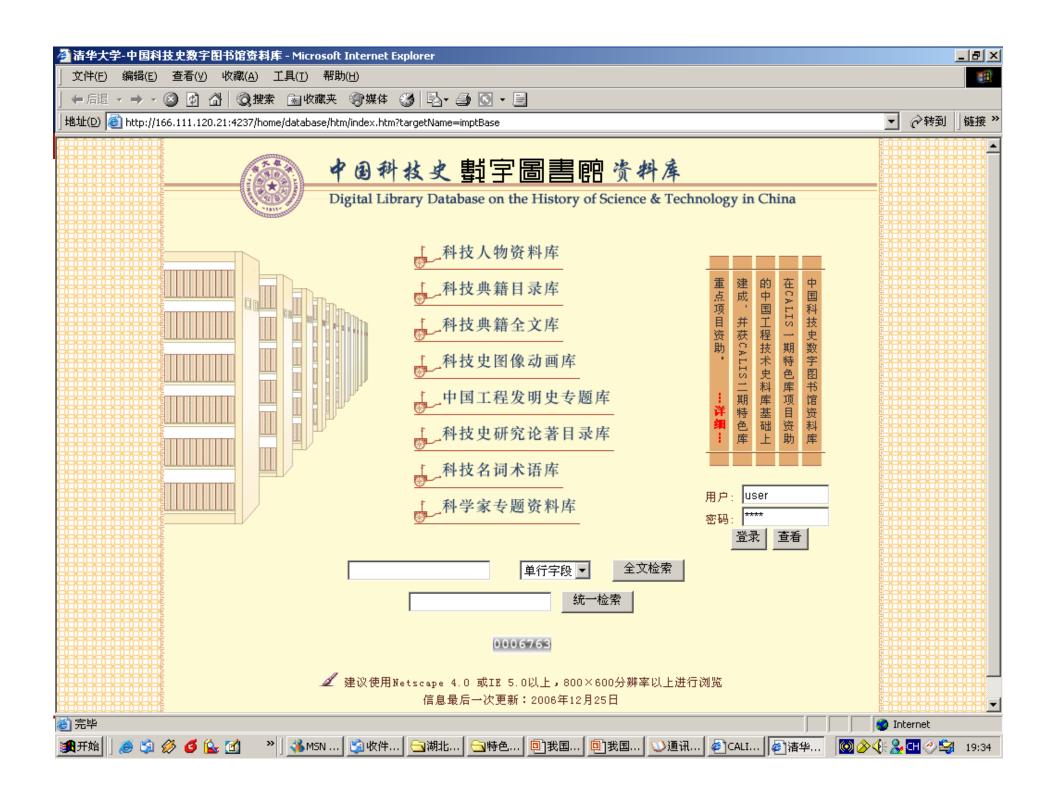
- Change supporting models, put more support emphasis on the technology, for example, provide free software to all participating libraries or at very low expense, in order to take the CALIS-RSC into sustainable development
- Discussing RSC products sharing and benefits reimbursing mechanism, establishing related service and charging system to the society.

# CALIS-RSC Sub-Project Demo

## 中国科技史数字图书馆资料库

- Digital Library Database on the History of Science & Technology in China
- By Tsinghua University Library
  - Types of resources: catalog, abstract, text, image, video, multimedia, etc
  - Divided into 8 sub-databases by subject or type of resources
  - Number of Records:127,335, 50% of which are full-texts.
- Website:

http://166.111.120.21:4237/home/database/htm

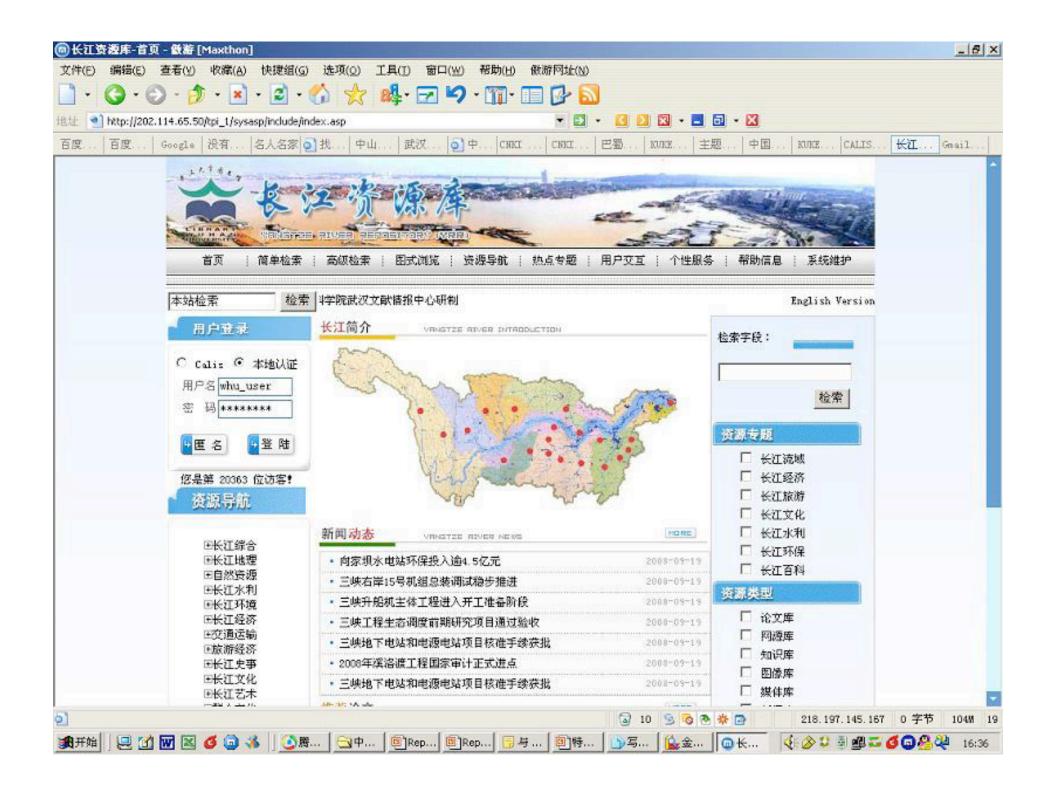


## 长江资源库

- Yangtze River Repository
- By Wuhan University Library
  - established together with the Library of the Yangtze Water Conservancy Committee and Wuhan Library of the Chinese Academy of Sciences.
  - Types of resources: catalog, abstract, text, image, video, multimedia, etc
  - Divided into 8 sub-databases by types of resources: articles, websites, knowledge, pictures, medium, news, books,index
  - Number of records:199,951, 34% of which are full- texts.

#### Website:

http://202.114.65.50/tpi\_5/sysasp/include/index.asp



# 高校古文献资源库

- Ancient Chinese Rare Book and Documentation in Academic Libraries
- By Peking University Library
  - Cooperated with libraries of Nanjing University, Beijing normal university and Sichuan University
  - Developed criteria for digitization process of ancient Chinese rare book and documentation, and developed an online cataloging system based on metadata. Number of records:201,055, 54% of which are fulltexts.

#### Website:

http://rbsc.calis.edu.cn/aopac/index.htm

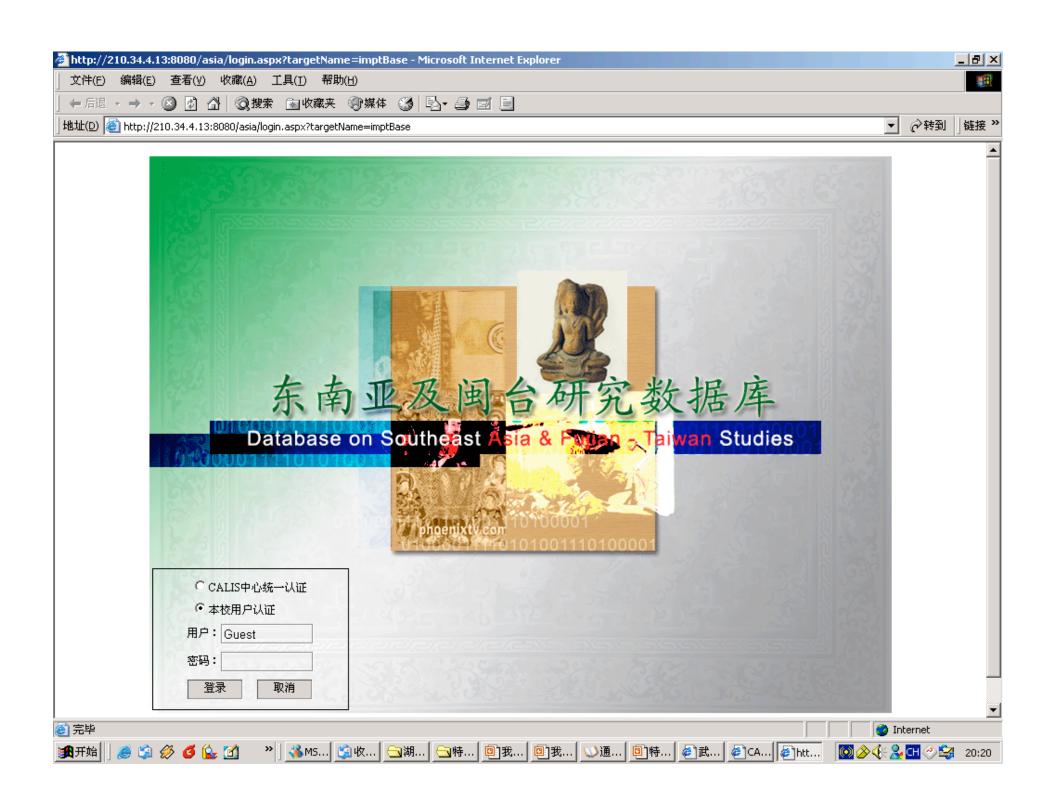


# 东南亚及闽台研究数据库

- Database on Southeast Asia & Fujian-Taiwan Studies
- By Xiamen University Library
  - Subject fields: politics, economics, society, history, culture, law, military affairs, folk-custom, clan, emigration, traditional opera, traditional arts and crafts
  - Types of resources: catalog, abstract, text, image, video, audio, etc
  - Number of records: 95,271, 51% of which are full-texts.

#### Website:

http://210.34.4.13:8080/asia/login.aspx?targetName=imptBase

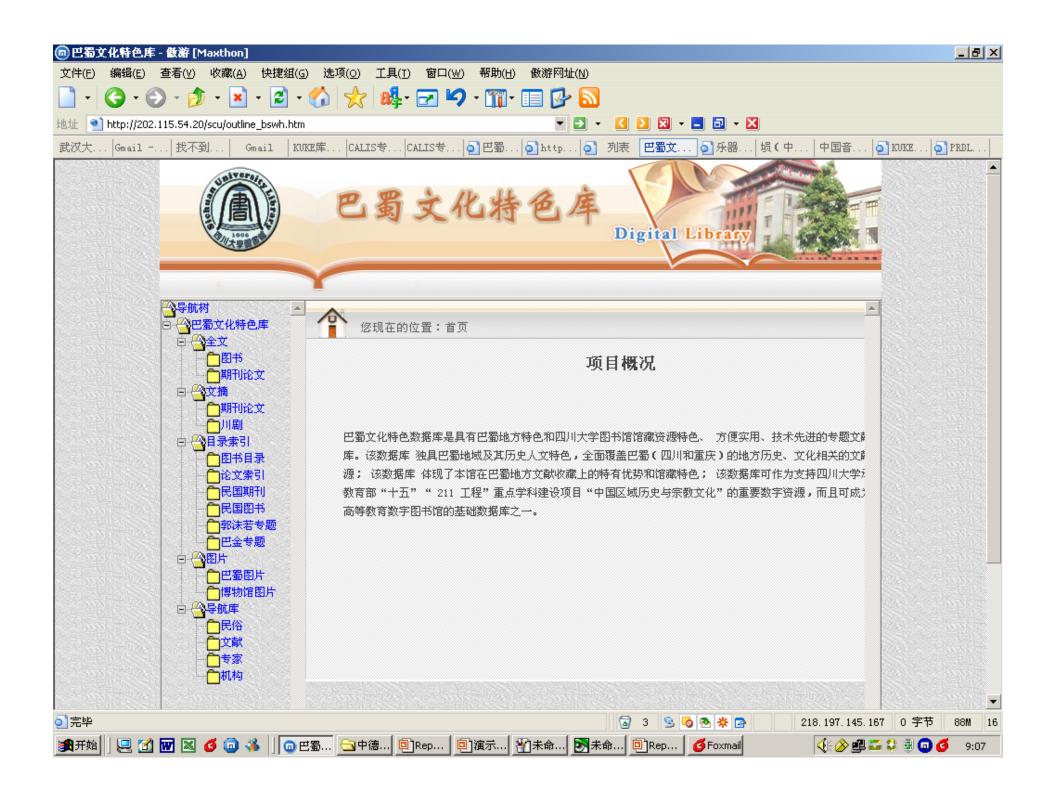


### 巴蜀文化特色库

- Special Collection on Bashu Culture
- By Sichuan University Library
  - The exclusive database on the history and culture of Sichuan province and Chongqing city.
  - Types of resources: catalog, abstract, text, image, etc
  - Number of records: 107,521, 21% of which are fulltexts.

#### Website:

http://202.115.54.20/scu/outline bswh.htm



# 书院文化数据库

#### Ancient Education System

#### By Hunan University Library

- Introduces the general pictures, histories and actuality of more than 8,000 colleges distributed in China and other Asian country, United State, Italy,etc.
- Number of records:47,000, 95% of them are full texts.

#### Website:

http://acd.hnu.cn:82/



書院文學數據庫

书院一览 书院人物 书院文献 书院揽胜 书院百科

书院是中国士人围绕着书,开展包括藏书、读书、教书、讲学、 校书、著书、刻书等各种活动,进行文化积累、创造与传播的文化教 育组织。自唐至清有千余年历史,遍布除西藏以外的全国所有省区, 总数在7000所以上。它为中国教育、学术、文化、出版、藏书等事业 的发展,对民俗风情的培植,国民思维习惯、伦常观念的养成等都作 出了重大贡献。明代开始,书院又走出国门,传到朝鲜、日本、东南 亚各国,甚至欧、美的意大利、美国等国,为中华文明的传播和当地 文化的发展作出了贡献。因此,自八十年代开始,在重振中华文化的 大势下, 其研究受到教育、历史、哲学、文化等各界重视。



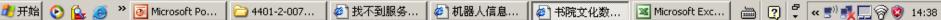




























Contact:

Liu Xia

Wuhan University Library

Te1:86-27-68754542

Email: <a href="mailto:liuxia@lib.whu.edu.cn">liuxia@lib.whu.edu.cn</a>