

题目： 兼融共享：一站式新平台检索多种中文名称规范档
作者： 何以业（香港中文大学 大学图书馆系统 编目组）
刘得光（香港中文大学 大学图书馆系统 资讯技术与策划部）
日期： 2006年9月25日

1. 前言

“中文名称规范联合协调委员会”于2003年成立，是为了推动中国大陆地区、港澳台地区和海外华人图书馆的书目规范工作；推动数字图书馆规范工作进程；并最大程度地实现资源的共建共享。该委员会的成员为包括：

- 中国国家图书馆
- 中国高等教育文献保障系统管理中心 (CALIS)
- 中文名称权威资料库 (台湾)
- 香港联校图书馆馆长咨询委员会，香港中文名称规范工作小组 (JULAC-HKCAN)

“中文名称规范联合协调委员会第三次会议” (2005年) 的其中一项会后共识决定，为了加强跨地区合作及资源共享，有需要开发一个多种中文名称规范档检索及下载之虚拟整合平台与系统。原则上由香港联校图书馆馆长咨询委员会 (JULAC) 支持，并由香港中文大学图书馆负责研发。

2. 建站方向及初期设计

2.1. 这个多种中文名称规范档检索及下载之虚拟整合平台与系统可称为一站式检索网页介面。在2003年，香港中文名称规范工作小组已经建立了一个类似的网页介面，该网页介面可以同时检索香港中文名称规范数据库 (HKCAN (Hong Kong Chinese Authority (Name))) 和美国国会图书馆规范 (Library of Congress Authorities)

2.2. 计划中的一站式检索网页介面的建站方向是扩展数据库网站范围，会包括：

- 中国国家图书馆 -- (NLC)
- 中文名称权威资料库(台湾) -- (CNAD)
- 香港中文名称规范数据库 -- (HKCAN)
- 美国国家图书馆规范档案 -- (LC)
- 下一步将加入 CALIS 的规范资料

3. 初期面对的问题

3.1. 建站初期所遇到的困难

3.1.1. 检索结果不全。检索常见的名称(如：冰心)却不是每个数据库都有回应。问题经过剖释后得悉中文字符检索时发生问题，因各数据库所使用字符集不同。请看各库所用的字符集：

- NLC : Unicode
- CNAD : Big5
- HKCAN : Unicode
- LC : Unicode

3.1.2. 各库所回应的普遍使用栏位也不多，只有款目，数量等。以下为各库回应显示页。

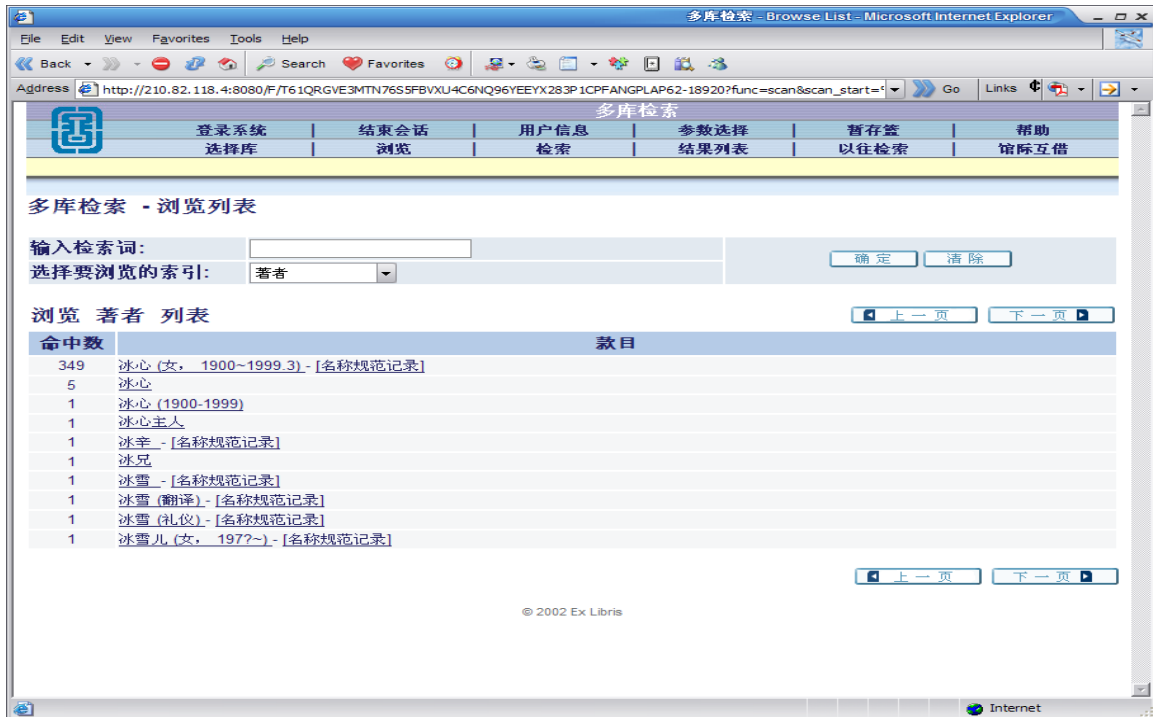


图 1：中国国家图书馆回应检索结果页



图 2：中文名称权威资料库(台湾) 回应检索结果页

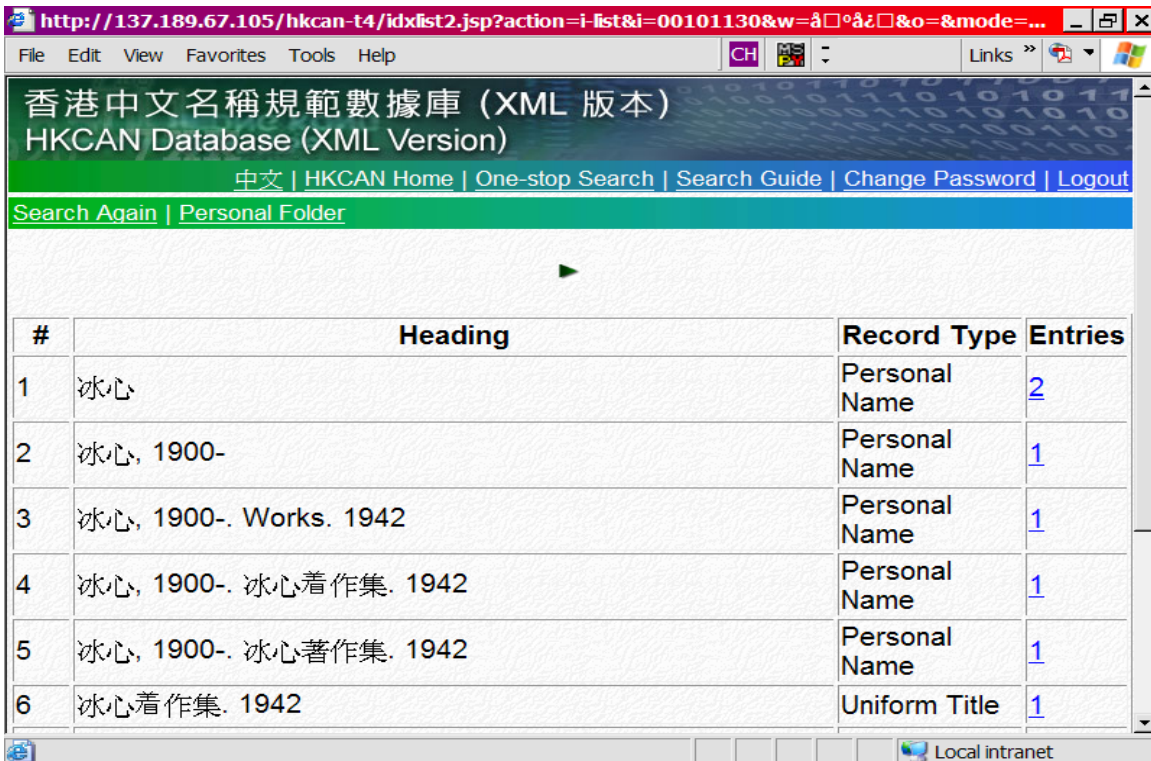


图 3：香港中文名称规范档数据库回应检索结果页

http://authorities.loc.gov - Library of Congress Authorities - Microsoft Internet Explorer

The Library of Congress >> [Go to Library of Congress Online Catalog](#)

LIBRARY OF CONGRESS AUTHORITIES

Help New Search Search History Headings List Start Over

SOURCE OF HEADINGS: Library of Congress Online Catalog
 YOU SEARCHED: Name Authority Headings = 冰心
 SEARCH RESULTS: Displaying 1 through 25 of 25.

< Previous Next >

#	Bib Records	<i>select icon in first column to...</i> View Authority Headings/References	Type of Heading
1	43	冰心, 1900-	personal name
2	1	冰昆.	personal name
3	1	冰秀.	personal name
4	1	冰莹.	personal name
5	3	冲中 纪念 成人病 研究所.	corporate name
6	2	冲中 重雄, 1902-	personal name

Done Internet

图 4：美国国会图书馆中文检索回应检索结果页，只回应书目，不是规范。

Library of Congress Authorities - Microsoft Internet Explorer

The Library of Congress >> [Go to Library of Congress Online Catalog](#)

LIBRARY OF CONGRESS AUTHORITIES

Help New Search Search History Headings List Start Over

SOURCE OF HEADINGS: Library of Congress Online Catalog
 YOU SEARCHED: Name Authority Headings = bingxin
 SEARCH RESULTS: Displaying 1 through 25 of 25.

< Previous Next >

#	Bib Records	<i>select icon in first column to...</i> View Authority Headings/References	Type of Heading
 1	60	Bingxin, 1900-	personal name
 2	0	Bingxin, Jiang	personal name
 3	0	Bingxiong, 1915-	personal name
 4	1	Bingxiu	personal name
 5	0	Bingxue, Shi.	personal name

Done Internet

图 5：美国国会图书馆拼音字符检索回应检索结果页，能显示规范

4. 一站式规范检索网站的特点

现已初步完成了这个一站式规范检索网站的测试版，初步解决了上述困难。这个测试版的特点如下。

4.1. 建立中文新功能。此网站能：

- 支援以中文字符检索
- 支援 Unicode
- ‘聪明’中文检索能力

下表列出以字符集转换及‘聪明’中文检索去解决支援中文的问题。

目标数据库	字符集	转换	‘聪明’中文检索	命中结果显示再转换
NLC	GB18030	UTF-8→ GB18030	繁体 → 简体 e.g. 蘇軾 → 苏轼	GB18030→ UTF-8
CNAD	Big5	UTF-8→ Big5	简体 → 繁体 e.g. 中国 → 中國	Big5→ UTF-8
JULAC- HKCAN	UTF-8	不转换	内含互换功能可支援 繁简体	直接输出
LC	UTF-8	不转换	复制为两个检索: e.g. 台湾 → 臺灣 臺灣 → 台湾	直接输出

4.2. 系统流程见如下的设计。

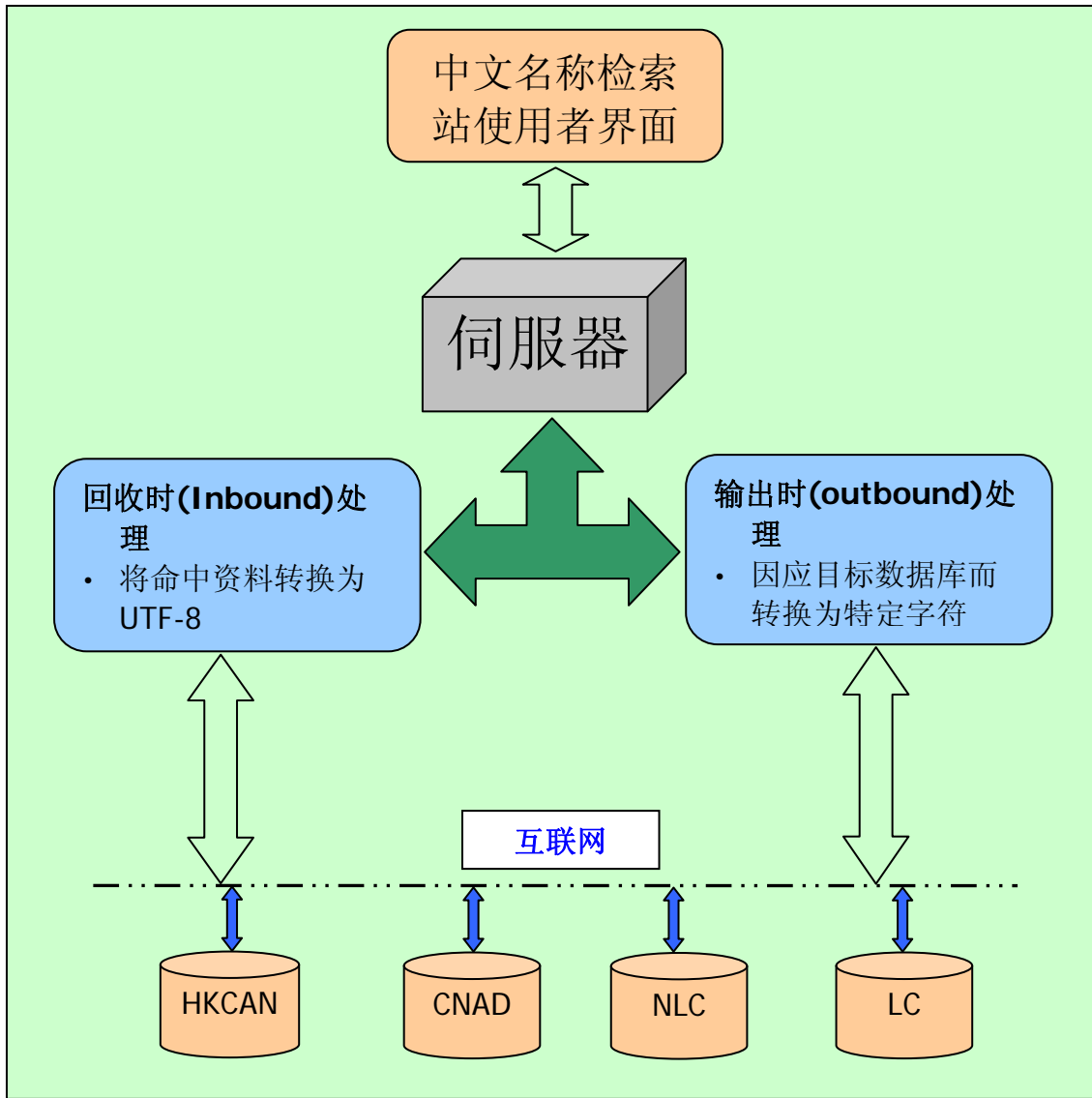


图 6：一站式规范检索网站系统流程

4.3. 测试版介面介绍

初步完成的测试版的读者介面有如下特色：

- 有三个语言介面，分别为简体中文、繁体中文、英文
- 检索中途状况页
- 汇合检索结果页

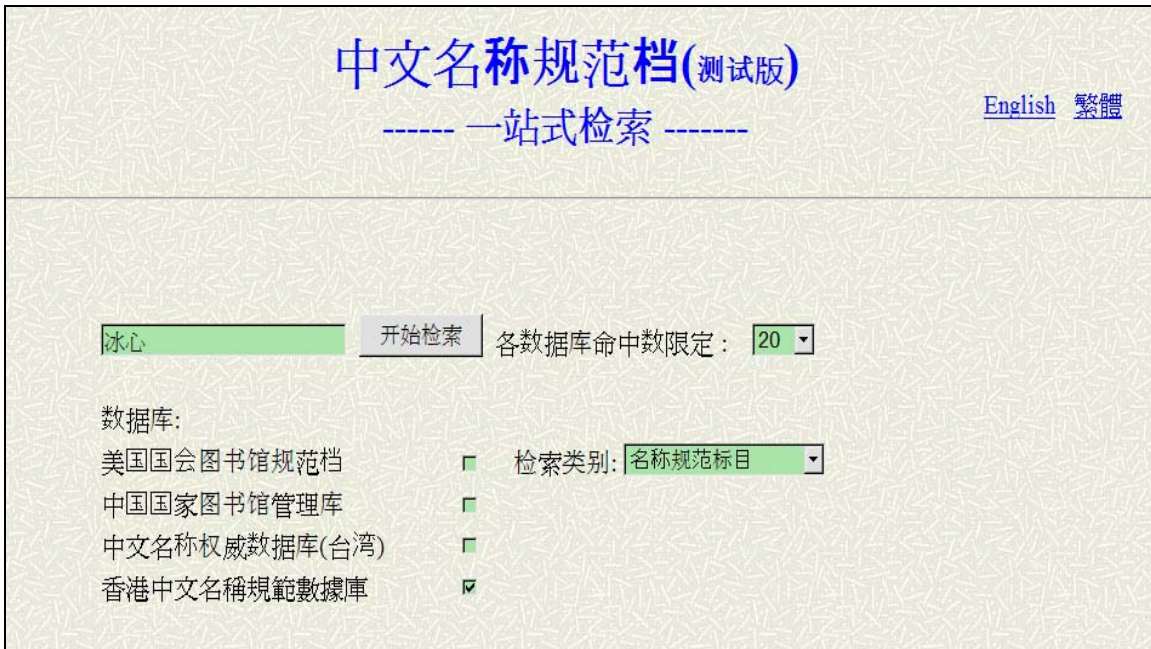


图 7：测试版主检索页



图 8：测试版检索中途状况页

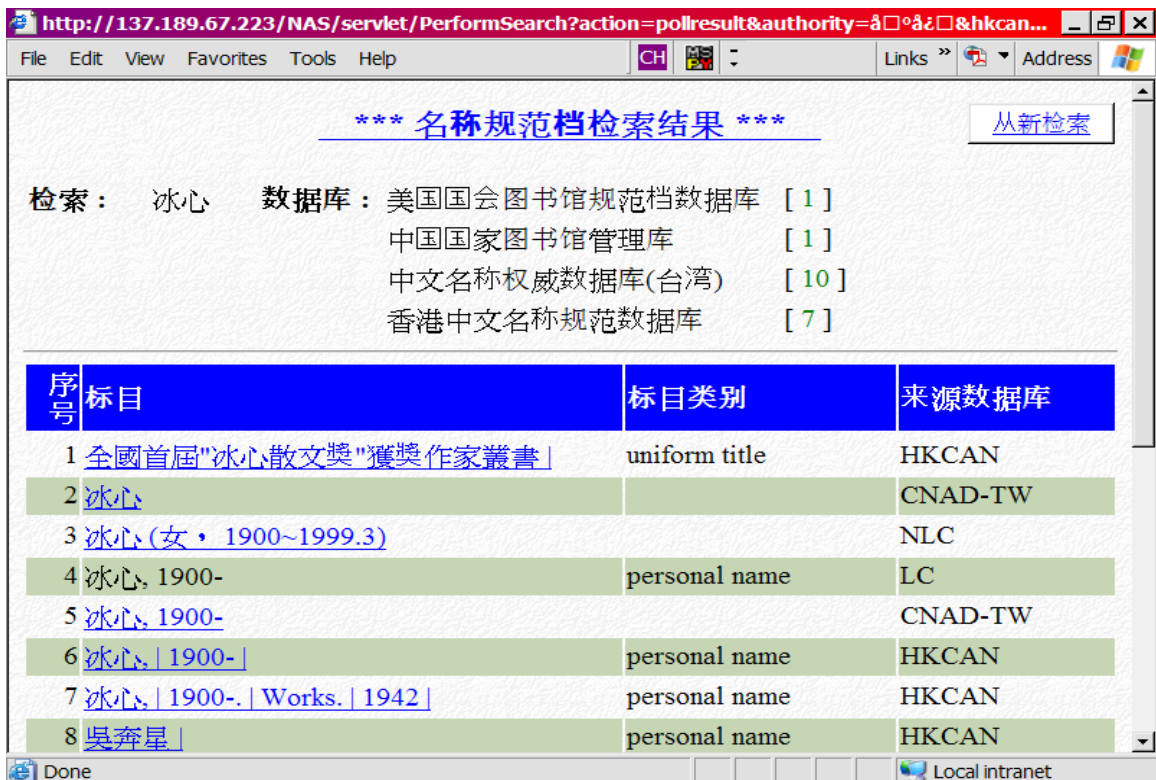


图 9：测试版中文字符汇合检索结果页

5. 挑战与展望

一站式检索中文名称规范档测试版有如下待改善之处。

5.1. 回应时间受限制。由于网络速度及能否进入网站检索均影响检索，而该等因素均受制于目标网站，因此测试版系统实无法控制及改善。解决方案是可考虑限制每站检索目录结果在 100 条目以下。

5.2. 目标网站有改动时会带来影响。各站若有任何改动，或会影响检索结果。解决方案是在系统中建立自行测试程序以确保系统运作正常及能应付可能发生的错误。

5.3. 继续完善规范数据共享的检索平台，例如下载功能，MARC 格式转换功能，精细检索功能等。