

# On the Way to Easy Discovery and Deep Integration

-----

## The Case of Peking University Library

聶華

Nie Hua

2011-11-10



北京大學圖書館  
PEKING UNIVERSITY LIBRARY

# Outline

- Background:
  - Changing Environment of Library
    - ~~ Resource, Systems, User
  - Need for Innovative & Competitive Solutions
    - ~~ What do our user want ?
- Case of PKUL
  - ~~ Evaluation & Selection for a solution of easy discovery and deep integration



# Changing Environment-Resource

## A Glance at PKUL collection development

Annual Budget: ¥ 30 millions (\$4.5 millions)

P-Resource: 6,635,181 copies

2010: Total spending: ¥ 20 millions (\$3 millions)

- Books: 55,228 Titles /118,187copies ( ¥ 8 millions)
- Journals: 5,689 Titles ( ¥ 12 millions)

E-Resource:

2010: Total spending: ¥ 10 millions (\$1.5 millions)

- Databases: 486 ( 7 of these are multi media databases )
- E-Books: 1,601,677 Titles
- E-Journals: 52,268 Titles
- Multi media: 20,019 titles 49,574 items

(含语音磁带录像带DVD/CD/随书盘)



北京大学图书馆  
PEKING UNIVERSITY LIBRARY

# Changing Environment-Resource

## PKUL collection development statistics 2005-2010

|       | 2005          | 2006          | 2007          | 2008          | 2009          | 2010          |
|-------|---------------|---------------|---------------|---------------|---------------|---------------|
| 图书    | 7,187,552.85  | 6,741,213.79  | 7,950,428.79  | 7,602,348.72  | 10,412,903.17 | 7,991,009.02  |
| 期刊    | 12,375,902.00 | 12,352,462.00 | 15,584,042.40 | 13,667,389.40 | 11,118,920.55 | 11,952,289.57 |
| 电子资源  | 4,540,000.00  | 6,450,000.00  | 8,353,162.00  | 6,353,704.00  | 7,991,269.00  | 8,839,895.00  |
| 多媒体资料 | 15,861.90     | 8,836.00      | 728,000.00    | 400,000.00    | 633,000.00    | 788,000.00    |
| Total | 24,119,316.75 | 25,552,511.79 | 32,615,633.19 | 28,023,442.12 | 30,156,092.72 | 29,571,283.59 |

Changes in last 6 years:

- P-Resource: minor decrease
- E-Resource: increase 10%
- Multi-media: increase 37%

Usage of E-Resource of 2010 to 2009

- number of search: increase 34%
- number of full-text download: increase 34%
- multi-media: increase 37%

# User's Need for: Deep Integration of Library Resource



significant level of frustration of discovery and access to library resource, esp. the full text of scholarly journal articles:

- Scattered as well as confusing and badly designed portals, slow federated search engines, and so on can make for a negative user experience.
- What our user need is a deep integration of those scattered library



[【返回】](#)[【限制检索】](#)[【新的检索】](#)[【前一页】](#)[【后一页】](#)[【跳至](#)

检索结果 --

122 找到你检索的记录 "global warming". 查看 1 到 20.  
使用校验框 ☐ 为打印 抓取标记馆藏清单.

#1 ☐ P467/R634

[查看](#)

Adaptation to climate change : a spatial challenge  
Roggema, Rob.

#2 ☐ P/Ad95s

[查看](#)

Advances in earth science : from earthquakes to global  
Sammonds, P. R. (Peter R.)

#3 ☐ X16/As41

[查看](#)

Asia, looking ahead : initial stages of national communica  
Ramakrishna, Kilaparti.

#4 ☐ Q16/B52

[查看](#)

The biodiversity crisis : losing what counts  
Novacek, Michael J.

#5 ☐ D771.222/B964

[查看](#)

Business and environmental policy : corporate interests



## Users expect and demand:

- same ease of use that they experience on freely available websites such as Google or Amazon.
- current, reliable, and pertinent scholarly information quickly and consistently through a simple search box interface

# 12 studies by OCLC in partnership with JISC:

to better understand information-seeking behaviors of library users  
to review the issues associated with the development of information services  
and systems that will best meet users' needs.

A synthesis of findings from these major user studies points toward a number of implications for libraries, including:

(1) **Library systems** must do better at providing seamless access to resources

无缝接入

(2) Librarians must increasingly consider a greater variety of digital formats and content  
更多种类型的数字格式和内容

(3) **Library systems and content** must be prepared for changing user behaviors  
图书馆系统和内容必须随需应变

(4) **Library systems** need to look and function more like search engines, i.e., Google and Yahoo, and Web services, i.e., Amazon, since these are familiar to users who are comfortable and confident in using them.

(5) **Information Seeker: Report of Findings From Selected OCLC, RIN and JISC User Behavior Studies**  
by Lynn Silipigni Connaway, Ph.D. and Timothy J. Dickey, Ph.D., OCLC Research,  
was funded by JISC, and was presented during the JISC Annual Conference in London in 2004.

# Call for a PKUL discovery & Integration Solution

- A single interface (一个唯一的界面)
- A comprehensive gateway to library resource that PKUL owned, subscribed, and created . (一个实现图书馆全部馆藏信息资源组成的整合的发现与获取)
- User Experience similar to famous web sites ( 读者习惯的于哪些著名检索站点相似的风格和成熟度)

=>“Google Like”:

Single, Comprehensive and

easy



北京大学图书馆  
PEKING UNIVERSITY LIBRARY



# Industry Innovations:

## Last Decade:

- Electronic Resource Integration: A-Z, Meta Search based federated search; Open URL based Linking Server

## This Decade:

- Next Generation Interface
- Unified Resource Discovery: Index of Metadata based

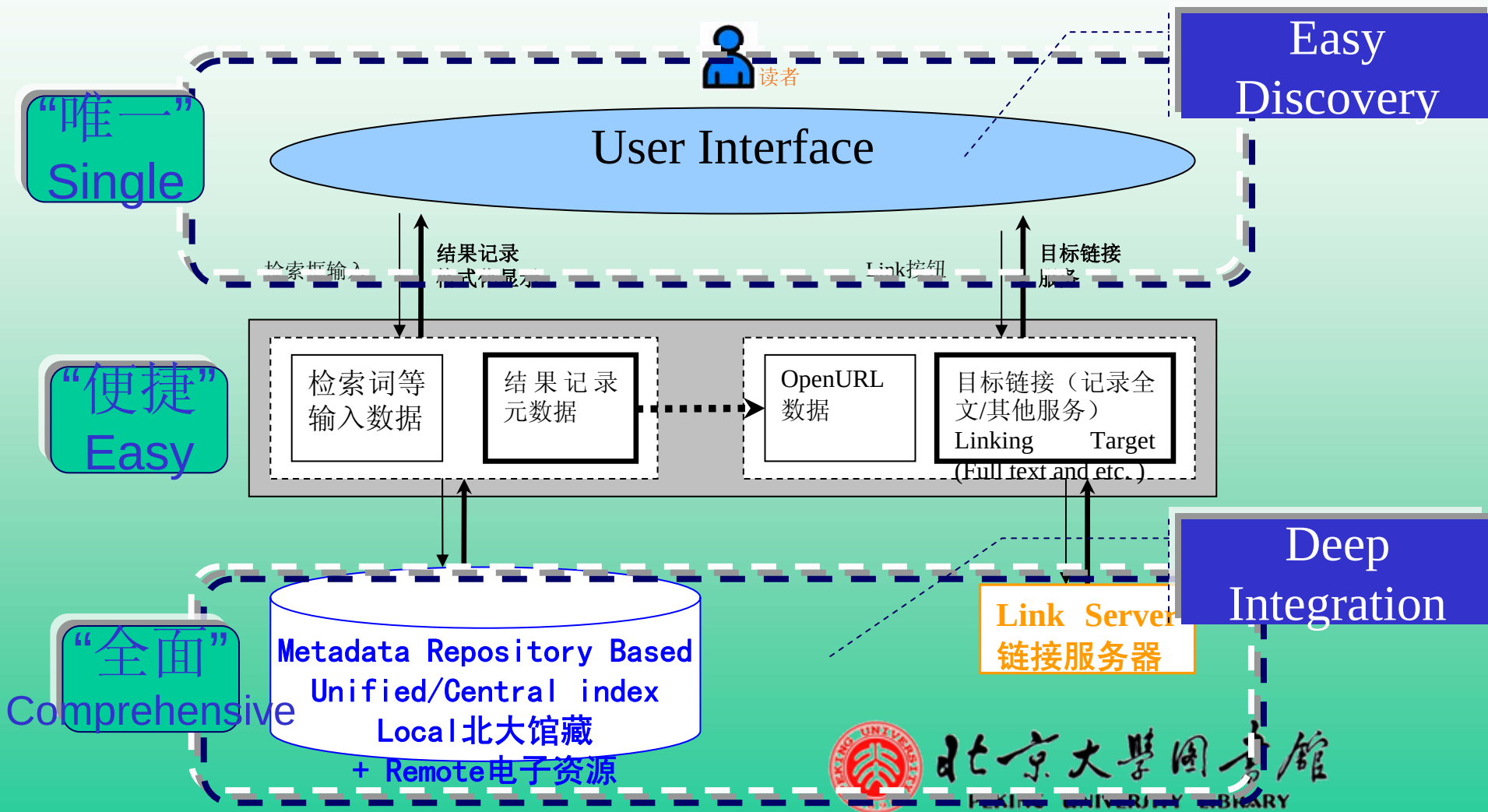
## Next Decade:

- Next Generation Library System



北京大学图书馆  
BEIJING UNIVERSITY LIBRARY

# Peking University Library's Solution on integration, discovery & Delivery



乱云飞渡仍从容 -

## Easy Discovery & Deep Integration

| 商家 | Exlibris      | Serials Solutions    | EBSCO             | 其它 |
|----|---------------|----------------------|-------------------|----|
| 产品 | Primo V1, V3  | Summon               | Discovery Service |    |
|    | Primo Central | Summon Unified Index | Discovery Service |    |
|    | SFX           | 360 Link             |                   |    |
|    | Metalib       | 360 Search           |                   |    |

# Unified Central Index based Web Scale Discovery: Advantages

## 基于元数据仓储索引的发现系统: 显著优势

“Single”: not only interface but also an architecture, Library's Google

“Comprehensive”: Coverage, Content, quality of the resource  
(资源的覆盖范围、内容和质量)

- library resource: P+E; Catalog+full text; Local+Remote
- **Multi source/format/elements** (多来源, 多格式, 多元素)
- Authority, Current, reliability (规范性, 时效性, 准确性)

“Easy”: Search, Display, Sorting (检索; 显示; 排序)

- 更快更强更深的检索 (faster, deeper, and more powerful Search)
- 更优化更人性化更可视化的显示 (better, more user friendly and visualized display)
- 更精准更可靠更优化的结果集相关度排序 (more accurate)



北京大学图书馆  
PEKING UNIVERSITY LIBRARY

# What we did with the trials

- Coverage Alignment for **Databases and E-Journals** (ISSN), Apr.-May, 2010
- **CNMARC** mapping and loading, July, 2010
- Summon **Chinese search and interface** beta version, Aug. 31st, 2010
- 1st round library trial, Sept.-Oct. 2010
- Refinement for content types, metadata mapping for CNMARC & **local digital contents, RTA**, Oct.-Nov., 2010
- Coverage Comparison between Summon and others for databases and E-journals: Nov.-Dec., 2010
- 2nd round library trial: Dec., 2010





# Major Focuses and Arrangements

|                     | Interface         | Searching              | Personalized Functions | Systems Reference              |
|---------------------|-------------------|------------------------|------------------------|--------------------------------|
| Interface Searching | Library Catalog   | Digital Content        | Commercial Databases   | Systems Collection Development |
| Resources           | RTA Unicorn       | Authentication         | Commercial Databases   | Systems                        |
| Local Systems       | Chinese Searching | Simplified Traditional | Character Codes        | Systems                        |
| Localization        | SaaS Mirror site? | Locally host           |                        |                                |
| Service pattern     |                   |                        |                        |                                |
| Price               |                   |                        |                        |                                |

# Criteria for Evaluation and Selection of a Good Discovery Service: Architecture & Searching

- Architecture:
  - It is not only interface , but a framework for knowledge discovery
  - Clean, simply, modern? Single box VS. Scope (s) ? (pre-defined resources scopes)

不仅是界面，而是发现框架，基于对读者需求的求证和确定
- Facets Navigation: Aligned positively with the quality and enrichment of metadata （与元数据的质量和丰富性成正相关）
  - Facets: content type, subject terms, author, time period, library location, language, full-text filter, peer-reviewed filter, if exclude newspaper article or add results beyond library's collection, etc.

分面：资源类型、主题、作者、时间、分馆、语言、在线全文、同行评审、是否排除报纸、是否扩展到本馆馆藏之外等

  - Is locally customized necessary and practical?



- Sorting: Optimized relevancy algorithm (优化相关度算法)
  - What to match on? subject headings from controlled vocabularies, article titles, author keywords, keywords within abstracts, keywords within full text.....
  - Chinese Relevancy Sorting: **a more difficult assignment?**
- Searching String: Auto-Correction, auto completion, etc.  
(纠错, 补全, 人名识别等)
- Full text access:
  - % of full-text indexing (全文索引比例)
  - Support for open URL linking (支持开放链接)
- Visualization:
  - Aquabrowser's Word cloud

# Criteria for Evaluation and Selection of a Good Discovery Service :

## Content: Metadata & Coverage

Metadata: core of Unified or Central Index, and the key to the resources' integration, discoverabilities & accessibilities

直接关系到资源的可整合性，可发现性和可获得性

- from the content provider, direct, reliable, and stable  
直接、可靠、稳定

- Refinement: de-duplicated, correction, merging & normalization 查重，归并，修正，聚类等规范化处理

- Enrichment: Cover page(books), No. of cited(article), peer reviewed status, taxonomies....

增强的元数据元素和内容，封面（书目），被引次数（文章），同行评议（期刊），多种分类法等

- Ensured regular & timely updating frequency

有规律有保障的及时的更新频次



北京大学图书馆  
PEKING UNIVERSITY LIBRARY

# Criteria for Evaluation and Selection of a Good Discovery Service :

## Electronic Resource

- Objectives: evaluations based on the data from Vendors:
  - No. of Signed publishers & content providers (签约出版商)
  - No. of Journals indexed (article level) (全文文章级期刊数)
  - No. of Indexed metadata (索引元数据)
- Subjective: Coverage comparison analysis on PKUL's holdings
  - 134 foreign language databases (外文数据库)
  - 27973 titles duplicated based ISSN (英文期刊, ISSN去重)
  - % of Journal coverage: 期刊覆盖率
    - High: 19444 titles, 95.2%



# Criteria for Evaluation and Selection of a Good Discovery Service: Local Digital Contents

- Data Loading: 支持多种数据接口  
OAI-PMH, FTP等
- Data Formats: 处理多种数据格式  
MARC, XML, Excel, ascii, etc.
- Character Codes: GB,GBK, BIG5, etc.
- Deep Integrated display (深度整合融会显示)
- Real Time Availability for ILS:



# Criteria for Evaluation and Selection of a Good Discovery Service: Localization & Customization

- Chinese:
  - Searching: eg: Segmentation based Character string matching?
  - Pinyin: search, sort and etc.
  - Chinese Relevancy boost
  - Dually Simplified and Traditional Chinese Character Mapping
- Chinese Commercial contents coverage  
中文电子资源内容的覆盖
- Strong Local Technical Support  
有力的本地技术支持





# Summon PKUL Launched:

The screenshot displays the PKU Library website with a blue header and a white sidebar. The main content area features a search bar, navigation links, and a list of resources. The sidebar on the left contains a '学科导航' (Subject Navigation) section with a list of disciplines. The main area includes a '未名学术搜索' (Wuming Academic Search) section with a search bar and a '开始' (Start) button. Below this is a '书刊目录检索' (Book/Journal Catalog Search) section with a search bar and a '开始' (Start) button. Further down is a '全文统一检索' (Full-text Unified Search) section with a search bar and a '开始' (Start) button. The right side of the page features a '留言板' (Message Board) section with a '留言簿' (Message Book) button and a '最新消息' (Latest News) section with a list of news items. At the bottom right is a '最新资源报道' (Latest Resource Report) section with a list of resource reports. The top right corner has a language selector set to 'English' and a login section with a '校园卡 / 图书证号' (Campus Card / Library Card Number) field and '登录' (Login) and '修改密码' (Change Password) buttons.

English

校园卡 / 图书证号:   [登录](#) [修改密码](#)

本馆概况 开馆时间 馆藏分布 借阅规则 校外读者 规章制度 帮助 联系我们

学科导航

- 所有学科
- 社会科学
- 哲学
- 管理学
- 政治学
- 法学
- 经济学
- 教育/文化
- 图书档案学
- 新闻传播
- 语言文学
- 艺术
- 历史学
- 自然科学
- 数学
- 物理学/力学
- 化学
- 地球科学/天文学
- 生物学
- 医学
- 农业
- 计算机/网络
- 工业技术

未名学术搜索 (试运行) **NEW**

图书,期刊,论文...一搜即得 [开始](#)

书刊目录检索

[开始](#)

全文统一检索

[开始](#)

常用资源

|         |          |
|---------|----------|
| 图书/电子图书 | 期刊/电子期刊  |
| 数据库     | 学位论文     |
| 教学参考书   | 多媒体资源    |
| 网络资源    | 电子书/随书光盘 |

咨询台

即时问答 | 邮件咨询 | 常见问题 | 新生空间 **New!**

我的借阅帐户

借阅状态 | 续借 | 推荐购买 | 修改邮箱 | 借阅历史 **New!**

新书导读

最新上架图书... | 名师推荐 | 教师分享 **New!**

一小时讲座

电子资源利用 | 网络培训 | 检索课 | 专场讲座

馆际互借

最新消息 | 提交申请 | 优惠政策

研究咨询

课题咨询 | 论文收录及引用 | 查新 | 个人参考文献

学科馆员

学科个性化服务 | 图书馆视界 **New!**

分馆

留言板

最新消息

- 未名学术搜索试用 **New!**
- 免费征订国外优秀 **New!**
- 纪念辛亥革命展览 **New!**
- 欢迎推荐中文报刊 **New!**
- 用电纸书体验经典 **New!**

更多..

最新资源报道

- 刘岷渭音乐导聆 **New!**
- ASP学术视频在线数 **New!**
- Taylor & Francis... **New!**
- Cairn法语数据库 **New!**

更多..

[本站索引](#)



霸王别姬

搜索館藏資源

☐ 繼續精確搜索

☒ 重新搜索

搜索結果: 搜索霸王别姬返回 12 項結果

### 精確搜索

- ☐ 只顯示有全文的結果
- ☐ 只限於來源於學術性出版物的文章，包括同行評議
- ☐ 不包括報紙文章
- ☐ 添加您的圖書館館藏之外的結

### ▼ 內容類型

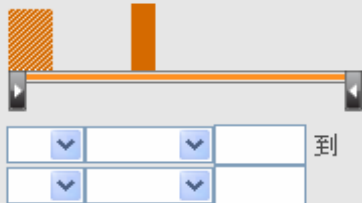
- ☒ 任何
- ☐ 專輯 (6)
- ☐ 書籍 (4)
- ☐ 電子書 (2)
- [更多選項](#)

### ▼ 主題詞

- ☒ 任何
- ☐ 电影评论 (2)
- ☐ 长篇小说 (2)
- ☐ 京剧 (1)
- ☐ 京剧 剧本 (1)
- [更多選項](#)

### ▼ 出版日期

任何



### 后窗看电影



電子書

作者: 卫西谛, 黄小璐, 徐震  
版次: 1 ISBN: 7563350047 日期: 2005  
主题: 电影评论  
圖書編目號碼: J905.1  
**獲取渠道:** 線上全文

### 绚烂与平淡 影片分析教程



電子書

作者: 黄琳, 杨尚鸿  
版次: 1 ISBN: 9787562435334 日期: 2005  
主题: 电影评论  
圖書編目號碼: J905.1  
**獲取渠道:** 線上全文

### 霸王别姬青蛇



書籍

作者: 李碧华  
ISBN: 7536035020 日期: 2001  
主题: 长篇小说  
圖書編目號碼: I247.5/1751f  
**獲取渠道:** [Check Availability](#)

### 霸王别姬



書籍

作者: 李碧华  
ISBN: 7020017290 日期: 1993  
主题: 长篇小说  
圖書編目號碼: I247.5/1751  
**獲取渠道:** 中文图书, I247.5/1751

### 京剧大观

團體作者: 北京宝文堂书店编辑部

日期 (最新)

Support Eng.  
S. & T Chinese

Sort by Date  
Or Relevance

# What's Next?

- Problems shooting
- Improvement for everything
- Local digital content
- Chinese content
- Chinese Localization
- Go ahead and .....enjoy!



Thank you!

[hnie@lib.pku.edu.cn](mailto:hnie@lib.pku.edu.cn)

