# OAI Pilot Project

## PRDLA Archive, 2007

**David Palmer**

**University of Hong Kong**

環太平洋數字圖書館聯盟簡稱 -- 태평양지역 전자도서관협의체 -- **環太平洋電子図書館会議**

# OAI Pilot Project
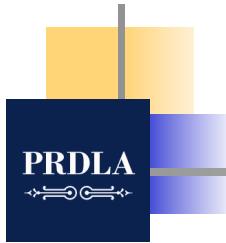
- Project Team & Chair
- HKU to host "PRDLA Archive"
- Harvest from 24 members
- OAI-PMH & DSpace platform
- Project: Apr 2006 ~ Mar 2007

HKUL > PRDLA Archive

Bibliotecas
Libraries 도서관 図書館
Pacific Rim Digital Library Alliance

# PRDLA Archive

Home / Search

Twenty-five academic libraries surrounding the Pacific are cooperating with each other to leverage the use of their digital resources. Metadata from member library local collections has been harvested and archived here to increase discovery and access, and to allow one search across all data.

Browse by:

• Authors

• Titles

• Dates

• Communities
  & Collections

PRDLA Home

Help

**Search:** All of PRDLA Archive

| Search type: | Search for: |
| --- | --- |
| Keyword | |

| AND | Keyword | |
| AND | Keyword | |

Search        Clear

Sponsored by HKU Libraries

Done        Internet

Start        PRDL...        16:12
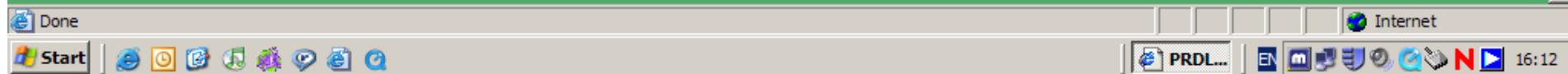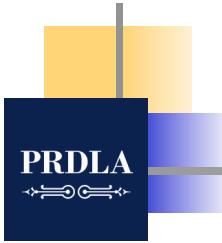
# DSpace Problems

- Slow browsing on 300,000 records
- Scaling would mean clustering of servers
- Problems in OAI harvesting & updating

# CDS Invenio

- CDS Invenio, CERN Document Server
  - http://invenio.cern.ch/
- Solved DSpace problems
  - Fast searching ~ 1,500,000 records
- GNU s/w, free license
- Easier to customize
- Multilanguage interface

# Harvesting, Loading, Mapping
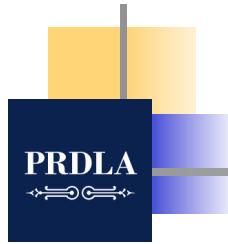
Remote Site ➔ XML file ➔ Convert ➔ Loader

Database

_____

Native ➔ oai_dc ➔ MARCXML ➔ MARC ➔ Invenio

Format                                                  Field

環太平洋數字圖書館聯盟簡稱 -- 태평양지역 전자도서관협의체 -- 環太平洋電子図書館会議

```xml
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/
      http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
 <responseDate>2007-09-07T03:11:02Z</responseDate>
 <request verb="ListRecords" metadataPrefix="oai_dc">http://
digicoll.manoa.hawaii.edu/sos/XML/oai/index.php</request>
 <ListRecords>
  <record>
   <header>
    <identifier>oai:streetprint.org:digicoll.manoa.hawaii.edu-sos:http://
digicoll.manoa.hawaii.edu/sos/Pages/viewtext.php?
s=browse&amp;by=title&amp;route=browseby.php&amp;tid=2</identifier>
    <datestamp>2007-06-28T16:15:04Z</datestamp>
   </header>
   <metadata>
    <oai_dc:dc
      xmlns:oai_dc="http://www.openarchives.org/OAI/2.0/oai_dc/"
      xmlns:dc="http://purl.org/dc/elements/1.1/"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai_dc/
      http://www.openarchives.org/OAI/2.0/oai_dc.xsd">
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<collection>
  <record>
    <datafield tag="041" ind1=" " ind2=" ">
      <subfield code="a">en</subfield>
    </datafield>
    <datafield tag="100" ind1=" " ind2=" ">
      <subfield code="a">Kelly, John</subfield>
    </datafield>
    <datafield tag="245" ind1=" " ind2=" ">
      <subfield code="a">A Brief on John Kelly and Save Our Surf</subfield>
    </datafield>
    <datafield tag="260" ind1=" " ind2=" ">
      <subfield code="b"></subfield>
    </datafield>
    <datafield tag="520" ind1=" " ind2=" ">
      <subfield code="a">'A Brief on John Kelly and Save Our Surf' by Kelly, John. 2
images. Has additional notes.</subfield>
    </datafield>
    <datafield tag="650" ind1="1" ind2="7">
      <subfield code="a">article</subfield>
    </datafield>
    <datafield tag="650" ind1="1" ind2="7">
      <subfield code="a">Hawaii -- History -- 1959-</subfield>
```

```
001__ 293121
041__ $$aen
100__ $$aKelly, John
245__ $$aA Brief on John Kelly and Save Our Surf
260__ $$b
300__ $$aimage/jpeg
520__ $$a'A Brief on John Kelly and Save Our Surf' by Kelly, John. 2 images. 65017 $
$aarticle
65017 $$aHawaii -- History -- 1959-
6557_ $$aText
720__ $$aStreetprint Admin
8564_ $$uhttp://digicoll.manoa.hawaii.edu/sos/Pages/viewtext.php?
s=browse&by=title&route=browseby.php&tid=2
8564_ $$qhttp://digicoll.manoa.hawaii.edu/sos/Libraries/imagelibrary/
2-265thumbnail.gif$$xicon
909CO $$uoai:streetprint.org:digicoll.manoa.hawaii.edu-sos:http://
digicoll.manoa.hawaii.edu/sos/Pages/viewtext.php?
s=browse&by=title&route=browseby.php&tid=2
970__ $$aoai:streetprint.org:digicoll.manoa.hawaii.edu-sos:http://
digicoll.manoa.hawaii.edu/sos/Pages/viewtext.php?
s=browse&by=title&route=browseby.php&tid=2
980__ $$aSave Our Surf (SOS)
```

http://prdlaarchive.lib.hku.hk/

Google

Bookmarks ▾  61 blocked  Check ▾  AutoLink ▾  AutoFill  Send to ▾  Settings ▾

PRDLA Archive: Home  ×  CDS Invenio: Overview

Page ▾  Tools ▾

Bibliotecas
Libraries 도서관 図書館
Pacific Rim Digital Library Alliance

# PRDLA Archive

**Search 274,407 records for:**

| any field ▾ | **Search** | **Browse** |

Search Help :: Advanced Search

**Pre-defined searches:**

Pacific Exploration
Chinese diaspora
Minority groups in China

PRDLA Archive :: Search :: Help
Powered by CDS Invenio v0.93.50.20070910
Maintained by systems@lib.hku.hk
Last updated: 07 Sep 2007, 11:57

This site is also available in the following languages:
Български Català Česky Deutsch Ελληνικά English Español Français Hrvatski Italiano
日本語 Norsk/Bokmål Polski Português Русский Slovensky Svenska Українська 中文
(简) 中文(繁)

Done  Internet  100%

http://prdlaarchive.lib.hku.hk/search?sc=1&p=exploration+or+discover*+or+expedition*+and+Pacific&f=&action_search=S

Google [ roar repository ] Go | Bookmarks | 61 blocked | Check | AutoLink | AutoFill | Send to | Settings

PRDLA Archive: Search R... | CDS Invenio: Overview | LibQUAL+(TM) Survey Preview | Page | Tools

**Bibliotecas**
**Libraries** 도서관 図書館
Pacific Rim Digital Library Alliance

Home > Search Results

# PRDLA Archive

**Search:**

exploration or discover* or expedition | any field | **Search** | **Browse**

Search Tips :: Advanced Search

**Search collections:**

*** any collection ***

**Sort by:** | **Display results:** | **Output format:**

latest first | desc. | - or rank by - | 10 results | split by collection | HTML brief

**Results overview:** Found **108** records in 0.03 seconds.

**University of Washington Libraries Digital Collections**, **94** records found
**University of Hong Kong Libraries** , **1** records found
**University of Southern California**, **13** records found

**University of Washington Libraries Digital Collections**          **94** records found  1 - 10 ▶ ▶▶  jump to record: 1

☐ 1.
**U.S. Revenue Cutter THETIS, n.d.**
n.d.

Internet                    100%

**University of Southern California**       13 records found  1 - 10 ▶  jump to record: 1

☐ 1.

**Lithograph by William McMautrie, showing abandoned ships in the San Francisco harbor, April 1, 1850**

*Abstract*: Photograph of a lithograph by William McMautrie, showing abandoned ships in the San Francisco harbor, April 1, 1850. [...]

Detailed record - Similar records

☐ 2.

**View of Rindge Ranch in Malibu from a rise above the Pacific Coast looking south across an estuary, ca.1935**

*Abstract*: Photograph of a view of Rindge Ranch in Malibu from a rise above the Pacific Coast looking south across an estuary, ca. [...]

Detailed record - Similar records

☐ 3.

PRDLA Archive: Record#172526: Lithograph by William McMautrie, showing abandoned ships in the S - Windows Internet Explorer

http://prdlaarchive.lib.hku.hk/record/172526    Google

Google  roar repository  Go  Bookmarks  61 blocked  Check  Settings
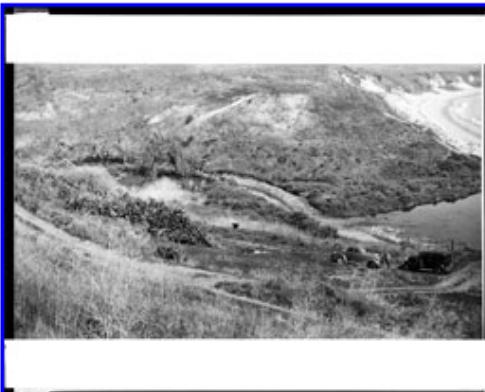
PRDLA Archive: Record#... ✕   CDS Invenio: Overview   LibQUAL +(TM) Survey Preview    Page  Tools

**Information**  References  Discussion  **Usage statistics**  Fulltext

/ Buildings

# Lithograph by William McMautrie, showing abandoned ships in the San Francisco harbor, April 1, 1850

Photographer: Pierce, C.C. ; McMautrie, William B.

## Caption

Photograph of a lithograph by William McMautrie, showing abandoned ships in the San Francisco harbor, April 1, 1850. The harbor is shown at left, cluttered with unmanned sailing vessels. To the right, the city is shown, with the most commercial area by the roads along the shore. In the distance, hills can be seen, while in the foreground, a man can be seen to the left and a few goats to the right along the ridge of Telegraph Hill. Picture file card identifies McMautrie as a draughtsman for the U.S. surveying expedition and reads "From left to right: Clark Point, Rincon Point, Happy Valley, Long Wharf, Pacific Wharf, Appolo Warehouse, Niantic Warehouse, Sampson Street and Portsmith Square".

## Resources



© PRDLA Archive

Record created 2007-04-29, last modified 2007-04-29

Similar records

(external link):

CHS-8981cis_1024_lres.jpg

➡ Export as BibTeX, MARC, MARCXML, DC, EndNote, NLM

Bibliotecas
Libraries 도서관 図書館
Pacific Rim Digital Library Alliance

# PRDLA Archive

**Search 274,407 records for:**

| All of the words: ▼ | | any field ▼ | AND ▼ |
| All of the words: ▼ | | any field ▼ | AND ▼ |
| All of the words: ▼ | | any field ▼ | **Search**  **Browse** |

Search Help :: Simple Search

**Added/modified since:**                      **until:**

| Added since: ▼ | any day ▼ | any month ▼ | any year ▼ | any day ▼ | any month ▼ | any year ▼ |

| **Sort by:** | | **Display results:** | | **Output format:** |
| - latest first - ▼ | asc. ▼ | - or rank by - ▼ | 10 results ▼ | split by collection ▼ | HTML brief ▼ |

**Narrow by collection:**

☑ **Beijing University Library** (19)
   Famous Scholar Database (19)

☑ **University of Hawaii at Manoa Libraries** (3,459)
   Hawaii War Records - U.S. Army Signal Corps Photographs (880) Save Our Surf (SOS) ( Steve Thomas Traditional Micronesian Navigation Collection (2,086) TRAIL-Technical Report Archive and Image Library (167) Hawaiian Photo Album (40)

☑ **National University of Singapore Library** (293)
   Technical Reports (293)

☑ **University of Auckland** (5,845)

Advanced Search

Bibliotecas
Libraries 도서관 図書館
Pacific Rim Digital Library Alliance

# 环太平洋数字图书馆联盟

关于这个网站

为了共建共享电子资源，位于环太平洋区内的二十五所学术图书馆合作组成PRDLA，元数据均来自各图书馆并典藏于香港大学图书馆，供读者一次过检索所需数据。

参见

PRDLA
CDS Invenio
CERN
Google

搜寻 274,407 笔记录:

| 包含以下全部的字词: ▾ | | 任何字段 ▾ | 和 ▾ |
| 包含以下全部的字词: ▾ | | 任何字段 ▾ | 和 ▾ |
| 包含以下全部的字词: ▾ | | 任何字段 ▾ | 搜寻 浏览 |

搜索帮助 :: 初级检索

**Added/modified since:** 至:

记录日期 由: ▾ | 任何日子 ▾ | 任何月份 ▾ | 任何年份 ▾ | 任何日子 ▾ | 任何月份 ▾ | 任何年份 ▾

**排序** **结果显示** **输出格式:**

-时间- ▾ | 顺序排列 ▾ | -或 排列- ▾ | 10 结果 ▾ | 以专辑分类 ▾ | HTML brief ▾

专辑:

☑ **北京大学图书馆** (19)
  北大名师 (19)

☑ **University of Hawaii at Manoa Libraries** (3,459)
  Hawaii War Records - U.S. Army Signal Corps Photographs (880) Save Our Surf (SOS) (286) Steve Thomas Traditional Micronesian Navigation Collection (2,086) TRAIL-Technical Report Archive and Image Library (167) Hawaiian Photo Album (40)

☑ **National University of Singapore Library** (293)
  Technical Reports (293)

☑ **University of Auckland** (5,845)
  Anthropology Photographic Archive (5,845)

*Other Interfaces*

http://prdlaarchive.lib.hku.hk/?ln=ja&as=1

Bibliotecas
Libraries 도서관 図書館
Pacific Rim Digital Library Alliance

# 環太平洋電子図書館会議

**274,407 のレコードを検索: :**

| すべての単語: ▾ | | すべての分野 ▾ | そして ▾ |
| すべての単語: ▾ | | すべての分野 ▾ | そして ▾ |
| すべての単語: ▾ | | すべての分野 ▾ | 検索 ブラウズ |

検索のヒント :: シンプルな検索

**Added/modified since:**                    **まで:**

| 検索期間 から: ▾ | すべての日 ▾ | すべての月 ▾ | すべての年 ▾ | すべての日 ▾ | すべての月 ▾ | すべての年 ▾ |

**並べ替え:**              **結果の表示**              **出力フォーマット:**

| - 新しいものから - ▾ | asc. ▾ | - または ランク - ▾ | 10 結果 ▾ | コレクションにより分離 ▾ | HTML brief ▾ |

**コレクション:**

☑ **Beijing University Library** (19)
  Famous Scholar database (19)

☑ **University of Hawaii at Manoa Libraries** (3,459)
  Hawaii War Records - U.S. Army Signal Corps Photographs (880) Save Our Surf (SOS) (286) Steve Thomas Traditional Micronesian Navigation Collection (2,086) TRAIL-Technical Report Archive and Image Library (167) Hawaiian Photo Album (40)

☑ **National University of Singapore Library** (293)
  Technical Reports (293)

☑ **University of Auckland** (5,845)
  Anthropology Photographic Archive (5,845)

# Invenio Possibilities
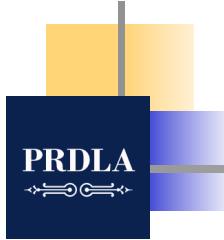
- Parallel searching on external collections
- Allow commenting/reviewing on items
- Save records to Basket
  - Allow batch export & email

# And…
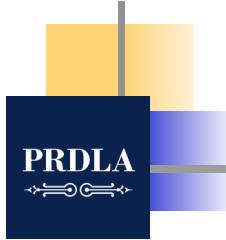
- Allow emailing of item
- Administrator can set up groups of similar-interest users
- Fulltext indexing
- Each collection, its own look & feel

環太平洋數字圖書館聯盟簡稱　--　태평양지역 전자도서관협의체　--　**環太平洋電子図書館会議**

# Included Libraries
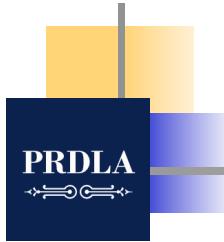
- Beijing University
- University of Hawaii at Manoa
- National University of Singapore
- University of Auckland
- University of British Columbia
- UCB

- UCLA
- UC Merced
- University of Washington
- University of Hong Kong
- USC

環太平洋數字圖書館聯盟簡稱 -- 태평양지역 전자도서관협의체 -- 環太平洋電子図書館会議

# Not Included
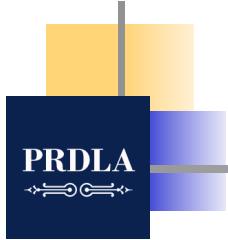
- No answer
- Or, no database
- Or, problems with OAI-PMH

# Survey Results (7 responses)
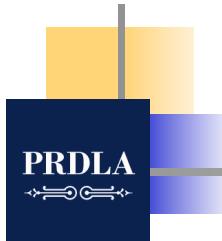
- Collection not large enough to produce relevant results
- Harvest complete metadata (not just oai_dc)
- "similarity" function needs to be explained
- "Browse" function confusing
- "Usage Statistics" page accurate?

# And..,

- Purposeful searches could be done
- Multi-language good
- Archive should be released to the public

# What Now?
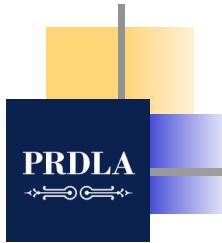
- Good pilot, but let's not continue?
- Or, Continue on? … and,
  - Include more sites
  - Continue with Invenio?
    - Further customize? How?
  - Pilot Project has completed?
    - New PRDLA committee?

環太平洋數字圖書館聯盟簡稱  --  태평양지역 전자도서관협의체  --  環太平洋電子図書館会議

# And..?

- Open to OAIster, Google & others?
- Identify themes & common topics
- Translate pages into other languages
- Scirus Respository Search Service?

環太平洋數字圖書館聯盟簡稱 -- 태평양지역 전자도서관협의체 -- **環太平洋電子図書館会議**

# SCIRUS
### for scientific information only

| About Us | Newsroom | Advisory Board | Submit Web Site | Help | Contact Us |

## Advanced Search
Basic Search   Search Preferences

All of the words ▼ | _____ | in | The complete document ▼

AND ▼

All of the words ▼ | _____ | in | The complete document ▼

Search

**Dates**
**Only show results published between**
before 1900 ▼ and 2008 ▼

**Information types**
**Only show results that are**

☑ Any information type        ☐ Conferences
☐ Abstracts                   ☐ Patents
☐ Articles                    ☐ Preprints
☐ Books                       ☐ Scientist homepages
☐ Company homepages           ☐ Theses and Dissertations

**File formats**
**Only show results that are**

☑ Any format                  ☐ HTML
☐ PDF                         ☐ Word

List more file types

**Content sources**

**Only show results from**

| Journal sources | Preferred Web sources |
|---|---|
| ☑ All | ☑ All |
| ☐ BioMed Central | ☐ E-Print ArXiv |
| ☐ Crystallography Journals Online | ☐ CogPrints |
| ☐ Institute of Physics Publishing | ☐ MD Consult |
| ☐ MEDLINE/PubMed | ☐ NASA |
| ☐ Project Euclid | ☐ Patent Offices |
| ☐ ScienceDirect | ☐ RePEc |
| ☐ Scitation | ☐ DiVA |
| ☐ Society for Ind. & App. Mathematics | ☐ Univ. Toronto T-Space |
| ☐ Pubmed Central | ☐ MIT OpenCourseWare |
| | ☐ NDLTD |
| | ☐ Caltech |
| | ☐ HKUST |
| | ☐ Organic Eprints |
| | ☐ PsyDok |
| | ☐ Wageningen Yield |
| | ☐ Curator |
| | ☐ IISc |
| | ☐ HKUTO |
| | ☐ Humboldt |
| | ☐ Digital Archives |
| List fewer sources | List fewer sources |

☑ The rest of the scientific web

Internet    100%

# Indirect Value?
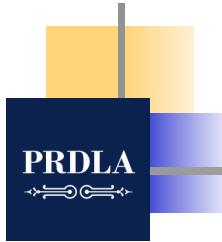
- Act as a unifying project of PRDLA
- Spearhead for other mutual development
- Impetus for all members to make all their digital content OAI-PMH compliant and searchable in Google, Yahoo, etc.

環太平洋數字圖書館聯盟簡稱 -- 태평양지역 전자도서관협의체 -- 環太平洋電子図書館会議
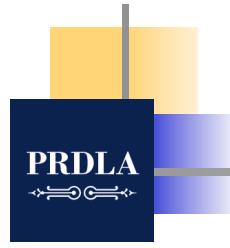
# 谢谢 ！

## Thank you ！

環太平洋數字圖書館聯盟簡稱 -- 태평양지역 전자도서관협의체 -- 環太平洋電子図書館会議