<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Institutional Repository at National Taiwan University</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>CHEN, Kuang-hua (National Taiwan University, Taiwan)</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Issue Date</strong></td>
<td>2008-01-30</td>
</tr>
<tr>
<td><strong>Text Version</strong></td>
<td></td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/11094/14147">http://hdl.handle.net/11094/14147</a></td>
</tr>
<tr>
<td><strong>DOI</strong></td>
<td></td>
</tr>
<tr>
<td><strong>rights</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Osaka University Knowledge Archive: OUKA*

http://ir.library.osaka-u.ac.jp/dspace/

Osaka University
Institutional Repository at National Taiwan University

Kuang-hua Chen  
Department of Library and Information Science  
National Taiwan University  
National Taiwan University Library  
khchen@ntu.edu.tw

Outline

- Background  
  - Overview of IR  
  - Research Archiving in Taiwan  
  - Overview of development of IR in Taiwan
- NTU’s experiences in IR  
  - NTU Repository (NTUR)  
  - System architecture and features  
  - Contents of NTUR  
  - Promotion for NTUR
- Copyrights in IR  
  - Separation of copyrights
- TAIR and IR30 project  
  - Goals and Visions  
  - Current state and Participating Institutions
- Conclusions
Essence of Institutional Repository

- IR system has to try its best to display **full texts** of data items for users
- IR system can preserve research histories and results
- IR system is an excellent platform for knowledge preservation and sharing

Necessities of IR – Researchers’ Viewpoints

- Traditional academic publishing models are too slow to keep up with fast-paced outputs of scientific researches
- Papers with full text available on web will have wide readership and citation rates
  - Take computer science for example
  - Offline paper cited times: 2.74
  - Online paper cited times: 7.03
  - Increase by 157%
  
- Open access will increase cited times from 50% to 250%
  
- As the Internet gains even wider usage, citation rates for online papers will continue growing quickly
Necessities of IR – Institutions’ Viewpoints

- Possible to analyze and evaluate research categories, trends, and results
- Demonstrate research capacity via a single gateway
- Strengthen academic influence and raise visibility

Cost to publish by publishers still increase, universities will need an alternative to disseminate research results

- ArXiv (http://arxiv.org) as a dissemination system, it operates at a factor of 100-1000 lower cost than a conventional peer-reviewed system.


Current State of IR’s Development

- Many universities around the world have already launched IR systems.
- Most IRs have a collection of thousands items, few have that of ten thousands items.

Currently only small number of countries have official policies to support IR development at the government level.

However, the support is very active at the university level

- Few universities in Taiwan have installed their IRs; many have started planning IR; some have started deploying IR systems
Higher Education in Taiwan

- 163 universities and colleges including polytechnics

<table>
<thead>
<tr>
<th></th>
<th>Junior Colleges</th>
<th>Independent Colleges</th>
<th>Universities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public</td>
<td>3</td>
<td>11</td>
<td>41</td>
</tr>
<tr>
<td>Private</td>
<td>13</td>
<td>42</td>
<td>53</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>16</strong></td>
<td><strong>53</strong></td>
<td><strong>94</strong></td>
</tr>
</tbody>
</table>

- Universities have to find ways to demonstrate academic or practical research results in terms of one identity

Research Archiving in Taiwan

- **Journal articles**
  - National library is responsible for services of table of contents and document delivery

- **Conference papers**
  - No systematic archiving plan

- **Technical reports**
  - No systematic archiving plan except few funding organizations

- **Theses & Dissertations**
  - Some universities have archived their own theses and dissertations with full texts
  - National library has archived printed theses and dissertations with digital abstracts for both. Some have digital full texts as well.
Why Need Institutional Repository
- especially for Taiwan or Asian Countries

- The number of published papers is large, but the citation rate is comparatively lower
- The following table shows statistics of paper published by the top ten universities in Taiwan from 1996 to 2006

<table>
<thead>
<tr>
<th>University</th>
<th>Number of papers published</th>
<th>Paper cited</th>
<th>Highly cited papers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Numbers</td>
<td>Rank of the world</td>
<td>Cited times</td>
</tr>
<tr>
<td>NTU</td>
<td>22,623</td>
<td>83</td>
<td>138,819</td>
</tr>
<tr>
<td>NCKU</td>
<td>12,374</td>
<td>218</td>
<td>60,292</td>
</tr>
<tr>
<td>NTHU</td>
<td>9,076</td>
<td>323</td>
<td>45,976</td>
</tr>
<tr>
<td>NCTU</td>
<td>8,706</td>
<td>331</td>
<td>31,168</td>
</tr>
<tr>
<td>YMU</td>
<td>5,724</td>
<td>486</td>
<td>44,975</td>
</tr>
<tr>
<td>NCU</td>
<td>5,288</td>
<td>506</td>
<td>25,563</td>
</tr>
<tr>
<td>NSYSU</td>
<td>4,858</td>
<td>534</td>
<td>18,666</td>
</tr>
<tr>
<td>NCHU</td>
<td>4,116</td>
<td>619</td>
<td>16,452</td>
</tr>
<tr>
<td>NTUST</td>
<td>3,424</td>
<td>727</td>
<td>12,518</td>
</tr>
<tr>
<td>CGU</td>
<td>3,000</td>
<td>800</td>
<td>15,541</td>
</tr>
</tbody>
</table>


Building IR to Solve Problems

- Research outputs in Taiwan are scattered without information systems to present a whole picture of research achievements and to feature searching and browsing functions for potential users.
- Many operational information services provide only metadata and abstracts rather than full texts.
- No systematic way to deposit important digital materials such as conference papers and technical reports
- Institution's self-consciousness makes it to host a system as a persistent repository and an exchange platform for research results.
Possible Solutions to IR Systems

- **Purchase commercial software or services**
  - Columbia University purchases the services provided by ProQuest, i.e., the Digital Commons.
  - However, Columbia University has decided to use DSpace in the near future.

- **Use open-source packages**
  - University of Michigan at Ann Arbor uses DSpace.
  - University of Southampton uses EPrints.

- **Modify open-source packages**
  - NTU Library modifies DSpace based on the local requirements to develop NTUR.
System Architecture of NTUR

- Modules with thick line are modified by NTUR developing team
- Modules with thin line are remained the same as the original DSpace’s design

**Application Layer**
- **DSpace Public API**
  - Core Tools (configuration, logging)
  - Search (Lucene Wrapper)
  - History Recorder
  - E-person/Group Manager
  - Administration Tasklet
  - Authorization

**Business Logic Layer**
- **Browse Tools**
- **Content Management**
- **Handle Server**
- **Stashplus**
- **Handle Manager**
- **Submission Workflow Manager**

**Storage API**
- **RDBMS Wrapper**
- **File System**
- **Bitstream Storage Manager**
- **MySQL**

**Storage Layer**

NTUR System Functions (1/2)

- **Community & Collection**
  - Community is defined as the department of NTU
  - Collection can be defined by each community

- **Main functions**
  - Browsing by community & collection, date, title, and author
  - Full-text search with rich features for post processing
  - Language switch on the fly
    - traditional Chinese, simplified Chinese, and English
  - Upload & download ranking by author and community
  - Single sign on to NTUR
  - Data Submission
NTUR System Functions (2/2)

- No need to login for data browsing, searching and downloading
- Need to login for data uploading
  - Single sign on policy
  - Interactive mode
    - Faculty and Ph.D. students can upload data by themselves or ask contact staff to help
    - Master students and undergraduates have to ask contact staff to help
  - Batch mode
    - Only open to contact staff of departments or subject specialists of library

Potential Items in IR

- Journal articles
- Conference papers
- Dissertations and Theses
- Teaching or training materials
- Technical Reports
- Speech materials
- Preprints and postprints
- Any digital object
Items Acquisition

- Collect items from Internet and other databases
- Cooperate with interested departments first
- Consider different methods for different departments
- Hold promoting activities and workshops
- Explain the copyright issues
- Do not force faulty to participate but persuade them of advantages of IR
- Do not bother faculty but expect staff

Number of Registered Items in NTUR

- Over 45,000 metadata records
  - Technical reports: 23,000
  - Journal papers: 4,000
  - Conference papers: 1,500
  - Theses and dissertations: 13,000
  - Others: 3,500 (from NTU press and websites)
- About 27,000 with searchable full texts
  - Technical reports: 9,000
  - Journal papers: 3,700
  - Conference papers: 1,300
  - Theses and dissertations: 11,000 (full-text indexing only)
  - Others: 2,000
Promoting Strategy

- **Procedural level**
  - The priority of archiving
  - Rights and mandatory
  - Interoperability among existent information systems
  - Users community
    - Setup contact window
    - Compile name list of faculty and graduates
  - Mechanism for long-term operation of IR
  - Cooperation with other unions or consortiums

- **Administrative level**
  - Administrative regulation for IR
  - Encouraging measures
行政院國家科學委員會專題研究計畫 成果報告

新富農業、跨國移工、認同政治(2/2)

計畫編號：個人型計畫
計畫編號：NSC93-2412-H-002-004-SSS
執行期間：93年08月01日－95年01月31日
執行單位：國立台灣大學社會學系研究所

計畫主持人：蔡佩君
計畫參與人員：蔡國

2008-01-30

2008-01-30
Copyrights in IR

Why can Archive Published Papers

- Separation of copyrights
  - Self-archiving right
  - Full-text indexing right
  - Access right

- Most publishers allow the first two rights at least according to the survey of SHERPA/ROMEO project

- If institutions have access right to some journals, IR systems should provide full texts for papers published in those journals

- Otherwise, IR systems have to carry out full-text indexes and then feature full-text search functionality for users

- If there are no full texts, IR systems have to provide metadata and informative links
Separation of Copyrights

- Self-archiving right
- Full-text indexing right
- Access right

<table>
<thead>
<tr>
<th>Author</th>
<th>Institutional Repository</th>
<th>Publisher</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Author retain the copyright" /></td>
<td><img src="image" alt="Archiving" /> • Self-archiving</td>
<td><img src="image" alt="Access" /> • Archiving</td>
</tr>
<tr>
<td><img src="image" alt="Journal paper" /></td>
<td><img src="image" alt="Full Text Indexing" /></td>
<td><img src="image" alt="Access" /></td>
</tr>
<tr>
<td><img src="image" alt="Copyright has been transferred to the publisher" /></td>
<td><img src="image" alt="Full Text Indexing" /> • Self-archiving</td>
<td><img src="image" alt="Access" /></td>
</tr>
</tbody>
</table>


Most Publishers Allow Self-Archiving in IR

- 93% publishers allow authors self-archive their papers in institutional repository based on the investigation of SHERPA/ROMEO project

Source: http://www.sherpa.ac.uk/romeo.php
## Policies of Copyrights of Publishers

<table>
<thead>
<tr>
<th>Publisher</th>
<th>Full-text index</th>
<th>Personal website</th>
<th>IR</th>
<th>Free public servers</th>
<th>Embargo period</th>
<th>Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackwell Publishing</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td>Varying by journal. Author can post their version of the Article on their personal website, their employer’s website/repository and on free public servers in their subject area.</td>
</tr>
<tr>
<td>Elsevier</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td>Authors of articles may also post the final version of their manuscript publicly on their personal website or their institution’s website (including institutional repository).</td>
</tr>
<tr>
<td>Nature</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td>Six months</td>
<td>Authors are encouraged to archive their version of the manuscript in their institution’s repositories and on their personal web sites.</td>
</tr>
<tr>
<td>Oxford</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td></td>
<td></td>
<td>Authors have the right to deposit the postprint and/or URL or PDF of the finally published version of the article into an institutional or centrally organized repository.</td>
</tr>
<tr>
<td>IEE/IEEE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td>Authors and/or their companies shall have the right to post their IEEE-copyrighted material on their own servers without permission.</td>
</tr>
<tr>
<td>Springer-Verlag</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td></td>
<td>An author may self-archive an author-created version of his/her article on his/her own website and his/her institution’s repository.</td>
</tr>
<tr>
<td>ACM</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td>Authors retain the right to post author-prepared versions of the work in a personal collection on their own Home Page and on a publicly accessible server of their employer.</td>
</tr>
</tbody>
</table>

## Copyrights in Taiwan - Situations

- 10% of academic publications are published by international publishers
- The copyright of 15% of academic publications was transferred to domestic publisher
- The copyright of 25% of academic publications was reserved by authors (non-exclusive license, or open access)
- The copyright of 50% of publishers was not specified owner
Copyrights in Taiwan - Solutions

- Papers collected in the domain of NTU
  - Send email to authors to get non-exclusive license
- Papers published by international publishers
  - Clarify the policies of publishers
- Papers with copyright transferred to domestic publishers
  - Discuss with these publishers
- Papers with copyright retained by authors
  - Get the non-exclusive license from authors when they upload papers to IR systems
- Papers with unspecified copyright owners
  - Discuss with publishers and get the non-exclusive license from authors when they upload papers to IR systems

Taiwan Academic Institutional Repository (TAIR) & IR30 Project

http://ir.lib.ntu.edu.tw/
TAIR

- Reach a consensus of “many institutional repositories, a centralized gateway”
- Raise visibility and respectability of Taiwan’s researches
- Present Taiwan’s research capacity as a whole
- Contribute to research communities all over the world

Goals and Visions of TAIR

- Build a mechanism for “many institutional repositories, a centralized gateway”
  - “Many institutional repositories” will remain the specialty of each institution and demonstrate its research outputs
  - “A centralized gateway” will present achievements of researches of Taiwan as a whole

- A gateway for IRs in Taiwan
  - Short-term goal
    - Federated search for metadata via OAI-PMH
  - Long-term goal
    - Federated search for full texts via centralized indexes or meta-search
IR30 project (1/2)

- Conform to “many institutional repositories, a centralized gateway” to push institutions in Taiwan to construct their own IR systems

- Help at least 30 institutions to build their own IR systems by May 2008
  - NTU selects 6 seed institutions with qualified IT technology based on the regional distribution and helps them to deploy IR systems
  - Each seed institution helps other institutions near by to deploy IR systems
  - Invite institutions with deployed IR systems to join TAIR

- Up to Dec. 2007, 22 universities and 36 polytechnics join IR30 project

IR30 Project (2/2)

- NTU release NTUR package with comprehensive manuals and training materials
  - Free IR software, NTUR, (DSpace enhancement)
  - Technical documents and user’s guides
  - System functional requirements
  - Standard Operating Procedure (SOP)
  - User’s FAQ
  - Training materials

- Hold workshops for promotion of IR concepts and construction of IR systems
Seed Institutions

Conclusions

- Government funds construction of IR
  - NTU is responsible for implementation of IR package
  - NTU will release IR package and help other institutions

- Implement our own IR package based on open-source software

- We have effective development strategy
  - Seed institutions go first
  - Participating institutions follow up
  - A tree scheme for effective and efficient development

- Polytechnics are interested in IR particularly
  - IR systems can demonstrate the specialties of polytechnics

- TAIR will be a gateway with federated search for full texts via centralized indexes or meta-search
Thank you

Please visit NTUR at
http://ntur.lib.ntu.edu.tw/
for our institutional repository