The Construction of Institutional Repository from Viewpoints of Dissemination and Citation

由傳播與引用的觀點看機構典藏之建置

Kuang-hua Chen 陳光華
Associate Professor, Department of Library & Information Science
National Taiwan University
臺灣大學圖書資訊學系副教授
Head, System and Information Technology, NTU Library
臺灣大學圖書館系統資訊組組長
khchen@ntu.edu.tw

(Draft)

Abstract

The meaning of Institutional Repository (IR) is to preserve the research outputs of research institutes. The preserved contents as a whole will demonstrate the achievements and influences of research institutes. Many investigations pointed out that an open-access IR system can decrease the cost in dissemination of scholarly information and increase the citation times of research outputs. As a result, the construction of institutional repository systems has been broadly receiving attentions. National Taiwan University (NTU) has been publishing many research outputs. How to make these researches available to the public is the spiritual goal of NTU Repository (NTUR) project; how to implement a user-friendly system is the technical goal. This article describes current developments of IR in the world from the viewpoint of dissemination and citation. The implementations and features of NTUR are discussed as well.

摘要

機構典藏的真實意涵是存取學術機構的研究產出，以展現機構的研究成果與學術影響。許多研究指出具有開放取用特性的學術機構典藏系統，能夠降低學術資訊的傳播成本，並提升研究成果的引用頻率。因此，機構典藏系統的建置已經受到學術機構以及學者們的廣泛重視。臺灣大學擁有豐富的研究成果及大量的學術出版，如何善用並促進這些研究能量的呈現、取用與交流，是臺灣大學機構典藏計畫的精神目標；為使用者建構一個易於使用的機構典藏系統是該計畫的技術目標。本文以資訊傳播與學術引用的觀點，說明目前全球機構典藏的發展現況，並探討臺灣大學機構典藏系統的實作與主要的功能。

Keywords: Citation, Institutional Repository, Scholarly information
關鍵詞：引用，機構典藏，學術資訊
1. Introduction

Nowadays, it is a main stream to present various digital documents, videos, audios, and multimedia via WWW browsers, since browser has been regarded as the de facto common platform for Internet applications. In addition, the roles of users and authors of Web resources have been intermingled and the roles of publishers have become less important than ever before. The quick development of WWW is a challenge and also an opportunity to universities and research institutes, which may create alternative ways to publish, disseminate, and acquire scholarly information without the intervention of publishers and aggregators. As a result, the cost derived from publication and acquisition of scholarly information will be reduced and the time of dissemination of scholarly information will be reduced as well.

The purpose of dissemination of information is to make information usable or citable. Citation has been regarded as a good way to clarify the roles and positions of researches in the whole research structure. Although the reasons to citing other papers are multiple, we all recognize at least there are some connections among cited papers and citing papers. Bibliographic couplings and co-citations are two obvious examples. Citation also demonstrates a very special function, i.e., the influence or impact of papers, researchers, and research institutes. In these years, academic evaluation has become an important topic, especially, the evaluation for universities or research institutes. Citation, therefore, has been used as a way to evaluate the academic performance. To increase citation times of research outputs will be the plus to academic evaluation for research institutes. Institutional repository seems to be one of the solutions.

Although many definitions for institutional repository could be found in literature, the idea of institutional repository is very simple. It is to deposit every part of researches from the very beginning to the final publication for universities or research institutes. With the observation to publishing research results in recent years, many researchers prepare their papers or articles using word processors and submit them in digital form rather than printed form. As a result, the volume of “born-in-digital” documents increases very quickly. However, Lynch (2003) pointed out we still have no good ways to preserve these research results in their digital form. The idea of institutional repository may be a good turning point to rethink the model, mechanism, and means for the preservation and dissemination of scholarly information. From the viewpoint of universities or research institutes, to encourage faculty to publish research outputs and to make dissemination of faculty’s papers effective are the better means to increase the citation times of published papers and then to increase universities’ influences. Institutional repository systems as academic portals of universities or research institutes, the preserved scholarly information as a whole will demonstrate the achievements and influences, make accessing scholarly information much more convenient, and share research results with other universities and institutes in the world.

National Taiwan University (NTU), one of important universities in Taiwan, has been obtaining lots of significant achievements in researches and publishing great deals of research results. It is necessary for NTU to fulfill a good model to preserve and disseminate NTU’s research outputs. NTU has initiated a task force for National Taiwan University Repository (NTUR) project and established an operational institutional repository system named NTUR. At present, NTUR contains 45,526 records of various scholarly information including journal articles, conference papers, lecture notes, technique
reports, etc. Among these 45,526 records, 27,003 records are bundled with full texts.

This article is to describe the dissemination and citation of scholarly information and to discuss the construction of IR systems. The structure of this article is shown as follows. Section 2 discusses the advantages of online papers in terms of dissemination and citation. Section 3 will briefly review the current status and development of institutional repository. Section 4 presents the NTUR and discusses major modifications to original DSpace. Section 5 is a brief conclusion.

2. Dissemination of Scholarly Information

Since Denis de Sallo published the first issue of the *Journal des sçavans* in 1665, the research outputs has been disseminated through journals for hundreds of years. (Hallam, 1970, p. 406) This seems a default model for dissemination of scholarly information. According to the statistics of ISSN International Centre (2007), there are about 905,090 registered periodicals as Table 1 shows. This statistics depicts that a great deal of journals were created to disseminate scholarly information based on the traditional model. Although the types of publishers and the types of publications have been changed along the time passing, this model is basically not changed too much. However, there are at least three problems in the default model from the viewpoint of universities. The first is a university is hard to know research results of its own when taking a university as a whole; the second is a university is difficult to trace research results from the temporal perspective; the third is the continuing increases in the cost of acquisition of journals. In addition, this model is not efficient and effective from the viewpoints of dissemination and citation of scholarly information.

<table>
<thead>
<tr>
<th>Table 1. Statistics of ISSN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type</strong></td>
</tr>
<tr>
<td>other types, annuals and other</td>
</tr>
<tr>
<td>updating Web site</td>
</tr>
<tr>
<td>periodicals</td>
</tr>
<tr>
<td>monographic series</td>
</tr>
<tr>
<td>newspapers</td>
</tr>
<tr>
<td>no attempt to code</td>
</tr>
<tr>
<td>Updating loose-leaf</td>
</tr>
<tr>
<td>Updating database</td>
</tr>
</tbody>
</table>

Figure 1 shows a scholarly information cycle for current information process from creation, evaluation, publication, acquisition, indexing, storing, retrieval, reading, to citation. (Cleveland & Cleveland, 2001, p. 23) We could examine the relative relations of each stage of information processing. Basically, from creation, evaluation to publication takes long time and takes high cost when we apply traditional model. Therefore, some reformations are adopted to speed up the processing of publications of scholarly information. The introduction of electronic journals (e-journals) is one of the efforts. At the present time, the number of e-journals continues to increase according to the report of Glose, Fletcher, & Bromberg (2005). The increasing rate is about 25% in two years. Please refer to Figure 2 for details. That means most researchers thinks e-journals will help the dissemination of scholarly information, otherwise the number of e-journals will not increase so quickly.

![Scholarly information Cycle](image)

**Figure 1. Scholarly information Cycle**


Many contributors are involved in this cycle including researchers, publishers, librarians, and information systems. The roles of researchers are multiple folds. They could be creators, reviewers, and readers. Therefore, citation behavior occurs naturally between the creators and users who become creators later on. Citation is regarded as the base of research development, since almost all researches are based on the preceding researches. Researchers usually “cite” the ideas and papers of pioneers. A citation includes two objects: one is the citing work; the other is the cited work. Although there are many reasons for a citation (Weinstock, 1971), the cited work has something to do with the source work. In addition, citation has been regarded as one of important factors for academic evaluation of
research institutes, universities, and researchers. That is to say, the publication and dissemination of scholarly information are not only for research purposes, but also for increase of influence of research institutes and researchers.

![Figure 2. The Growth of Online Electronic Journal](image)


Let’s take a look at the statistics of publication and citation of papers written by researchers in Taiwan as shown in Table 2. National Taiwan University is the top one university in terms of both published papers and cited papers. However, she only ranks 83 in the world in terms of number of papers and ranks 231 in terms of cited times of papers, respectively. Although some researchers would not like to admit there exists a kind of western academic hegemony, statistics shows a little bit tendencies. The reasons might be many folds. Two of the reasons may be these papers are not broadly disseminated and not easily accessed. Actually, on the one hand, research outputs in Taiwan are scattered in many different information systems without some kinds of integrated information systems to present a whole picture of research achievements and to feature searching and browsing functions to potential users. In addition, many operational information services provide only metadata and abstracts rather than full texts. On the other hand, there is no systematic way to deposit important digital materials such as conference papers and technical reports, which are very important for a few disciplines. As a result, for research institutes or universities in Taiwan or other Asian countries, whose academic impacts are relatively lower with comparison to American and European research institutes, to make their research results freely available to public in form of full texts via Internet seems a very good strategy to increase their academic influences and earn their reputations.
Table 2. Statistics of the Top Ten Universities in Taiwan from 1996 to 2006

<table>
<thead>
<tr>
<th>University</th>
<th>Published Papers</th>
<th>Cited Papers</th>
<th>Highly Cited Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Numbers</td>
<td>World Rank</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>508</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>4,118</td>
<td>619</td>
<td>18,452</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>


Universities and research institutes in Taiwan are confronted to the challenge and thrust to become one of top 100 universities in the world. They apply budgets, settle the goals and draw up strategies for pursuing the goal. As mentioned above, the e-journals will speed up the processing of scholarly information. However, the cost of acquisition of e-journals is not reduced actually, because many publishers bundle e-journals and printed journals together. This makes an obstacle in some sense for general public to access scholarly information. In recent years, two concepts, open access and institutional repository, have great influences to dissemination of scholarly information. Open access is the answer to the third problem mentioned in the very beginning of this section. In contrast, institutional repository is the answer to the first and the second problems. In general, open access and institutional repository are almost mentioned together, but not necessary. Open access depict a model for scholarly information, i.e., a model of freely online scholarly information available to general public. The major goal of an institutional repository, in contrast, is to demonstrate the research results of university or research institute as a whole and the information system of institutional repository is regarded as an academic portal for university or research institute.

One research conducted by Lawrence (2001) pointed out the cited times of online papers is 157% larger than those of offline papers. Another research said freely available online papers will increase their cited times from 50% to 250%. (Hajjem, Gingras, Brody, Carr, & Harnad, 2005). If we consider the cost of dissemination for scholarly information, the traditional model is even less efficient and less effective. Ginsparg (2001) reported ArXiv (http://arxiv.org) as a dissemination system; it operates at a factor of 100-1000 lower cost than a conventional peer-reviewed system. Of course, ArXiv is an open-access repository. Bergstrom (2001) has reported that the journals purchased by using 91% of budget only created 38% citations, but freely available online scholarly information and other resources purchased by using the rest 9% of budget created 62% citations. These
statistics shows the following insights.

- Open access greatly decrease the cost in dissemination of scholarly information
- Online papers greatly increase the cited times
- Freely available online papers contribute major part of citations

The insights together make clear the importance of an open-access institutional repository. To sum up, open-access institutional repositories can disseminate scholarly information of universities and research institutes in an efficient and effective way. They will also reinforce influence and reputation of universities or research institutes.

3. Current Development of Institutional Repository

Institutional repository has been recognized as an information infrastructure for scholarly information from the viewpoint of research institutes. Many universities and research institutes all over the world have started building their own IR systems. 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. The following briefly discusses the current development of IR.

(1) Europe and America

In USA, there have already had over 40% of universities putting their own institutional repository systems in operation. 88% of the rest universities are thinking the construction of IR systems. (Lynch & Lippincott, 2005). That is to say, over 90% of USA’s universities take serious steps to construct their own institutional repository. Van Westrienen & Lynch (2005) have surveyed the institutional repository in Europe and USA. The results showed that almost every university in Germany, Netherlands, and Norway has exploited IR systems. Furthermore, some universities built more than one IR systems for different purposes. These statistics showed the importance of IR has been fully recognized by the advanced academic or educational organizations in the developed countries. Although only few countries have draw up government-level policies, the promotion of IR, the construction of IR systems and the application of using IR systems as platforms for preservation and dissemination of scholar information have become a global trend.

(2) Asian Countries

Table 3 shows the current status of construction of IR systems in Asian countries. It shows that the numbers of IR systems in Japan, India, Hong Kong, and Taiwan are more than other Asian countries. The percentage of construction of IR Systems in Asian countries is about 4% to 10% except mainland China. The percentage in Hong Kong is higher than other countries, since the number of universities in Hong Kong is relatively much smaller. In fact, it is possible that the statistics are different in different investigations executed by different researchers. For example, some investigations identified 2 IR systems in mainland China. (Oliver & Swain, 2007) CNKI (http://www.cnki.net) in China is a special case among these IR systems. CNKI is the one centralized IR system for about 300 universities in China and the number of universities is increasing. As a result, the specialty of each university could not known by simply using CNKI. However, taking the CNKI as the
presentation of integrated research outputs of China is a very useful and helpful from users’ perspective.

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of IR Systems</th>
<th>Country</th>
<th>Number of IR Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>1</td>
<td>Russia</td>
<td>2</td>
</tr>
<tr>
<td>Japan</td>
<td>26</td>
<td>Pakistan</td>
<td>1</td>
</tr>
<tr>
<td>India</td>
<td>25</td>
<td>Philippines</td>
<td>1</td>
</tr>
<tr>
<td>Taiwan</td>
<td>7</td>
<td>Singapore</td>
<td>1</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>3</td>
<td>Indonesia</td>
<td>1</td>
</tr>
<tr>
<td>Israel</td>
<td>2</td>
<td>Korea</td>
<td>1</td>
</tr>
</tbody>
</table>


Although Table 3 shows the numbers of IR systems in Japan and Korea are 26 and 1, respectively, those numbers increase very quickly according to the reports of DRFIC2008 conference held in Osaka, Japan. Most universities in Korea develop their IR systems using a product of joint efforts, dCollection. As a result, the number of IR systems has increased quickly to 215. (Lee, 2008) However, the major contents are dissertations or theses which are transformed from existent digital libraries systems. The same situation occurs in Japan, too. Figure 3 shows the growth of number of IR systems in Japan. National Institute of Informatics is responsible for the development of IR in Japan, which reviews applications, allocates budgets, monitors performance, etc.

(3) Taiwan

With comparison to other countries, the promotion and construction of IR in Taiwan start relatively late. In the early stage, only 7 universities started to implement their IR system. After NTU applied an IR project granted by Ministry of Education, she organized an IR30 task force which plan to install IR system in 30 universities in one year. At present, 73 universities or institutes join IR30 task force. Please visit ir.org.tw for the complete list of participating universities. Table 4 shows a few universities which have installed their own IR systems. They are National Taiwan University (NTU), National Chengchi University (NCCU), National Chiao Tung University (NCTU), National Tsing Hua University (NTHU), National Sun Yat-sen University (NSYSU), Feng Chia University (FCU), and Kun Shan University (KSU). Among these universities, NCCU, NCTU, and KSU adopted NTUR which is enhanced by NTU based on DSpace package. In contrast, FCU and NTHU have used DSpace system directly. However, NTHU will use NTUR in the near future.

NTU had decided to use DSpace beta version 1.4 implemented by HP and MIT as the base to develop NTUR with modification and enhancement. DSpace has been used by many universities. This shows that DSpace meets the common needs of many universities.
However, users in Taiwan have their special requirements and these features are not provided by the DSpace package. NTUR system has been launched in early 2007 and presented to many presidents of universities in Taiwan for the purpose of IR promotion. NTU also provided system documents such as Standard of Procedure, Reference of System Specification, User Manual, System Manual, Training Courses, FAQ, etc. All of these are for one main goal, which is to persuade universities in Taiwan to construct their own IR systems and to pave a smooth way for them with the precious NTU’s experience.

![Figure 3. The Growth of IRs in Japan](source)


### Table 4. IR Development in Taiwan

<table>
<thead>
<tr>
<th>University</th>
<th>Number of Records</th>
<th>IR system</th>
<th>Web Site</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTHU</td>
<td>11,499</td>
<td>NTUR</td>
<td><a href="http://140.114.72.31/dspace/">http://140.114.72.31/dspace/</a></td>
</tr>
<tr>
<td>NCTU</td>
<td>1,057</td>
<td>NTUR</td>
<td><a href="http://140.113.39.78/dspace/">http://140.113.39.78/dspace/</a></td>
</tr>
<tr>
<td>NCCU</td>
<td>1,708</td>
<td>NTUR</td>
<td><a href="http://140.119.115.26/">http://140.119.115.26/</a></td>
</tr>
<tr>
<td>NSYSU</td>
<td>2,929</td>
<td>NTUR</td>
<td><a href="http://140.117.120.62:8080/nsysuir">http://140.117.120.62:8080/nsysuir</a></td>
</tr>
</tbody>
</table>

Source: Authors collected data from the IR system of each university. The numbers are subject to change.
4. Institutional Repository at NTU

There are many solutions to construction of IR systems. We could purchase turnkey solutions, develop our IR systems by implementing totally new systems, exploit our IR systems using open-source packages, or design IR systems by modifying open-source packages. Nevertheless, once decide to construct an IR system, we have to analyze the requirements and features, consider the available resources, and make the best decision. With the observation on the current development of IR in the world, the three major approaches to constructing IR systems are shown as follows.

- **Purchase commercial software or services**
  For example, 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 package**
  For example, University of Michigan at Ann Arbor uses DSpace. University of Southampton uses EPrints.

- **Modify open-source package**
  For example, NTU modifies DSpace based on the local requirements.

Each solution has its own right way which research institutes will take. Research institutes intending to exploit IR systems have to consider the costs of each potential solution including short-term cost and long-term cost. NTUR developing team has carefully surveyed each solution and considered resources available. Finally, the third solution is adopted and DSpace is used as the base for NTUR. Among many available open-source packages, the reasons we use DSpace are shown as follows.

- DSpace is implemented using Java. This meets the policy of NTU Library for system development.
- DSpace is platform independent. It could be installed on Windows-based systems or Linux-based systems. A lot of Java APIs are available. This makes us much easier to modify original DSpace package.
- DSpace has been adopted by more than 200 institutions. (DSpace, 2007) From the operational viewpoint, it has been proved its stability and effectiveness for the large volume of data with variant types.
- DSpace supports OAI-PMH. It is convenient to fulfill integrated search or federated search later on.
- DSpace integrates handle mechanism. The digital object registered in IR systems has a unique URN. As a result, researchers could cite registered object in IR systems.
- DSpace has complete frontend and backend. The time to develop a system based on DSpace will be shortened.

As a matter of fact, DSpace has been continuingly developed by HP and new versions of DSpace will be released ceaselessly. NTUR developing team has decided to use DSpace 1.4 beta version and will not follow the new version of DSpace. As a result, NTUR...
developing team will develop the new features based on DSpace 1.4 beta version and local requirements.

NTUR has the similar system architecture as DSpace 1.4 beta version shown as Figure 4. It is composed of storage layer, business logic layer, and application layer. Each layer consists of many functional modules. Those modules with thick line are revised by NTUR developing team and the rest modules with thin line are remained unchanged as original DSpace’s modules. In general, storage layer is responsible for management of the registered documents and submission of documents; business logic layer is responsible for administration, authorization and work flow of submission; application layer is responsible for interaction between outer systems and users. The main parts of our modifications to each layer are discussed in the following.

(1) Web UI

In order to improve the degree of user friendly, NTUR developing team has been adding new features for our potential users based on local requirements, e.g., the post processing for search results, ranking for uploads and downloads, multi-language interface, identification of document formats. NTUR developing team will continue modifying user interface and hopes users could use NTUR with ease and by instinct.

The homepage of NTUR is shown as Figure 5. The right up corner shows switches for multi-language interfaces and numbers of “items with full text/total items”, number of “visitors”, and number of “online users”. Users could switch to one of multi-language interfaces on the fly while browsing items or searching items. The message of numbers of “items with full text/total items” clearly declares the current items registered in NTUR. To our knowledge, NTUR is one of larger IR systems which stores items with full texts in comparison to other IR systems in the world.

The right-hand side of NTUR’s homepage shows information about top upload and top download of communities and authors. This information will help users have an overview of NTUR’s contents. In fact, the communities of NTUR are departments rather than colleges. However, in order to avoid a long list of communities, the center of NTUR’s homepage shows the list of colleges with the number of their own items, respectively. Users could drill down from each college to see the details of communities by clicking on the “+” symbol on the left of each college. The left-hand side of NTUR’s homepage shows the feature of browsing “Communities & Collections”, “Titles”, “Authors”, and “By Date” and information about “News and Activities”. Clicking on the “Communities & Collections” will pop up the details of communities and the corresponding collections as Figure 6 shows. While clicking on the “By Date”, we will have a temporal experience on the registered items in NTUR. For example, the publication date of the earliest item registered in NTUR is 1963 as shown in Figure 7. Reviewing items by this mode will help users, especially, scientific historians interested in NTU, to trace the development of researches in NTU.

Another little but not the least modification to the user interface is that NTUR promptly shows icons of items with the corresponding file format. For example, Figure 7 shows PDF-format items with the PDF icon. Users can access items by clicking on the icon without many clicks. The corresponding PDF document will be displayed directly.
Figure 4. System Architecture of NTUR

Source: A modified version based on http://www.dspace.org/index.php?option=com_content&task=view&id=145

Figure 5. Homepage of NTUR
Figure 6. Snapshot of Clicking on “Communities & Collections”

Figure 7. Snapshot of Clicking on “By Date”
(2) Media Filter

Media filter is responsible for filtering out the contents of registered items. The major modification for this part is to enhance the features of parser. An institutional repository contains large volume of items with various formats. The most of items are in PDF format comparing to other file formats. Therefore, it is very important to process PDF files correctly. An open source Java PDF library, PDFBox, is bundled with DSpace package. However, the functions of PDFBox are not good enough, especially for Chinese. Therefore, while users search items or browse items, the contents may be scrambled. This makes users confused and inconvenient. General speaking, PDF files are created by using many different ways. For example, a PDF file could be created from MS Word file, Postscript file, Tex file, or the likes. In addition, there are too many solutions to created Chinese PDF files early time. This delivers a challenge to process Chinese PDF files for an operational IR system. In fact, the correction rate of PDFBox is about 0.3 for Chinese. Therefore, the modification to Media Filter is a very important task.

(3) Search Features

In order to make users of NTUR have a much more structured view to the retrieved results rather than just a list of items, NTUR developing team has been focusing on the post-processing features. Figure 8 shows a snapshot for search results. The search results are post-processed based on “Category”, “Author”, and “Year” as shown in the left-hand side of the snapshot. Category demonstrates the classifications of search results based on communities and collections. Users could know the distribution of the research results and have an idea about which departments have carried out the similar researches. While clicking on “Author” as shown in Figure 9, users will have a view of which researchers focus on the related research topics. The feature is very helpful for decision makers of institutions to find the experts in their institutions. While clicking on “Year” as shown in Figure 10, users will have a view of the temporal distribution of search results. It is useful for users to trace the development and to know the peak of related research in time.

Actually, it takes time to post-process the retrieved items. NTUR developing team decided to apply Ajax to make post-processing efficient and to makes users enjoy new features at the same time.

(4) Browsing Features

The browsing feature designed by DSpace is to search designated items among all items and to show them in the current “page”. In order to make this feature much more efficient, DSpace applied sliding-window-like method which works like to drag the user’s view to a particular section of data items. Therefore, there is no obvious paging for browsing items in the original design of DSpace. However, users are used to the paging mode while they use the search engines or other search features of database systems. NTUR developing team has modified the browsing features by adding clear paging mode and the option of number of items per page. Please take a look at the thick rectangle in Figure 8.

We also make users to easily see the full text by just one click rather than many clicks. That is to say, the retrieved items are not only paged, but also associated with designated icons to denote the file formats of full texts. In contrast, those items without associated icons have metadata only.
(5) Database

NTUR developing team will release NTUR to help other research institutes in Taiwan to deploy their own IR systems according to the project granted by Ministry of Education. We have to consider the different situations of various research institutes. For example, NTUR may be installed in Linux or Windows Server; the database management system could be open-source software and adopted widely in Taiwan; NTUR could be also compatible to commercial database management systems. In fact, PostgreSQL is bundled in DSpace package, but most users and database administrators in Taiwan are not familiar with it. On the contrary, MySQL is popular and we decide to integrate it into NTUR. However, in order to make the NTUR effective when the volume of data becomes much and much larger, ORACLE database management system could be used as well.

The database schema of original DSpace package is very simple. Only four relations are designed by DSpace developing team: basic elements table, basic element relationships table, object browsing table, and other elements table. In order to support many new implemented features, NTUR developing team designs new tables, e.g., authority control table, term mapping table, author-document table. In addition, new attributes are proposed for the existent tables, e.g., attributes for users and communities, attributes for file linking, attributes for access control, and attributes for other languages.

Although, NTUR works stably and features useful functions, there are still many problems and challenges. We continue to think about the following points: the modification of system architecture, the management of system resources, the effectiveness of information retrieval, the difficulty of item acquisition, the mode of item visualization, and the authorization of data items.

Figure 8. Snapshot of a Search result
Figure 9. Result of Clicking on Author

Figure 10. Result of Clicking on Year
5. Conclusions

Institutional repository has received much more attention from researchers and research institutes globally. IR systems can preserve research outputs, demonstrate research results, and reinforce academic influence. In addition, connection among various IR systems in the world under a virtual umbrella will create global platform for open access, reduce the cost in access of research resources, and speed up the dissemination of scholarly information. Many investigations have concluded that freely online papers available to the public will increase their citation times. The advantages of constructing IR systems are very obvious to all of us.

National Taiwan University has fully recognized aforementioned advantages and built up the NTUR to serve our faculty and potential users from all over the world. Currently, the number of registered items in NTUR is 45,526. Among these registered items, 27,003 items are with full texts. It is one of the bigger IR systems in the world. In addition to enhance the features of NTUR, NTUR developing team will promote IR in Taiwan and release NTUR package under the support of government granted project. Hopefully, we will have a virtually integrated IR system to demonstrate the research outputs of Taiwan and to serve as an academic portal for Taiwan. Therefore, researches carried out by Taiwanese researchers will have increasingly influence and make great contributions to academic researches.

Acknowledgments

Author thanks University Librarian, Professor Jieh Hsiang, for his support to NTUR project. Author is also grateful to the members of NTUR developing team at National Taiwan University Library. They are Su Chuan Chang, Mei-lin Kuo, Yu-ting Chiang, Ting-fai Ho, Cipher Kao, Chiu-chu Liu, Hsiao-ying Hung, Hou Ieong Ho, Shih-Li Chen, Feng-i Lin, Ya-tzu Liu, Hsin-Yi Yeh, Yi-chun Liu, Bai-hsuan Lee, Ren-shiang Liu, Yun-jung Young, Tsan-juin Ho, Ya-fang Weng, Wen-chi Huang, Ying-fang Huang, Ying-lin Liu, Yu-Chieh, Lin, Feng-ju Chen, Ching-hua Liao, Li-yun Chiang, and Yi-ling Tu.

References


