
A Study Analyzing Reference Linking Service in Digital Libraries

Younghee Noh*, Patricia R. Ladd**, Kyoungsik Na***

ARTICLE INFO

Article history:

Received 23 March 2015

Revised 30 March 2015

Accepted 9 April 2015

Keywords:

Reference Linking Service,
Digital Libraries,
Information Resource,
Linking Methods

ABSTRACT

This study seeks to present an option for including a variety of information resources in a reference linking service system. Discussion on reference linking service in journals began in 1999 and applied to real systems, but linking to research papers and patent registers has more recently emerged. This study further suggests that system construction include various information resources such as books and dissertations. This paper also presents principles and prerequisites that should be considered essential when building such systems, including expansion of resources and media formats, development of identifiers on every sort of information resource, linking methods to existing DB, publication revolution and managing copyright-holders, use and improvement of OPACs, and building perfect authority DB.

1. Introduction

1.1. Need and purpose of study

With the development of information technology and the Internet, intellectual works are increasingly presented in digital formats. Web information resources have become important sources of recent information for researchers. Most proceedings are published online among expanded unofficial online publications, and authors upload their papers on the web. Currently, a number of online journals are 'born digital,' meaning that no print version exists.

In 1997, Bide predicted that 90% of journals' references would be digitized in 20 years (Bide, 1997), thus linking to the original text for citation would be a significant service (Van de Sompel and Hochstenbach, 1999).

Reference linking service is designed to enhance user convenience based on an identification system. Recently, identification systems like DOI have been imposed on most academic journals, which brought about a landmark change in digital content distribution, establishing the foundation to provide users with reference linking service.

* Professor, Department of Library and Information Science, Konkuk University (irs4u@kku.ac.kr)

** MLS University of North Carolina at Chapel Hill (patriciarladd@gmail.com)

*** Assistant Professor, Department of Library and Information Science, Konkuk University (kn05d@kku.ac.kr)
(Corresponding Author)

International Journal of Knowledge Content Development & Technology, 5(1): 69-88, 2015.

<http://dx.doi.org/10.5865/IJKCT.2015.5.1.069>

Although the service is referred to by various names, such as reference linking service or original text linking service, materials to be linked may be original texts, websites, journals, or books, all with the aim to link the literature to related materials. Using this system, all of the stakeholders including publishers, subscription agencies, researchers, and libraries are determining ways to create value-added content from related information objects, especially by linking text with related information sources. As a result, many information providers use reference-linking services for their digital content. In general, most of the services consist of links of built-in metadata.

There are many ways to provide reference linking service, most of which depend on specific identifiers such as DOI or OpenURL. In general, the methods involve either reference linking service provision after building a separate abstract and bibliography DB or journals and publisher-centered linking service where publishers impose DOI on each journal site. The majority of the linked literature focuses only on papers in academic journals, leaving out a wide range of other information formats, including books.

This study seeks to present ways to include various types of information resources in suggesting a system for reference linking service. Though papers in journals were the first to be reference linked, other formats such as patents have recently become available to linking services. In the future, it will be important to support even more formats, such as books or dissertations, and this paper suggests a system for accomplishing this.

1.2. Expected effects and application plan

This study is expected to contribute to maximized user convenience, the generation of bibliometric statistics, and improved use of domestic literature by synthesizing various existing linking methods.

First, user satisfaction levels can be improved through reference linking services. The research process will be more thorough and efficient when the literature retrieved by users is linked with the full-text of other related literature.

Second, this study could contribute to increased citation of domestic literature and journals. Linking to related literature could increase citation of domestic literature, which could, in turn, improve international citation indices of domestic literature. Bibliometric methodology is mainly focused on validating the computations of international citation indices, while this study will improve the international citation index of domestic literature by providing links to the cited original text in references.

2. Theoretical background

2.1. Literature review

Simply put, reference linking services use an identification system as a base, which means improved service using link information instead of URLs and linking to the abstracts, full-text, and bibliographies of the cited references in the literature itself. This design is possible through the already-imposed

identification system within the literature. Types of the references to be linked in this way can be diversified to include papers in journals, books, reports, and websites. However, previous research mainly focused on papers in academic journals or other research papers.

Such research began 20 years ago when Van de Sompel and Hochstenbach (1999) published three papers on reference linking in the hybrid library. The first paper compared reference linking to other approaches to solve problems. In addition, dynamic and static solutions, just-in-case and just-in-time linking, as well as open and closed linking frameworks were introduced. Their second paper introduced SFX, a dynamic, real-time linking solution built to achieve set goals. However, the concept was still quite general, especially in its applicability to digital libraries. The third paper demonstrated proven methods in the “SFX@Ghent & SFX@LANL” experiment. As a starting point for SFX, some part of the local and remote information resources among the digital library resources in the Research Library of the Los Alamos National Laboratory and University of Ghent Library were used. The SFX framework was more generalized to realize technologies to easily transfer one digital library environment to the other and minimize overload in constructing distributed libraries compatible with SFX.

Kang et al. (2006) suggest linking to commercial academic information DB and full-text linking services provided by portal sites like Naver in the research on national scientific technology and information resource development. The research differs from imposing a unique identifier in that it links to data through metadata building on the content of each cooperative institution.

Lee, Yoo, & Lee (2010) intended to suggest an appropriate model for domestic citation services based on analysis of various citation information services to explore information resources and progress with recently increased interest in the tool. As a result, differences in functions and resource types of the citation information service were found and modeled. Eventually, the study suggested service priorities and individual services for the development of a domestic citation index service through expert surveys, along with three principles to provide academic information service using useful citation information.

Li, Burnham, Lemley, & Britton (2010) compared popular databases Web of Science, Scopus, SciFinder, and Google Scholar based on, among other categories, their reference linking abilities. After a detailed breakdown of each database’s strengths and weaknesses, the authors concluded that all the databases missed linking to some references included in the other databases.

Many studies have noted the inefficiency of manually entering metadata information, calling for the need for an automated tool to extract metadata from uploaded content. One of the earliest of these studies was Bergmark (2000), which identified the limitation of manual work among other problems in providing reference linking services. The author thus developed an algorithm to extract a link to the original text, demonstrating the high intelligence of the algorithm through experiments. Dennis et al. (2003) also suggested a method to automatically recognize bibliographic data of the cited literature from papers in journals in mathematics, and provide original text in PDF or DjVu format.

Lee & Lee (2010) argued that building a reference DB and matching identifiers were required to provide reference linking services to link between different information resources. CrossRef, PubMed, and Web Of Science built a reference linking system based on identifiers like DOI and

PMID using automation tools including Inera's eXstyles and Parity Computings' Reference Extractor, and domestically, several institutions such as KISTI (Korea Institute of Science and Technology Information) and the NRF (National Research Foundation of Korea) are building reference DBs. However, building an inclusive DB is impeded by the various forms of reference description. Therefore, improving efficiency in building a reference DB was attempted by extracting metadata through automatic parsing of references and developing Citation Matcher, a system for matching identifiers such as DOI, PMID, and KOI. By developing Citation Matcher for automatically parsing and linking references, they sought to automate the work of building references and identifier matching, processes that relied on manual work in the past. The solution parses citation information including journal name, author, and volume number from references through analyzing the reference pattern if the reference is "Copied & Pasted" as it is, and acquires and link identifiers by matching the parsed information with metadata such as CrossRef, PubMed, and yesKISTI in the standardized manner (Lee et al., 2007).

Groza, Grimnes, & Handschuh (2012) also explored a method for extracting semantic metadata efficiently and fully. The authors point out that many earlier methods of metadata extraction were either manually done by article authors or captured only fragmentary aspects of metadata, often leaving out important elements like references. The authors also stressed the need for an extraction and linking system able to expand to handle multiple different data formats.

Blake and Knudson (2002) focused on analysis of linking cases of A&I DB and linking from bibliography to the original text. To design such a system, they argued that a system using metadata to determine identifiers and components in various 'link=to' phrases linking different systems should be designed.

Jeong et al. (2004) sought to build environments where users could access materials regardless of conditions, where digital full-text could be freely distributed between KISTI and other institutions with digital information.

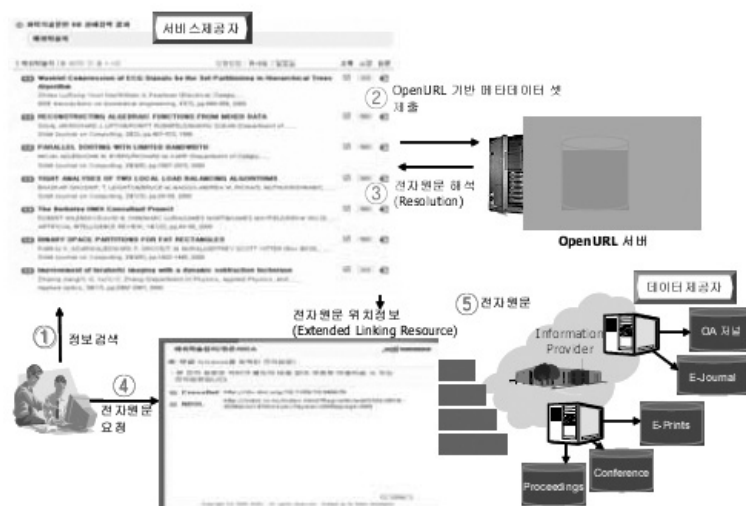


Fig. 1. Information linking based on OpenURL (Source: Lee et al., 2005)

On the other hand, Kim & Jeong (2005) suggested a reference DB construction plan acting as information sources and evaluation instruments, along with analysis of the ongoing reference DB construction and problems. For the purpose, reference DB in the specialized research information centers (KSCI of KISTI, KCI of National Research Foundation of Korea, and KoMCI of Korea Academy of Medical Science) were analyzed. The study suggests original text linking through separate reference building while this study is aimed at original text linking through references in the literature.

Lee et al. (2004) explored linking methods from specific literatures to referenced literatures and pointed out that, despite the wide distribution of digital content, the URL service with location information for digital content is still problematic. They also raised a concern about the sufficiency of the system for digitizing physical works since its main purpose is for purely digital content. Accordingly, they developed KOI (Kisti Object Identifier) as a unique identifier for KISTI, based on analysis of foreign cases of identification system applications in major digital content service agencies, and designed a digital original text linking system in the area of science technology based on KOI. Later, Lee et al. (2005) designed and built an open architecture for academic information distribution incorporating the international framework of OpenURL to improve linking between metadata of OAI architecture and full-text and provide scalability to similar resources. They sought to reinforce operability between information resources and scalability to similar resources in the distributed environment in the study.

Grogg (2006) also argued for OpenURL as the solution to context-sensitive linking difficulties. After providing a complete analysis of OpenURL's features, especially compared to other systems, the author explores how the OpenURL can be used for further standardization, integration of more accurate linking, and adaptability to future changes.

Bergmark, Arms, & Lagoze (2000) suggested a linking system based on API for reference linking services, presenting a structure that a referenced literature is linked to its original text as follows.

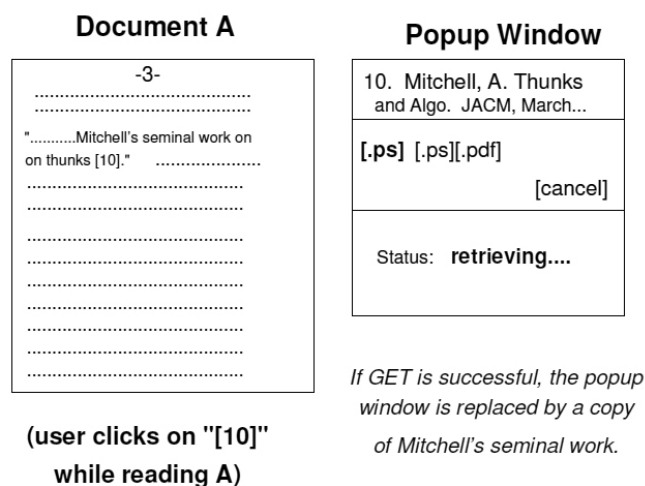


Fig. 2. API reference linking framework

Kim (2005) described an overview of digital resource linking systems using OpenURL and conducted a comparative analysis 16 OpenURL link resolvers based on the server installation status, number of journals provided, service customization status, and service for use statistics.

While the vast majority of studies focus on reference linking systems in academic journals, other studies have attempted reference linking in other environments with different types of resources. Kwon & Kim (2007) conducted research on building a national R&D reference linking system. National R&D information includes information on national projects, participants, tools or actions generated from the research, and research equipment. Based on a recognition of the usefulness of reference linking services for researchers, users can see related information, including results of research projects, information on research participants, related equipment, and human resources related to performance by direct link, without searching. Kwon & Kim also developed a national R&D reference linking system based on an identification system to promote information linking.

Choi et al. (2000) sought to build a linking center for convenient link information for related institutions. They tried to improve information accessibility for users and provide customer-centered services by forming intellectual link information and networks linking related resources through this link center.

Ryu & Lee (2002) built a systematic framework for foreign technologies to provide efficient searching services for users, foreign technology resource collection information for related institutions, role allocations for original text service, and a collaboration framework for efficient management and national sharing of imported and distributed foreign technologies.

In research on building a patent register linking system based on DOI, Kwon, & Noh (2003) looked into the status of patent numbers as the unique identification system for national patent documents, and sought a linking system using the patent information digital identification framework. Their work describes the standard digital identification system unifying diversified patent number systems for each country. In the suggested linking system, simple index information on patents are built into the metadata format using each country's patent numbers, and identifiers and location information for each patent register are centrally controlled.

Considering the above mentioned research results, each study focuses on linking service between academic papers, research papers, and science technology information. However, various types of references are cited in a single article, including books. Therefore, it is necessary to seek ways to expand reference linking services to encompass more references after identifying the most commonly cited information resources, which is where this study differs from previous research.

2.2. DOI and OpenURL

Major publishers in 1999 founded the Publishers International Linking Association (PILA), a non-governmental organization to build a reference linking system which links cited references on the Internet based on DOI and OpenURL. They thus launched the CrossRef service.

The movement was led by publishers in the early stage in advanced countries such as the US and the UK while recently digital content-related agencies, including IP, CP, digital libraries, and S/W firms have been leading the DOI business. In Japan, CID and metadata management systems

have been developed to describe properties of content and distribution in CIDF (Content ID Forum). JST implements a linking service to CrossRef for domestic distribution of academic papers and application of foreign digital journals through J-Stage service.

The above-mentioned services, based on DOI and OpenURL, are briefly introduced as follows.

1) *DOI*

As a representative digital identification system, DOI is a sort of barcode imposed on every type of digital content distributed through the computer, through which content location and distribution route can be tracked. DOI was developed in 1997 using the URN principle and certified as the US standard of ANSI/NISO in May of 2000. DOI includes copyright information of digital content and access via the Internet. It has proved useful for its permanency and uniqueness, used in a number of companies in the world. DOI is also used in the electronic commerce of digital content, automatic tracking of content distribution, copyright management, and prevention of unfair use.

2) *OpenURL*

OpenURL is a resolver linking to information resources provided by digital information institutions using metadata. OpenURL is a NISO international standard linking, which is the framework developed by Herbert Van Desompel and his colleagues at Ghent University to resolve problems of the mono-inflexible existing electronic linking. OpenURL retrieves metadata from sources and provides multiple information resources after interpretation of the data. Also, it can be linked to other identification systems inserting various identifiers into the transmitted metadata. The phrase structure of OpenURL is composed of BASE-URL and Query. OpenURL is applied to academic information linking in SFX, UKLON, CrossRef, KISTI, and KINS.

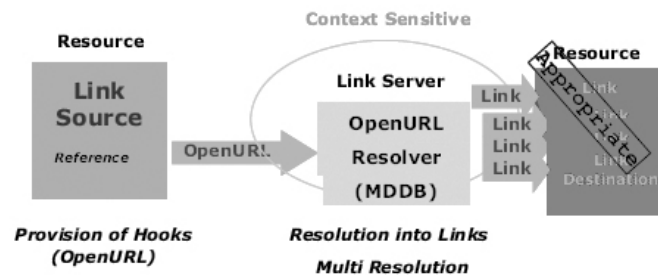


Fig. 3. OpenURL Framework

3. Trends and cases of reference linking

The most common international instances of reference linking services based on identification systems are CrossRef and JST (Soichi, 2005), and both cases provide reference linking services based on DOI identification. CrossRef provides reference linking services within the same journal, having imposed DOI numbers on around 30 million articles in 19,181 academic journals. The five

cases below were selected for review due to their early adoption of reference linking techniques and widespread use today.

1) CrossRef

CrossRef was launched by PILA in 1999 to link cited references on the Internet based on DOI and OpenURL. As of January 2001, CrossRef is run by PILA as an independent organization, under which 2,260 journal publishers, many partners, libraries, and related institutions participate as members.

Currently, CrossRef, Content Direction, Enpia system, Learning Object Network, TSO, and Copyright Agency Ltd are registered DOI-RA among the International DOI Foundation, of which only CrossRef is run on a business model, with service range over articles in science journals, books, and proceedings.

The purpose of CrossRef is to provide accessibility to the full-text of references by linking the cited papers published in other publishers through DOI. CrossRef provides reference linking for digital journals by imposing DOI on 9,000 academic journals and 13.5 million papers. The service does not retain full-text, but provides linking to the full-text through DOI imposed on papers in journals which participate in CrossRef.

2) KCI of the National Research Foundation of Korea: <http://www.kci.go.kr/>

KCI is able to compute the impact of journals according to statistics and number of citations needed for research resource management, thus it can be used as a tool to evaluate the qualitative level of journals published in a specific area. By studying the relationship between articles and works cited in domestic journals, the system can identify trends and evaluate research results. However, there will be significant synergy effects if full-text linking is added.

The screenshot shows the KCI search results for the article: "A Study on Development and Utilization of Specialized Programs in Curriculum of Library and Information Science". The article is by Lee, Hyeon-yeon and was published in the Journal of Library and Information Science, Volume 38, Number 1, in 2008. The page also displays a table of related articles.

NO	저자명	연월일	주제명	학술지명	면(호)	페이지
1	최남주	2008-09	공적정보기관의 정보유통의 중요성에 대한 연구 - 행정정보학회 학술대회 발표논문 -	한국도서관정보학회지	38 (3)	83 - 71
2	노동희	2008-06	한국정보과학회 2008년 학술대회 논문집 - 정보유통의 중요성에 대한 연구	한국정보과학회지	40 (2)	420 - 480
3	조남주	2008-12	한국정보과학회 2008년 학술대회 논문집 - 정보유통의 중요성에 대한 연구	한국정보과학회지	40 (4)	243 - 289
4	정영희	2011-10	공공정보기관을 위한 정보 유통 서비스 개발 방안 - 정보유통학회 학술대회 발표논문 -	한국정보과학회지	42 (1)	205 - 233
5	김영준	2011-12	공공정보기관을 위한 정보 유통 서비스 개발 방안 - 정보유통학회 학술대회 발표논문 -	한국정보과학회지	42 (4)	517 - 540
6	최남주	2011-12	공공정보기관을 위한 정보 유통 서비스 개발 방안 - 정보유통학회 학술대회 발표논문 -	한국정보과학회지	42 (4)	148 - 187

Fig. 4. KCI Search Results Display

3) DBPia: <http://www.dbpia.co.kr>

DBPia launched a service linking to the cited literature's full-text, but it is not yet in the stage of vitalization. If the system is modified, it is expected to generate great achievements due to the huge amount of literature registered in the DB.



Fig. 5. DBPia search results

■ 연관자료

참고문헌

- * 이 논문의 참고문헌이 13건 있습니다. 현재 논문의 참고문헌을 찾아 신청해주세요!
- Claire Creaser, Sally Maynard and Sonya White, LISU Annual Library Statistics 2006 : Featuring Trend Analysis of UK Public & Academic Libraries 1995-2005(Leicestershire : Library and Information Statistics Unit, 2006), p.176.
- <http://www.library.yale.edu/consortia/> [Link](#)
- Mark Rowse, "The Consortium Site License : A Sustainable Model?" Libri, Vol.53, No.2(2003), p.3.
- A. Keyhani, "The Online Journal of Current Clinical Trials : An Innovation in Electronic Journal Publishing," Database, Vol.16(1993), pp.14-23.
- 한국교육학술정보원, 대학도서관 해외전자정보 구축현황 및 수요조사 보고서(서울: 한국교육학술정보원, 2004), p.114.
- Ketcham-Van Orsdel and Born, "E-journals Come of Age," Library Journal, Vol.123, No.7 (1988), p.42.
- Walter L. Warnick, "The Search for Hidden Treasure : R&D Results in the Deep Web," In STIP Meeting (April 30, 2003) [Link](#)
- BrightPlanet.com LLC, The Deep Web : Surfacing Hidden Value (July 2000), pp.2-4. [Link](#)
- Bin He, et al., "Accessing the Deep Web : A Survey,"(2004) [Link](#)
- Bonnie Shucha, "Conquering the Invisible Web,"(Aug, 2995) [Link](#)
- <http://romeo.eprints.org/stats.php> [Link](#)
- <http://roar.eprints.org/?action=browse> [Link](#)
- Sherpa/RiMEO, "Publisher Copyright Policies & Self-archiving," [Link](#)

연용논문

- * 이 논문을 인용한 다른 논문이 3건 있습니다.
- 남태우, 김은주, '학술도서관 장서관리정책 사례 연구 한국정보관리학회, 정보관리학회지, 27(1), 2010, 79-102' [Link](#)
- 이지연, '공공도서관 장서관리 현황 분석 한국문헌정보학회, 한국문헌정보학회지, 43(3), 2009, 247-267' [Link](#)
- 윤희윤, '공공도서관 장서관리의 정론과 역설 한국도서관정보학회, 한국도서관정보학회지, 40(1), 2009, 5-22' [Link](#)

Fig. 6. DBPia search results display

4) *NHN: <http://academic.naver.com/>*

Naver, a popular Korean search engine, started to provide a ‘citation information’ search service between articles as a foundation for academic research development in 2012, under which citing and cited information for 2.1 million papers published in 90,000 journals in the humanities and 4.8 million papers published in 140,000 journals in social science is provided.

The service range for academic citation information covers references in journals registered or candidates for registration in KCI, and is applied to papers published in the humanities (since, 2001), social science, natural science, and engineering for the past 5 years.



Fig. 7. Naver search display

5) *EBSCO Host*

EBSCO is a database comprised of 375 full-text and secondary research databases with over 550,000 eBooks and 360,000 e-journals and print journals. When searching for an article in an EBSCOhost database, users have access to EBSCO’s “SmartLinks+” engine, which continuously manages the availability and full-text of articles available to users (depending upon their institution’s subscriptions). Users then receive one-click access to the full-text of related articles and references through an OpenURL framework.

Search results are displayed as shown in <Figure 8>. Detailed listings with reference linking are displayed as shown in <Figure 9>.

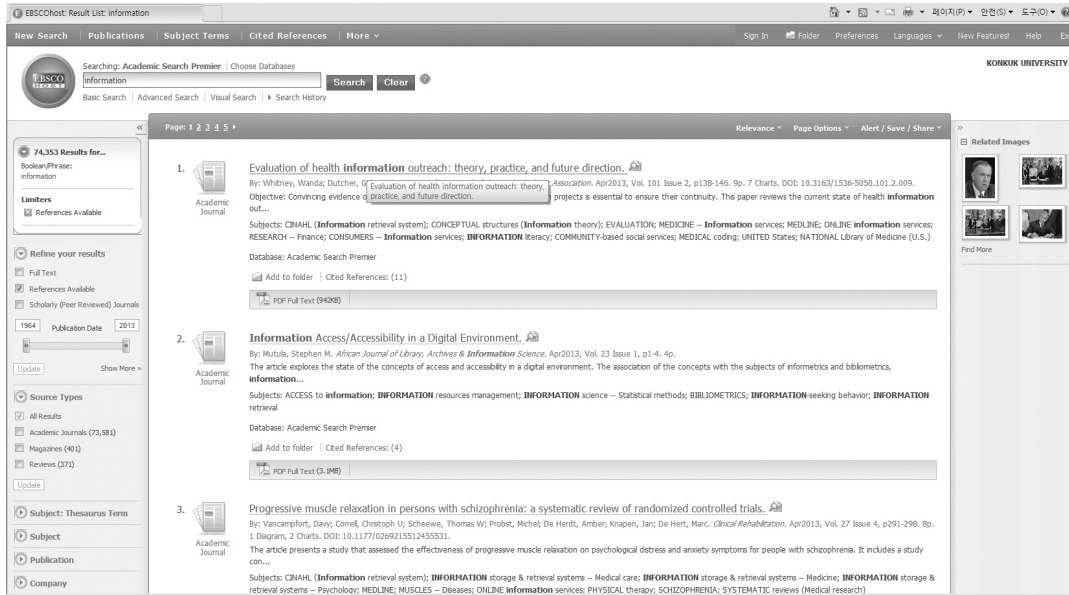


Fig. 8. EBSCO Host search results display

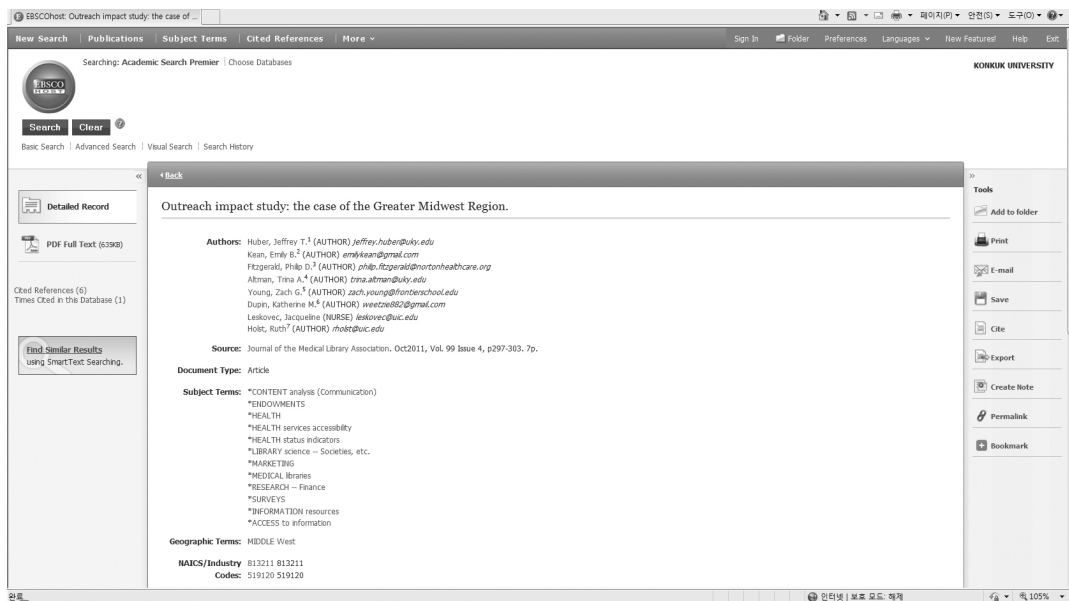


Fig. 9. Detailed listing of entry in EBSCO host

6) Results of case studies and implications

Although some research here in Korea has proposed ways to link to full-text based on DOI or identification codes, national libraries, public digital libraries, NDSL, and commercial DB (DBPia,

etc.) have no system linking to full-text of cited literature. By contrast, a number of institutions have established a full-text linking system and provided the service in foreign countries. However, ways to build more systematic and effective full-text linking services should be considered in Korea, given that the service has already been launched. The example below shows linking a specific paper retrieved by the user to the full-text of those it cites.

4. Reference linking system building strategy

As use of digital information in libraries and digital journals has rapidly increased, demand for solutions to provide an optimized linking service to access various information sources for the end user has also increased. Currently, although SFX and CrossRef are used based on DOI and OpenURL, both are limited to meet the needs of libraries. In other words, these resources are focused on journals, but types of information resources in libraries are more varied, including books, dissertations, research papers, proceedings, and multimedia materials. To tackle the problem, efforts in the publication phase are required, and this study seeks to present directions for future development.

4.1. Principles and prerequisites for building a reference linking system

Among many conditions to build a reference linking system, principles and prerequisites that must be considered essential are as follows.

1) Expansion of resource and media type

Although the existing systems are mostly focused on papers in periodicals or papers in academic journals, there are various types of works cited in research. Thus, the resource types able to be linked should be expanded into academic journals, dissertations, proceedings, and research papers. Also, material media types can be separated into electronic full-text or virtual document formats. The former includes PDF, TIF, XML, HWP, DOC, PPT, TXL, and TPD (DVI) files and the latter includes BAS and DET.

2) Development of identifiers for all types of information resources

Methods to impose identifiers on the various types of information resources mentioned above should be sought. The body of research on systems imposing identifiers on papers in periodicals and analyzing citations were discussed in the literature review. However, ways to impose identifiers on more varied information resources should be explored. To link to full-text based on identification systems, linking techniques such as DOI, OpenURL, eCID, UCI, and KOI are required. Patent information is identified by the patent number, and a way to impose an identifier on research papers is also needed.

For books, identification systems can be developed for linking services using ISBN (Lee, 2004). For example, 'Suffix' can be utilized in developing DOI identification for books.

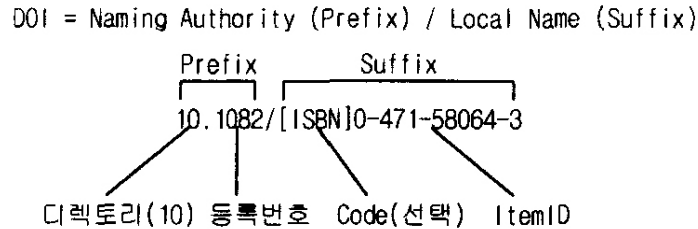


Fig. 10. Examples of Identification system for books

As for research papers, generally a unique number is imposed on each research project. In particular, the national R&D identification system is the only permanent number on each national R&D project. According to the “rules on management of national R&D projects,” information identifiers imposed on the inclusive project number and related information objects are used as a managing system, which can identify and access information objects (Kwon & Kim, 2007). Therefore, identifiers should be imposed and used based on the project number of the papers. Any URLs used should be actively maintained.

3) *Linking to existing DB*

Many institutions already provide full-text linking service, and future efforts should seek cooperation with them. Collaboration with foreign and domestic portals such as CrossRef, EBSCOhost, Google, Naver, DB firms like DBPia, national information distribution institutions, and libraries should be considered.

4) *Publication revolution and control on copyright holders*

Publication of information resources begins, naturally, with publishers. When publishers produce on- or offline literature, identifiers should be inserted on every reference. If the identifier is imposed for each work published, users will be able to directly access the related literature. Also, authors should be able to provide perfect references on their cited works, allowing the same identifier to be imposed on the cited literature. In that work, subjects to be linked for full-text should include all types of information resource as well as foreign and domestic materials.

5) *Use and improvement of OPACs*

The advantages of OPAC reference linking are demonstrated and realized through Ex Libris. Currently, national libraries in Korea are trying to collect and keep every information material published in Korea, and therefore have the largest amount of material of any institution. If their OPAC systems were applied to a reference linking system, many advantages would follow, including 1) increased use of internal information resources in national libraries, 2) increased rate of citation and use of domestic information resources, 3) provision of instant links to resources in and outside of libraries, and 4) competitiveness gained against commercial DB through proven link provision.

6) *Perfect authority DB building*

It is necessary to build various authority DBs in order to build a reference linking system, which can also be applied to the construction of cited literature analysis systems. Building authority DBs to include the name of authors, journals, and institutions is an essential part of providing efficient service.

4.2. *Effects for system building*

This study seeks to present considerations and directions in building a reference linking system in digital libraries. Advantages of such system building are summarized below.

First, digital libraries will connect significant electronic resources through links, from which will increase use of the internal information resources. Second, users' satisfaction levels will improve through direct linking to external open information resources. Linking to PubMed and CrossRef can also be considered. Third, users are able to view the full-text of all DBs subscribed to through full-text linking within DBs provided by commercial firms. Fourth, the rate of citation and use of domestic materials will increase since retained information resources are mostly domestic. Fifth, the system can function as an information source and evaluation tool (Kim & Jeong, 2005). Currently, many resources are issued in electronic forms, thus the number of full-text resources linked to will increase. Therefore, spillover effects will be considerable after development of the reference linking system.

For example, as a result of studying online material types in public digital libraries, it was found that 3 million digital journals and books were available, and holdings of video materials, websites (after resolving copyright issues), and documents were respectively 1,000, 40,000, and 27,000. Therefore various types of information resource are connected efficiently through full-text linking, details of which are displayed in the table below.

Table 1. Examples of online material types in public digital libraries

Types of materials		Number	
Text	Electronic journal	3,253,474	2,587,935
	Electronic book		665,539
Video		1,104	
Image	Image	236,153	236,152
	Electronic journal		1
Sound		1,605	
Software		4	
Website		39,968	
Substitute materials for the disabled		8,934	
Documents		27,076	
Others		38	

In addition, RISS DB provided by KERIS includes 118,000 dissertations, 168,000 academic journals, 3.379 million papers in domestic journals, 221,000 public lectures, 41.069 million papers in foreign journals, 8.229 million books, and 1.406 million other materials, which shows that the potential effects of reference linking are high, considering the scale of data provided.

Based on the above statistics, effects from the reference linking service and citation information service are summarized as follows.

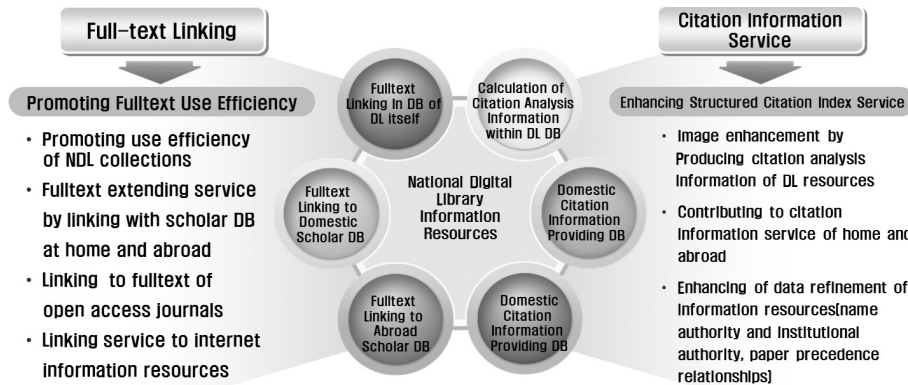


Fig. 11. Effects of full-text linking service and citation information service

4.3. Direction for system construction

As shown earlier, a number of researchers have already studied reference linking systems, experimental building cases, and tests of existing systems. This section presents the current status and direction for system construction in terms of full-text linking method, full-text linking range, and direction for system building, based on the analyzed results.

1) Full-text linking method

The types of full-text linking systems fall generally into four categories: metadata construction, reference DB construction, link center construction, and linking using OpenURL. Metadata construction means that institutions in collaboration build metadata into their content and then link the metadata (Kang, 2006), similar to CrossRef, PubMed, and NDSL. Simple index information for patents is also available in the metadata format using patent number. Identifiers and location information on each patent original register are centrally managed to link the patent material anytime (Kwon & Noh, 2003).

Reference DB construction involves building a reference linking system based on identifiers like DOI and PMID using automation tools like 'Reference Extractor,' which is operated by many institutions such as KISTI and NRF (Lee & Lee, 2010; Kim & Jeong, 2005). Building link centers for generating intellectual link information between science information and related resources improves users' information accessibility and provides customer-oriented services (Choi, 2000). A linking structure connecting cited literature to its full-text using an API-based linking system was also suggested (Bergmark, Arms, & Lagoze, 2000).

OpenURL is helpful in resolving copyright issues and fosters open academic information distribution (Lee, 2005), increased interoperability in distributed environments and scalability to similar resources

(Lee et al., 2005), and analysis of digital resource linking systems using OpenURL (Kim, 2005).

2) *Full-text linking range*

Tools such as ScienceDirect, SAGE Journals online, and PNAS online are limited in range, providing full-text only to internal journals with full-text already within the DB. Tools such as Google Scholar, CiteSeer X, CiNii, and ACM Portal use an expanded range to external journals and web information sources. The status of reference and full-text linking service for domestic and foreign journals is as follows.

Table 2. Status of reference and full-text linking service

Category	DB	Service	Reference linking	Linking service
Foreign academic DB	ScienceDirect		§ Reference list provided § CrossRef and linking to SCOPUS provided	§ Cited number and literatures in SCOPUS provided
	Sage Journals Online		§ Reference list URL provided § CrossRef provided	§ Cited literature linking by connecting to SCOPUS and Google Scholar § Searching author's other works and other information through Google Scholar
	PNAS Online		§ Reference list and abstract provided § CrossRef and ISI, linking to PubMed provided	§ Linking to ISI, Google Scholar, PubMed § Searching Author's other works and other information through Google Scholar and PubMed
	Google Scholar		§ Reference list and link provided § Similar materials § Citing work's references	§ Google Web Search § Library link § Open WorldCat DB
	CiNii		§ Reference list and link provided § Related author and serial list provided	§ Searching available on papers in J-STAGE, Journal@rchive § Providing information in journals by WebcatPlus linking
	CiteSeerX		§ Cited literature list provided § Cited literature link	§ Expanded searching to Google, Yahoo, MSN, CSB, DBLP § CiteULike, Feedback linking
	ACM Portal		§ Cited literature's bibliography and link provided § Reference bibliography and link provided	-
Foreign citation info-DB	Web of Science		§ Reference numbers and internal reference linked § Visualized citation map on reference (citation relation)	§ Reference number provided § References' full-text link
	SCOPUS		§ Internal library OPAC electronic journal full-text link § NBI link	§ Link to internal library's OPAC e-journal § Link to SSCI, SCI site
Domestic citation info-DB	KSCI		-	-
	KCI		-	-
	KoMCI		§ Cited literature link § Cited literatures' full-text link	-
Domestic academic DB	KoreaScience		§ Reference list provided	§ NDSL link provided § CrossRef provided (including other firms)
	NDSL		§ Reference list provided	§ Google Scholar
	DBPIA		§ Reference list provided	§ Reference (identified by UCI) link (limited)

3) System construction

Because portal sites have begun to provide citation DB, academic DB, and even citation analysis, broadly integrated linking is required including open access sites, webpages of researchers and projects, commercial DB, and journal sites providing full-text service. Most importantly, greater effects can be expected if the reference linking service system is built based on the OPAC of digital libraries not only because many books are currently being digitized but also because this system will provide easy access to information sources through metadata structures incorporated into the OPAC. Therefore, full-text linking service can be implemented in domestic libraries like the full-text linking service based on SFX, which is expected to contribute to increased awareness of digital libraries and their offerings.

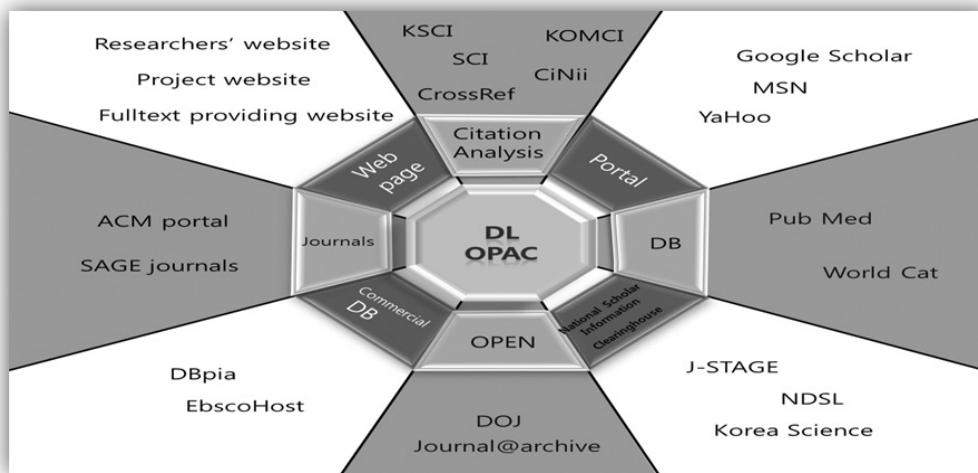


Fig. 12. Prediction of materials for reference linking service

As shown in the <Figure 12>, to create linking to full-text in various institutions for digital libraries, first, a linking system for all of the materials in the digital libraries should be built. For this purpose, a strategy to link full-text and bibliographic information should be established based on analysis on every material in digital libraries, and the strategy should consider integrated access to various media including physical books (ISBN, ISSN), digital journals (ex DOI), Internet resources, and video materials.

Second, it is impossible even for large libraries like a public digital library to retain every material's full-text, thus a strategy for Open Access material collection and application development is necessary. Digital full-text integrated search and a linking system should be developed along with building cooperative systems with foreign Open Access journal service agencies. It is necessary to get authority for free access to foreign Open Access journals through collaboration with PMC and DOAJ.

Third, integrated linking service of full-text is provided by many DB including ScienceDirect, SAGE Journals online, PNAS online, Google Scholar, CiteSeerX, CiNii, and ACM Portal. Through the integrated linking service on full-text DB, efficiency in material use in the digital libraries

and users' satisfaction level will increase, thereby improving the status of domestic digital libraries.

5. Conclusion and Future Research

With the rapid development of information technology, information resources are increasingly being published exclusively in digital formats and existing information sources are rapidly being converted. Users want easy access to full-text and use of similar literatures to retrieved resources online. In this regard, information providers including publishers make efforts to generate value-added products and improve user satisfaction levels through full-text and linking services. As a representative service, reference-linking services link abstract, full-text, and bibliography of cited references using an identification system, which improves connections by applying link information instead of URLs.

However, material types referenced in a single paper can include journals, books, dissertations, reports, dictionaries, encyclopedias, and various types of information resources that are not reflected in current linking systems, which focus on papers in academic journals. Considering the increased production and use of reports, proceedings, and e-books, a strategy to include these literatures in the reference linking service should be sought. This study gives some suggestions on a system enabled to include various information resources such as books or dissertations. As essential principles and pre-requisites in building the reference linking system, expansion of resource and media type, development of identifiers for all types of information resource, linking to the existing DBs, publication revolution and controlling copyright holders, use and enhancement of OPAC, and perfect authority DB construction were presented in the above sections. In particular, building a reference linking system based on the OPAC of the digital libraries was suggested, because it will lead to easy access to information resources outside libraries from metadata structures incorporated into the OPAC since libraries have become more digitized.

On the other hand, a publication revolution is required to build the suggested system. Currently, large foreign publishers are checking the accuracy of cited references and linking potentials when the paper review process is finished, meaning that linking cited references' full-text is relatively easy for them after such a verification process. Given that, domestic publishers and academies here in Korea should make efforts to secure a strict linking reference system in the publishing phase. Linking to reports, books, and dissertations also requires efforts in the publishing phase for accurate reference linking. Therefore, follow-up studies should seek reference linking methods depending on each publication format.

References

- Arms, W., Bergmark, D., & Lagoze, C. (2000). An architecture for reference linking. (Technical report, TR 2000-1820). Cornell University, Digital Library Research Group.
- Bergmark, D. (2000). Automatic extraction of reference linking information from online documents. (Technical Report. TR 2000-1821). Cornell Digital Library Research Group.
-

- Bide, M. (1997). *In search of the Unicorn*. London: Book Industry Communication, BNBRF 89. Available at <<http://www.bic.org.uk/bic/>> Retrieved 2009.08.01.
- Blake, M. E., & Knudson, F. L. (2002). Metadata and reference linking. *Library Collections, Acquisitions, and Technical Services*, 26(3), 219-230.
- Choi, S. B., Park, Y. W., Hing, S. W., & Kim, K. H. (2000). A Study for the link service system of science & technology information. *The Journal of the Korea Contents Association*, 7(1), 731-736.
- CrossRef System specification. Retrieved from <http://doi.crossref.org/doc/SystemInterface.html>
- DBpia About. Retrieved from <http://www.dbpia.co.kr/>
- Gang, M., Lee, S., Park, J., Choi, S., Shin, J., Choi, M., & Park, B. (2006). *Development and Linking of National Science & Technology Information Resources*. Seoul: Korea Institute of Science and Technology.
- Grogg, J. E. (2006). The development of context-sensitive linking. *Library Technology Reports*, 42(1), 14-23.
- Groza, T., Grimnes, G. A., & Handschuh, S. (2012). Reference information extraction and processing using conditional random fields. *Information Technology & Libraries*, 31(2), 6-20.
- Jung, T., Park, D., Kim, J., Seok, J., Lee, S., Cho, S., & Kim, M. (2004). *Development of Linking Applying Technology for the Digital Information* (K-04-IS-02-00S-4). Seoul: Korea Institute of Science and Technology.
- Kim, H., & Joung, K. H. (2005). An analysis on the operations of reference databases in Korea. *Journal of Korean Society for Information Management*, 22(2), 23-39.
- Kim, S. H. (2005). The comparison and analysis of linking system using OpenURL. *Journal of the Korea Society for Information Management*, 22(4), 221-234.
- Korea Citation Index. Retrieved from <https://www.kci.go.kr/>
- Kwon, L. N., & Kim, J. S. (2007). A study on the establishment of reference linking system for national R&D information. *The Journal of the Korea Contents Association*, 8(1), 195-202.
- Kwon, O. J., & Noh, K. R. (2003). A study on implementation of patent fulltext linking system based on patent DOI. *The Journal of the Korea Contents Association*, 1(1), 319-322.
- Lee, J., Yu, S. Y., Im, D. H., Jeong, S. J., & Lee, J. Y. (2009). Strategy for scholarly information service using citation information. *Proceedings of the 16th Conference of Korea Society for Information Management*, 93-98.
- Lee, S., Choi, S., Cho, S., Jhun, S., & Chung, T. (2005). Design and implementation of an open academic information dissemination architecture based on OAI and OpenURL. *Korea Computer Congress 2005*, 32(1), 670-672.
- Lee, S. G., Kim, S. T., Lee, Y. S., & Yi, T. S. (2007). Research and development of citation matcher for reference parsing and cross-reference linking. *The Journal of the Korea Contents Association*, 5(2), 426-429.
- Lee, S. H., Sin, D. G., Kim, J. S., Choi, J. Y., & Jung, T. (2004). Design and implementation of a digital contents service system based on the identification system. *Journal of the Korea Society for Information Management*, 21(3), 15-29.
- Lee, S. H., Sin, D. G., Kim, J. S., Choi, J. Y., & Jung, T. (2004). Design and implementation
-

- of a digital contents service system in science and technology field based on identification system. *The Korean Institute of Information Scientists and Engineers*, 31(1), 415-417.
- Lee, Y. S., & Lee, S. G. (2010). The reference identifier matching system for developing reference linking service. *Journal of Information Management*, 41(3), 191-209.
- Li, J., Burnham, J. F., Lemley, T., & Britton, R. M. (2010) Citation analysis: Comparison of Web of Science, Scopus, SciFinder, and Google Scholar. *Journal of Electronic Resources in Medical Libraries*, 7(3), 196-217.
- McCarthy, D., & Petit, G. (2003). IERS Conventions. *International Earth Rotation and Reference Systems Service (IERS Technical Note)*, 32, 1-172.
- NHN (Naver). Retrieved from <http://academic.naver.com/>
- Soichi, K. (2005). *How Japan Science and Technology Agency uses DOIs*. Japan Science and Technology Agency (JST), 1-42.
- Van de Sompel, H., & Hochstenbach, P. (1999). Reference linking in a hybrid library environment: Part 3: Generalizing the SFX solution in the 'SFX@ Ghent & SFX@ LANL' experiment. *D-lib Magazine*, 5(10), 1-42.
- You, B. J., & Lee, E. B. (2002). A study on the construction of meta database and management system of the foreign technical information. *Journal of the Korean Society for Library and Information Science*, 36(3), 201-214.
-