

# 국내 학술지 논문의 DOI 기반 연계시스템 구축에 관한 연구

한혜영(한국교육학술정보원)

## 1. 서 론

### 1.1 연구의 목적

최근 해외의 주요 학술지 출판사들은 자체 출판저널의 일부 혹은 전체를 전자저널(electronic journal 또는 e-journal)의 형태로 웹(World Wide Web: WWW)상에서 제공하기 시작하였다. 또한 최근 등장한 웹의 하이퍼텍스트 기술은 학술지 논문을 읽을 때 발생할 수 있는 인용 연계, 구/단어/문장 연계, 참고문헌 연계 등 다양한 종류의 연계를 가능하게 하는 '하이퍼텍스트 저널'(hypertext journal 또는 h-journal)의 개념을 이끌어내었다. 전자저널보다 진보된 하이퍼텍스트 저널의 개념은 정보를 담고 있는, 혹은 정보에 대한 포인터를 제공하는 소스 노드와 목적 노드를 이용자에게 연결하는 연계라는 기본적 하이퍼텍스트 단위를 포함한다.

여러 출판사간, 동종 혹은 이종의 학술지간 연계는 특정 정보제공기관의 한정된 자원의 효율을 극대화시킬 수 있는 한 방법으로써, 가치있는 학술지에 대한 각종 인용문헌으로의 연계를 지칭하는 '참조연계'(reference linking)는 점차 전자저널 제공서비스 환경의 필수 요소로 여겨지고 있다.

국내 대학도서관, 학회 및 기타 정보서비스 기관은 해당기관이 생산한 저작물 혹은 보유자료에 대한 색인과 원문을 자체적으로 구축하여 서비스하고 있다. 이들 개별기관이 구축한 자료는 서로 다른 시스템에 분산되어 있으므로, 이용자는 네트워크를 통하여 개별기관 웹 사이트를 방문하여 여러 번의 검색을 통하여 원문을 찾는 불편함을 겪는다. 검색 결과로 얻은 원문의 내용과 연관된 정보를 얻는 해결방안은 현재 제시하고 있지 않으므로 참조연계 서비스를 이루기 위해서는 인터넷 자원의 효과적인 통합 관리가 필요하다.

따라서 본 연구의 목적은 다음과 같다. 첫째, 전자화된 저작물간 연계의 개념, 연계를 제공하는 수단, 그리고 연계유형을 조사한다. 둘째, 국내 분산된 학술지 논문의 통합 검색, 원문의 식별 그리고 원문간 참조연계 서비스에 적합한 모델을 연구한다. 셋째, 이를 바탕으로 국내 원문연계 서비스를 제공하기 위한 중앙집중식 모델인 '학술지 논문 연계시스템'(E3R/LS: Electronic Research Resource Reference Linking System)을 구현한다. 본 시스템을 구성하는 요소로 URN(Uniform Resource Name)의 일종인 DOI와 DC를 이용한 메타데이터 형식을 정의하며, 이를 이용한 메타데이터 구축과 원문 검색, 식별 그리고 참조연계가 이루어지는 과정을 설계한다.

## 1.2 연구의 범위와 방법

본 연구에서는 이용자가 학술지 논문 연계시스템에서 제공하는 서비스를 통하여 원하는 학술지 논문에 대한 접근, 원문 획득과정 그리고 학술지 논문내 인용된 원문으로의 연계방안을 도출하기 위하여 관련된 기술과 사례를 분석하였다. 관련 문헌연구에서는 연계에 관한 전반적인 개념, 학술지 논문 연계시스템 구축시 필요한 구성요소를 조사하였으며, 연계에 필요한 구성요소로서 디지털 형식으로 표현된 정보자원을 식별할 수 있는 저작물의 식별자, 연계서비스를 가능하게 하는 핵심 요소인 메타데이터, 그리고 식별자를 위치정보로 변환하고 선별하는 체제에 대하여 고찰하였다.

국내 분산된 디지털 자원에 적절히 활용 가능한 모델로서 URN의 한 종류인 DOI를 이용, 학술지 논문 연계시스템에 대한 전체적인 구성도와 상세한 내용 및 절차를 연구하였다. 중앙집중식인 E3R/LS에는 원문에 관한 서지정보를 기술한 Dublin Core 형식의 메타데이터 정보를 구축하여 통합 검색이 가능하도록 하고 원문의 위치정보를

확인할 수 있는 식별자 정보를 제공하였다. 또한 원문내의 참고문헌은 인용된 원문으로의 연계가 이루어질 수 있는 PDF(Portable Document Format) 형식으로 구축하여 하이퍼링크가 가능하도록 하였다.

본 연구에서는 시범적으로 서로 다른 두 개의 서버에 저장된 대한수학회 발간 *Communications of the Korean Mathematical Society*와 한국문헌정보학회 발간 『한국문헌정보학회지』에 게재된 논문을 대상으로 시스템을 구현하였다.

## 2. 원문 연계서비스

### 2.1 연계의 개념 및 도구

연계에 대한 이론적 근원은 Bush(1945)의 연상 색인(associative indexing) 모델인 Memex 개념에서 출발한다. 인간의 연상 작용(Hitchcock et al. 1997)은 하나의 정보로부터 다른 정보로 이동하는 원리와 유사하며 이와 같은 개념을 컴퓨터에 도입한 것이 하이퍼텍스트 또는 하이퍼미디어 시스템이다.

최근 인터넷 환경에서 검색시 디지털 정보 대상물의 위치를 찾아내고 원문으로 연계하기 위한 도구로서 URL(Universal Resource Locator)을 사용한다. URL은 일종의 위치식별자로서 대상물을 발견하고 특정 서비스로 이동할 수 있는 주소와 매개변수를 포함하고 있는 단순한 인터넷 주소 체계이다. 그러나 URL 정보는 시스템 주소 변경 혹은 대상물의 위치가 변경되면 더 이상 존재하지 않는 URL을 가리키게 되므로 계속적으로 관리하지 않을 경우 이전 URL의 이용자들은 원하는 대상물을 얻을 수 없게 된다. 디지털 대상물에 대한 영구적인 식별자로서의 기능을 URL이 제대로 수행하지 못하게 됨에 따라 영구식별자 역할을 할 수 있는 IETF(Internet Engineering Task Force)의 URN과 같은 체계가 필요하게 되었다(Sollins 1994; Moats 1999).

URN은 인터넷 자원에 부여된 고유 식별기호로 다음과 같은 기본적인 특징을 가지고 있다(Paskin 1999).

- 포괄성 : URN은 하나의 시스템이나 특정 위치를 의미하지 않고 어느 곳에서도 동일한 의미를 갖는다.

- 유일성 : 동일한 URN이 결코 두 개의 다른 자원에 부여되지 않는다.
- 영구성 : URN은 영구적으로 유일해야 하며, URN이 식별하는 자원이 네트워크상에 존재하는 동안, 혹은 그 이름을 부여한 명명 기관이 존재하는 기간 동안 자원에 대한 참조정보로서 사용될 수 있다.
- 규모성 : URN은 네트워크 상에서 이용 가능한 어떠한 정보에 대해서도 부여할 수 있다.
- 기존시스템 지원 : URN 구조는 기존의 명명시스템을 지원해야 한다.
- 확장성 : URN 체계는 추후에도 확장할 수 있어야 한다.
- 독립성 : 이름을 만드는데 있어 명명기관의 독립성을 보장한다.

URN 사양에 따른 대표적인 구현 사례로 OCLC의 PURL(Persistent URL)(<http://purl.oclc.org/OCLC/PURL>)과 IDF의 DOI(<http://www.doi.org/>)가 있다. PURL은 일반적인 인터넷 자원의 명명(naming)작업과 변환서비스를 위해 OCLC에 의해 개발되었으며 인터넷 자원의 위치로 직접 이동하는 URL과는 달리, PURL 서버에 의한 중간 변환과정을 통하여 클라이언트에게 URL 정보를 알려준다.

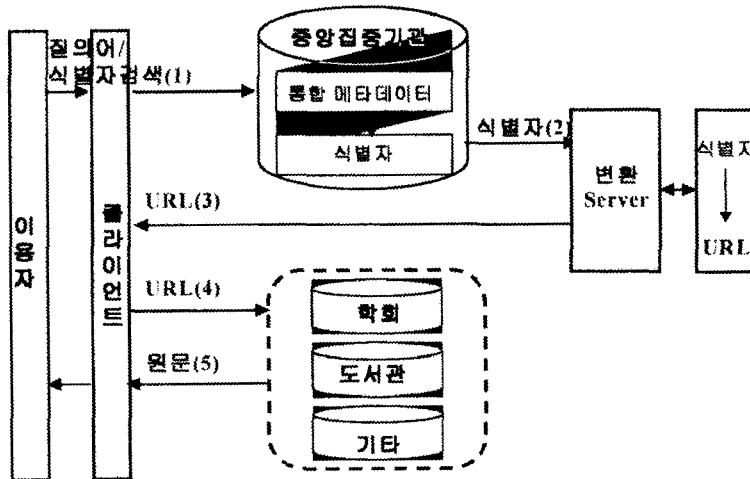
DOI는 모든 디지털 내용을 위한 식별자로서 고안되었다. URN 및 PURL과 같이 DOI도 유일한 영구적인 식별자로서 활용 가능성은 물론 추가적으로 기존 인쇄물 식별자와 기타 신규 식별자를 모두 포용하여 사용할 수 있으므로 저작권자와 이를 이용하는 이용자를 연결할 수 있다. DOI는 미국 출판협회와 CNRI(Corporation for National Research Initiative)에 의해 개발되었고 현재 많은 관련기관에서 채택하고 있는 URN의 한 방식이다(International DOI Foundation 1999).

## 2.2 연계의 유형

연계를 성공적으로 수행하기 위해서는 연계에 참여하는 출판사 혹은 원문소장기관들간 학술지 논문에 관한 서지정보의 공유, 연계정보의 정확성 유지, 그리고 이용자에게 연계기능을 제공하는 인터페이스가 필요하다. 이를 위해서는 원문소장기관들간에 공통적으로 활용할 수 있는 연계 도구, 공동 활용할 수 있는 연계 방법, 연계에 필요한 법적 동의가 필수적이다. 연계서비스에 참여하는 기관이 권한을 가지고 특정 범위 내에서 연계를 제공하는 권한부여는 정보의 경제적 문제와 더불어 이용제한에 관한

측면을 함께 고려해야 한다. 만약 연계에 대한 권한 여부가 필요한 경우, N개의 기관 사이에 완전한 양방향 연계를 위해서는  $N \times (N-1)$ 개의 연결 서비스에 대한 사전 동의가 있어야 한다. 중앙집중식 연계인 경우에는 전체 N개의 기관에 대하여, 중앙집중기관과 나머지 기관 사이 (N-1)개의 사전 연계 동의가 필요하다. 그것은 <그림 1>에서와 같이 이용자가 중앙집중기관에 검색을 보낸 후 검색 결과인 원문을 학회, 도서관, 또는 기타 기관으로부터 받기 때문에, 이들 기관으로부터 사전 연계 동의를 필요로 하는 것이다.

<그림 1>의 중앙집중식 연계 모델에서 이용자는 중앙집중기관의 통합 메타데이터 베이스를 대상으로 질의어 혹은 식별자(1) 검색을 행한다. 시스템 내부적으로 식별자의 변환과정을 거친 후(2) 원문이 있는 URL(3) 주소를 제공하면, 클라이언트는 획득한 URL(4)상에 존재하는 원문(5)을 이용자에게 제공하는 형태로 전체 시스템은 구성되어 있다. 여기서 이용자는 질의어나 식별자 검색을 통하여 곧바로 디지털 원문을 얻을 수 있다. 중앙집중기관은 사전에 각 기관으로부터 메타데이터 정보를 수집한 통합 메타데이터를 보유하고 있다.



<그림 1> 중앙집중식 연계 모델

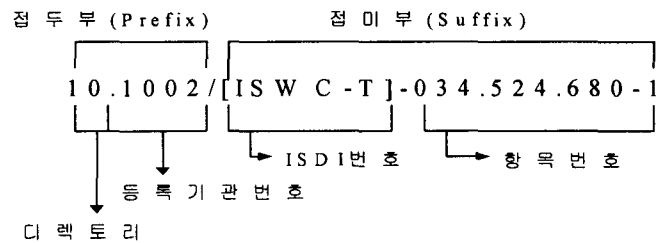
### 3. 연계시스템 구성 요소

#### 3.1 식별자

##### 가. DOI

현재의 모든 시스템들은 특정 아이템을 식별하기 위한 기본키를 요구하므로 식별자는 참조연계에 반드시 필요한 항목이 된다. 디지털 환경에서 대두되는 대표적 식별자인 DOI를 참조연계를 위한 식별자로서 활용하였다.

미국 출판협회는 전자적 환경에서 디지털 자원의 저작권 관리와 식별을 위한 개념으로 DOI를 제안하였으며 1997년 10월 프랑크프르트 책 박람회에서 CNRI의 Handle System(Sun and Lannom 1999)으로 첫 선을 보였다. DOI는 '접두부'(prefix)와 '접미부'(suffix)로 구성된 번호 체계를 가지며 접미부는 개별 등록기관(registrant)의 책임 하에 부여된다. <그림 2>는 DOI 구성요소인 접두부와 접미부에 대한 사례이며, 전체 DOI 번호의 유일성을 보장하기 위하여 등록기관은 자체 자료에 대한 DOI 접미부 번호를 일정한 원칙 하에 부여해야 한다(Lynch 1997).



<그림 2> DOI 예

DOI의 장점은 영구적인 식별자로 활용 가능하다는 것이다. 즉, DOI는 한번 등록되면 결코 인터넷상에서 사라지지 않는다(Davidson and Douglas 1998). 만약 출판사가 자체 웹 서버를 재구조화 하거나, 새로운 내용의 관리 시스템을 설치할 경우, 혹은 다른 방법으로 내부 조직을 변화시키면 변경된 URL 정보만 DOI 서버에서 갱신하면 된다. 즉, DOI는 출판사가 다른 기관에 해당 정보의 소유권을 팔거나 권한을 이전한다고 해도 변화하지 않는 장점이 있다.

DOI의 접미부는 대상물 식별기호 부분으로, 기존 식별체계인 BICI(Book Item and Contribution Identifier), ISRC(International Standard Recording Code), ISSN, ISWC(International Standard Work Code), ISBN, SICI 등과 같은 국제표준을 선택하여 사용할 수 있으므로 ISDI(International Standard Digital Identifier)라 한다.

본 연구에서의 DOI 번호 부여 방식은 다음과 같다.

#### 나. E3R/LS의 DOI 번호

##### (1) DOI 구분

본 DOI 구분은 NISO(National Information Standards Organization)에서 정의한 ANSI/NISO Z39.84-XXXX(NISO 1999)<sup>1)</sup>에 기반하여 정의하였다. 현재 이름부여기관 코드(Agency Code)로 유효한 값은 '10' 이며 DOI 스트링의 최대 길이는 128글자(UTF-8 코드)로 대소문자를 구분하여 처리한다. 본 연구에서 사용할 DOI 번호 매김 형식을 BNF(Backus-Naur Form)로 표시하면 <그림 3>과 같다.

```

E3R/LS_Number ::= <DIR>.<REG>"/"<ISDI>
    <DIR> ::= *(<이름부여기관>)
    <REG> ::= UTF-8로 표현된 모든 글자
    <ISDI> ::= (ISBN | ISSN | BICI | SICI ) <ItemNo>+
    <ItemNo> ::= 한글,숫자,영어를 포함한 UTF-8로 표현된 모든 글자
    
```

<그림 3> E3R/LS의 DOI 구분

<DIR>은 현재 '10'으로 정의된 DOI 변환기 한 개밖에 없으므로 '10'으로 표기된다. <REG>는 CNRI에 등록하면 배당되는 등록기관 번호를 사용하며 <ISDI>는 다음과 같이 정의한다. <ISDI>는 전통적 인식번호인 ISBN, ISSN, BICI, 혹은 본 논문에서

---

1) ANSI/NISOZ39.84-2000로 책정됨(NISO. 2000. *Syntax for the Digital Object Identifier: An American National Standard Developed by the National Information Standards Organization*. Approved May 12, 2000 by the American National Standards Institute.)

정의하는 확장형 SICI 형태의 체계를 가진다. ISSN을 사용하는 경우 예로 『한국문헌정보학회지』는 '[ISSN]1225-598X'이며 한국문헌정보학회지 1999년도 33권 4호 113쪽에서 시작하는 '이용남, 홍현진'의 '우리나라 공공도서관에 대한 평가지표 연구'라는 논문의 SICI는 [SICI]1225-598X(1999)33:4<113:우공대>2.0.CO;2-A가 된다. 본 연구에서는 학술지 논문에 관한 연계만을 다루므로 <그림 3>의 표기방법 중 ISDI는 SICI의 번호 체계만을 적용한다.

## (2) ISDI : SICI

기존 연속간행물의 식별자는 연속간행물 서명 단위, 권호 혹은 게재논문을 식별할 수 있는 코드로 구분된다. SICI는 학술지 논문의 식별자로서 연속간행물 출판물의 어떠한 구성 요소도 식별할 수 있으며 ANSI NISO 표준 Z39.56으로 개발되었고 1996년 새로이 개정되었다.

SICI는 영문자 문화권에서 개발(NISO 1996)되어, 국내 학술지에 적용하기 위해서는 새로운 확장된 개념의 구문법 개발이 필요하다. 또한 게재논문의 서지정보 혹은 서지정보를 기술한 메타데이터로부터 자동으로 SICI를 구축하기 위해서는 별도의 SICI 구문법이 정의되어야 한다. 그것은 영문(ASCII 코드) 환경에서 정의된 SICI가 한글(KS C 5601 또는 Unicode 2.0) 사용시 새로운 문제를 발생시키기 때문이다. 본 연구에서는 국내 학술지 논문에 대한 SICI 코드에는 한글 서명코드를 부여하여 SICI의 각 구성요소를 확장된 구문법으로 재정의하여 사용하였다(한혜영 1999).

## 3.2 메타데이터

### 가. Dublin Core(DC)

연계서비스를 위한 핵심요소로 메타데이터가 있다. 메타데이터는 데이터에 대한 구조화된 데이터로 웹상의 전자자원의 내용과 위치를 기술하며 네트워크상에서 효과적인 탐색과 검색을 가능하게 하는 데이터 요소를 포함한다(Xu 1997). 메타데이터의 데이터 요소에 식별자를 공통적으로 포함하므로, 메타데이터는 원문에 관한 일반 서지정보와 식별자 정보를 포함하여 서지정보 검색을 가능하게 하고 검색된 정보 중 원하



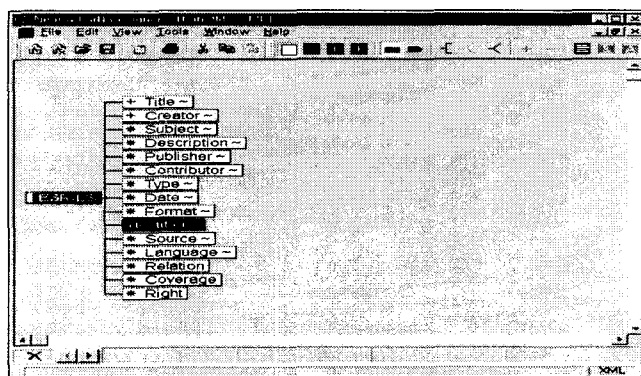
는 원문과의 연계를 제공할 수 있게 한다. 본 연구에서는 메타데이터로서 널리 활용되고 있는 DC를 선정하여 본 연구의 요건에 맞게 수정하여 사용하였다.

본 연구에서는 DOI 1 수준인 단순 변환시스템을 활용하므로 대상물에 관한 식별자를 정확히 알고 있다면 중앙집중기관의 메타데이터 데이터베이스(MDDB)를 검색하는 과정없이 곧바로 DOI 변환기를 통하여 관련 원문이 제공되는 '원문저장소'의 웹 사이트로 이동할 수 있다. 하지만 DOI를 하나의 요소로 하는 메타데이터를 기술하여 중앙집중기관의 MDDB에 구축하는 이유는 메타데이터의 내용으로 이용자가 찾고자 하는 아이템이 어떤 것인가를 식별하도록 하는 충분한 정보를 제공하여 실제 아이템에 대한 물리적인 조사없이도 유일한 식별자로서 활용할 수 있게 해야 하기 때문이다. 원문을 제공하는 기관의 경우도 중앙집중기관의 MDDB에 메타데이터를 구축하는 것이 이용자들에게 성공적인 참조의 기회를 줄 수 있으므로 메타데이터를 MDDB에 등록하는 것이 유리하다. 본 연구에서 사용한 메타데이터는 DOI 열람 기능을 제공하므로 비교적 간단하게 기술하였다.

#### 나. DC 메타데이터의 요소

본 연구에서는 학술지 게재논문의 데이터 요소로서 논문의 식별이 가능하고 SICI의 각 구성요소와 일치하는 항목을 선정하여 SICI 코드가 없는 경우 자동 생성도 가능하도록 지원하는 형식을 제안하였다. 즉, 학술지 서명수준 뿐 아니라 권호 혹은 게재논문 단위 데이터 요소까지 포함할 수 있도록 구조화하고 개별 학술지 논문에 관한 대부분의 질의에 답변할 수 있는 데이터 요소 중 최소한의 세트로 구성하였다. 구성 내용은 Caplan과 Arms(1999) 그리고 Martin과 Bide(1997)를 기초로 식별자와 출판사항목을 추가하였으며 데이터 요소는 학술지 논문명(표제), 학술지 논문의 저작자, 학술지명(표제), 학술지 논문의 출판년/월/일, 학술지 논문이 포함된 특정 권호의 번호표시, 학술지 논문의 시작/끝 쪽 위치, 자료 유형(학술지 논문, 초록, 권호 등), 식별자, 그리고 출판사이다.

이 최소한의 데이터 요소를 바탕으로 본 연구에서 정의한 DC 1.0에 대한 문헌유형 정의(Document Type Definition)는 <그림 4>와 같으며 DTD 대상은 국내 학술지 논문의 서지정보로 한정하였다.



<그림 4> Near & Far® Designer. Version 3.0으로 작성한 DC Metadata DTD

### 3.3 DOI 변환기

IETF는 단순한 주소 체계만을 제공하는 URL을 대체하기 위하여 완전하고 영구적인 인터넷 자원 식별구조로서 영구적인 이름을 활용하기 위한 표준 URN을 제안하였고 하나의 URN은 자원이 저장된 주소로 변환된다. 변환시스템은 식별자를 저장하고 식별자를 변환하는 네트워크상의 서비스로 하나의 식별자를 하나 이상의 URL로 변환한다.

<그림 1>과 같은 중앙집중식 참조연계 모델에서 변환서버는 통합 메타데이터 데이터베이스의 검색 결과인 식별자 정보를 해당 원문의 위치정보로 변환한다. 본 연구에서 식별자를 자원의 위치정보로 바꿔주는 변환절차는 CNRI Handle 시스템을 이용하여 변환작업을 수행하였다.

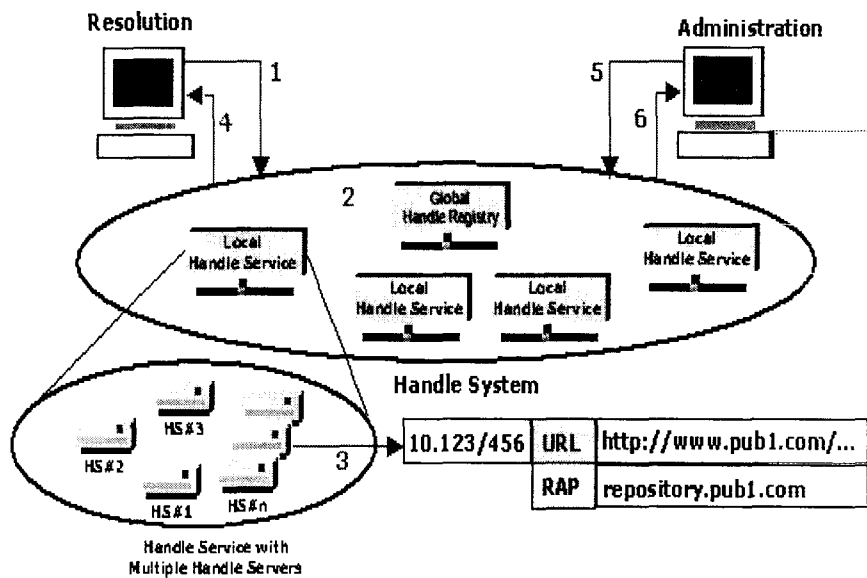
#### 가. 핸들시스템

핸들시스템(Sun and Lannom 1999)은 디지털 자원 혹은 인터넷상의 자원에 대한 영구적인 식별자인 핸들을 부여·관리·변환하기 위한 총체적인 시스템이다. 핸들은 URN으로서 사용되며 핸들시스템은 오픈 프로토콜, 이름공간, 그리고 프로토콜의 수행을 처리한다. 핸들시스템에서 수행되는 프로토콜은 분산 컴퓨터시스템에 디지털 자원의 핸들을 저장하고 이 핸들을 이용하여 디지털 자원의 위치확인 및 접근에 필요한

정보로 변환할 수 있게 한다. 변환과 관련된 정보에 변경이 필요한 경우 핸들은 변경하지 않고도 식별하고자 하는 자원의 현재 상태를 반영할 수 있다. 그리하여 찾고자 하는 아이템의 이름은 정보의 위치 및 상태 변화와 무관하게 영구적으로 사용할 수 있다.

핸들시스템은 핸들변환서비스와 핸들관리서비스, 두 종류의 서비스를 제공한다. 핸들변환서비스는 클라이언트가 요청한 핸들을 변환하고 그 값을 송부한다. 핸들관리서비스는 클라이언트가 요청하는 핸들의 추가, 삭제 혹은 갱신과 같은 핸들관리에 관한 사항을 처리한다.

<그림 5>는 핸들시스템의 기본 구조와 동작을 보여준다. <그림 5>의 (1)부터 (4)까지는 변환과정을 보여주며, (5)와 (6)은 핸들의 관리절차를 보여준다.



<그림 5> Handle System의 구조와 동작

자료: <http://www.ietf.org/internet-.../draft-sun-handle-system-def-01.tx>

#### 나. CNRI 핸들시스템

웹클라이언트 프로그램을 활용한 핸들 변환에서 클라이언트는 ‘핸들-인식’용 웹 브라

우저가 된다. 클라이언트 이용자는 중앙집중기관의 웹 화면에서 핸들을 만난다. 예로 핸들 구문법에 의거한 『한국문헌정보학회지』에 실린 ‘장혜란’의 ‘대학도서관 질적 평가를 위한 모형 개발’의 핸들은 ‘10.1064/[SICI]1225-598X(1997)31:4 <165:대질평>2.0.CO;2-D’로 DOI 변환서버내 핸들데이터베이스에 저장되어 있다. 핸들데이터베이스의 레코드는 효율적인 검색이 가능하도록 구축된다. 핸들데이터베이스에 저장된 핸들은 핸들관리데이터베이스에도 동일하게 저장되므로 해당 관리데이터와 연결하여 핸들이 누구에 의해 수정되었는지, 어떻게 처리되었는지를 보여준다.

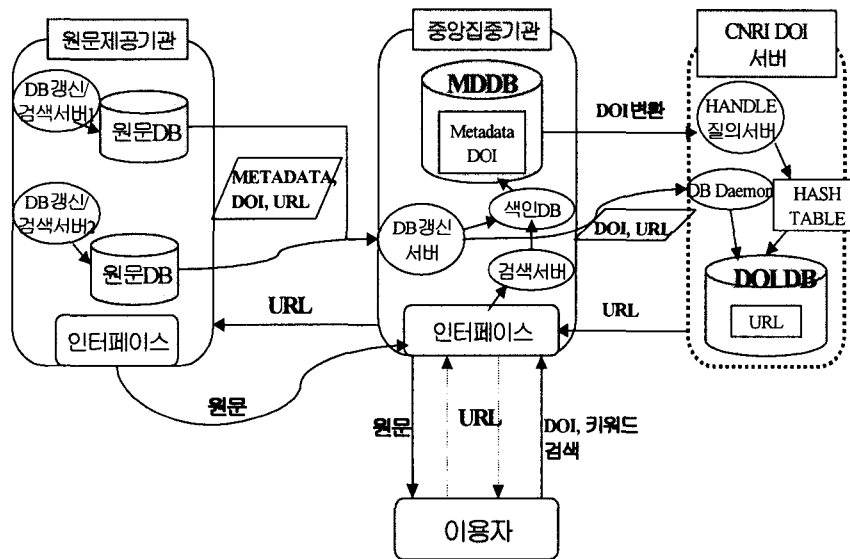
## 4. 국내 학술지 논문 연계시스템 구현

### 4.1 국내 프로토타입 시스템 구현

본 연구에서는 분산된 학술지 논문의 원문서비스 제공을 위한 검색 및 위치확인파 원문의 참고문헌내 인용된 특정 원문으로의 연계를 위한 시스템을 <그림 6>과 같이 구성하였다.

<그림 6>의 전체 구성도에는 원문제공기관, 중앙집중기관, CNRI의 DOI 서버, 그리고 이용자가 포함되어 있다. 이와 같은 시스템을 구성하기 위하여 우선적으로 원문제공기관에서는 학술지 논문을 PDF와 SGML의 두 가지 형태로 구축하여 물리적으로 분산된 두 개의 서버에 저장하였다. 또한 디지털 학술지 논문의 위치를 확인할 수 있는 SICI 기반 DOI 식별자 정의와 DC 기반 메타데이터를 정의하여 개별 원문에 대한 식별자 번호 및 메타데이터를 생성하였다. 원문에 관한 서지정보로서 식별자를 포함한 메타데이터는 원문의 URL 위치정보와 함께 중앙집중기관으로 송부하였다. 중앙집중기관은 메타데이터를 MDDDB에 구축하고 DOI와 URL 정보는 CNRI의 DOI 변환서버에 등록하였다.

디지털화된 두 개의 학술지 논문들간의 직접적인 연결을 제공하는 참조연계를 위하여 원문제공기관은 원문의 인용정보에 그 인용된 원문으로 연계할 수 있도록 표시하는 연계정보를 삽입하여야 한다. 본 연구에서는 『한국문헌정보학회지』의 게재논문을 한글 작업이 가능한 “훈Q PDF Version 1.0<sup>2)</sup>”으로 구축하고 원문내 연계정보를



<그림 6> E3R/LS 전체 구성도

SICI 기반 DOI 번호를 삽입하였다. PDF 화일에 Acrobat PDF에서 제공하는 ‘링크제공’ 기능을 이용, MDDB에 송부한 DOI 번호를 연계정보로서 삽입하였다.

실제 시스템 구성을 위하여 한국교육학술정보원 ‘학술정보서비스시스템(Research Information Service System)’의 장비를 이용하여 원문제공기관과 중앙집중기관, 그리고 이용자 인터페이스를 프로토타입으로 구현하였다. 학술정보서비스시스템 내에는 다양한 서비스를 위한 서버가 존재하고 있으며 그 중 웹서비스가 가능한 ‘RISS웹서버’(서버 1, <http://riss.keris.or.kr>)와 ‘DL서버’(서버 2, [http:// dl.riss4u.net:8888](http://dl.riss4u.net:8888) 혹은 <http://dl.keris.or.kr:8888>, 두 개의 URL을 혼용하여 사용함)를 이용하여 Communications of the Korean Mathematical Society와 『한국문헌정보학회지』의 원문 데이터를 두 개의 분산된 서버에 분리 저장하였다. 중앙집중기관의 MDDB는 서버 2에 구축하였고 이용자의 검색 화면을 웹에서 제공하도록 구현하였다.

- 2) 혼QPDF Version 1.0 활용하여 아래아 한글의 조합형 한글을 PDF 파일로 변환.  
Window 95 또는 98, Adobe Acrobat 4.0 환경에서 사용 가능.

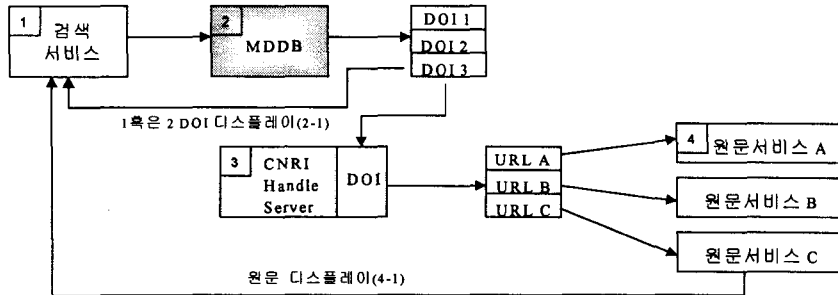
## 4.2. 실행 사례

본 장에서는 학술연구자인 이용자가 학술지 논문의 통합된 서지정보를 이용하여 원문을 획득하는 방법과 논문에 포함된 참고문헌의 정보로부터 인용된 원문을 찾아가는 방법에 대한 두 가지 시나리오에 근거하여 실행한 사례를 보여주었다.

### 가. 서지정보를 이용한 원문연계

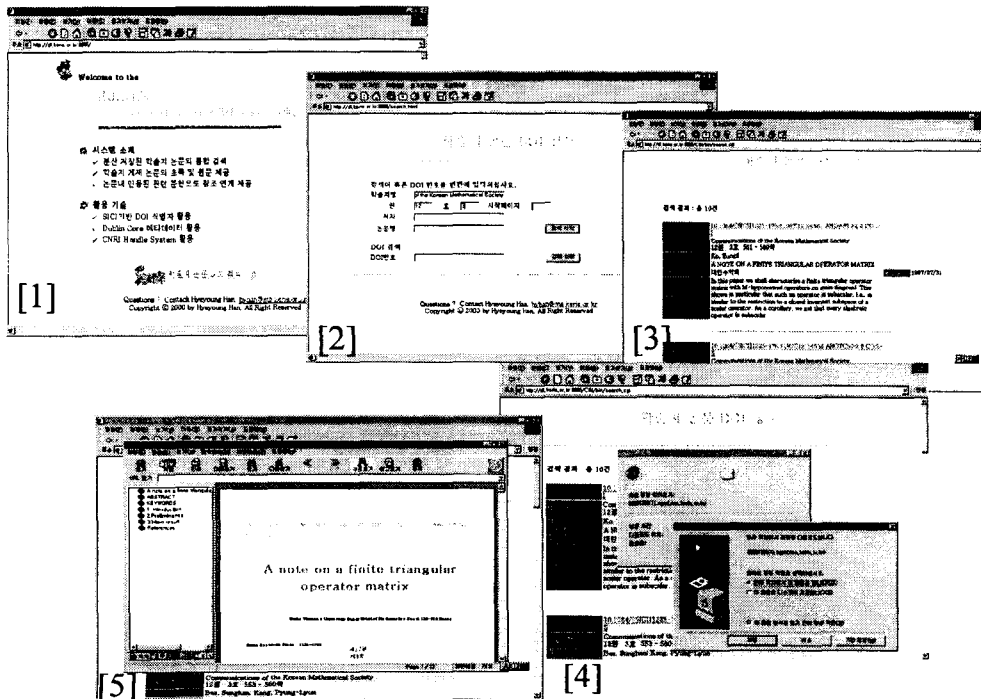
서지정보를 이용한 원문 연계방법은 연계사례 중 가장 기본적이고 단순한 정적인 방식을 사용하여 대학도서관, 학회, 혹은 정보서비스 기관들이 분산 소장한 원문에 대한 서지정보를 중앙에 통합하고 미리 계산되어 변환서버에 저장된 식별자를 이용하여 해당 기관의 위치정보인 URL을 파악하게 된다. 이 시스템의 구현을 위해 원문 구축 기관은 원문에 대한 서지정보를 메타데이터 양식에 의거하여 작성하고 메타데이터는 반드시 식별자 번호를 포함하도록 한다.

분산된 원문의 연계서비스 절차를 그림으로 표시하면 <그림 7>과 같다. 이용자는 검색서비스를 제공하는 중앙집중기관의 웹 화면을 통하여 MDDB를 대상으로 원하는 원문의 서지정보에 대한 메타데이터를 찾고자 한다(<그림 7>의 1번). 이용자 질의어와 일치하는 서지정보에 관한 메타데이터 정보는 식별자 정보와 함께 검색되어(2) 검색 결과가 이용자의 검색 화면에 디스플레이 된다(2-1). 검색 결과로 보여진 메타데이터 정보 중 이용자는 찾고자 하는 원문에 관한 서지정보의 식별자를 클릭하고 식별자 정보는 원문소장기관의 위치정보로 변환(3)하는 작업이 DOI 서버를 통하여 이루어진다. 그리하면 여러 기관에 분산 저장된 원문 중 해당 URL의 위치로 이동하여(4) 이용자에게 원하는 원문 정보를 화면에 보여주게 된다(4-1).



<그림 7> 분산된 원문의 연계서비스 절차

<그림 8>은 E3R/LS를 이용하여 실제로 서지정보를 이용한 원문연계를 실현한 예를 순서대로 보여주고 있다.

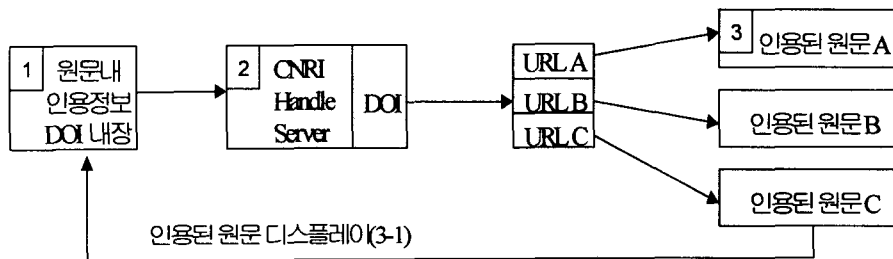


<그림 8> 서지정보를 이용한 원문연계 실현

나. 참고문헌의 인용정보로부터의 원문연계

이용자가 참고문헌내 인용된 원문을 보고자 하는 경우, 본 연구에서 구현한 E3R/LS에서는 인용문헌에 부여된 식별자 번호를 이용하여 DOI 서버와 연결, 변환작업을 거쳐 곧바로 인용된 원문 소장기관의 URL로 연결시킴으로써 원문을 화면에서 보여지게 된다.

<그림 9>는 원문내 참고문헌 인용정보의 식별자 정보를 이용하여(1) DOI 서버를 거쳐(2) 해당 URL로 이동하여(3) 원하는 인용된 원문으로 연계하는 절차를 보여준다 (3-1).

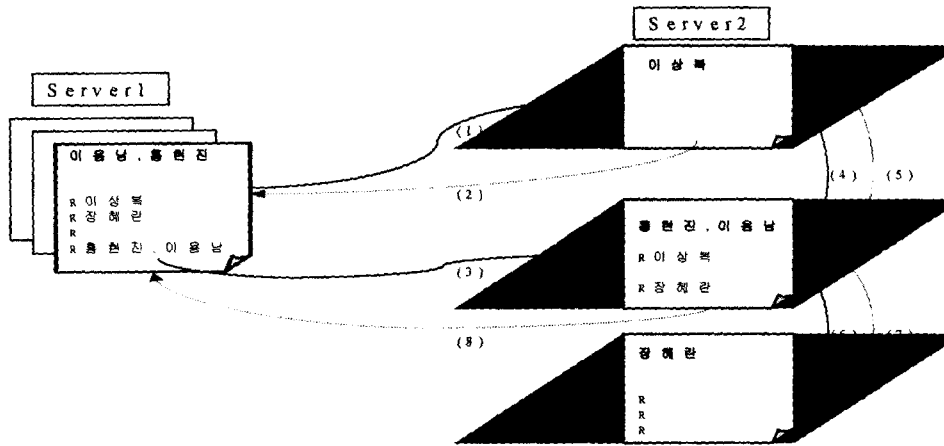


<그림 9> 원문내 참고문헌 연계서비스 절차

<그림 10>은 서버 1에 위치한 '이용남, 홍현진' 원문내 참고문헌으로부터 가능한 참조연계의 순서를 그린 것이다. <그림 10>의 (1)은 참고문헌 중 서버 2에 위치한 '이상복'의 원문헌으로 연계가 이루어지는 과정이며, '이용남, 홍현진'의 참고문헌으로 돌아와(2), 서버 2의 '홍현진, 이용남'의 인용된 문헌으로의 참조연계(3)를 만들 수 있다. 이러한 참조연계는 네트워크상에 시스템적으로 분산되어 있더라도 식별자 정보를 이용하여 영구적인 연계를 가능하게 한다.

'홍현진, 이용남'의 원문내 참고문헌에 '이상복'과 '장혜란'의 원문으로 이동할 수 있는 연계기능이 추가되어 있으므로 '이상복'(<그림 10>의 (4))와 '장혜란'(<그림 10>의 (6))의 인용된 문헌으로의 연계가 가능하다.





<그림 10> 참조연계 순서도

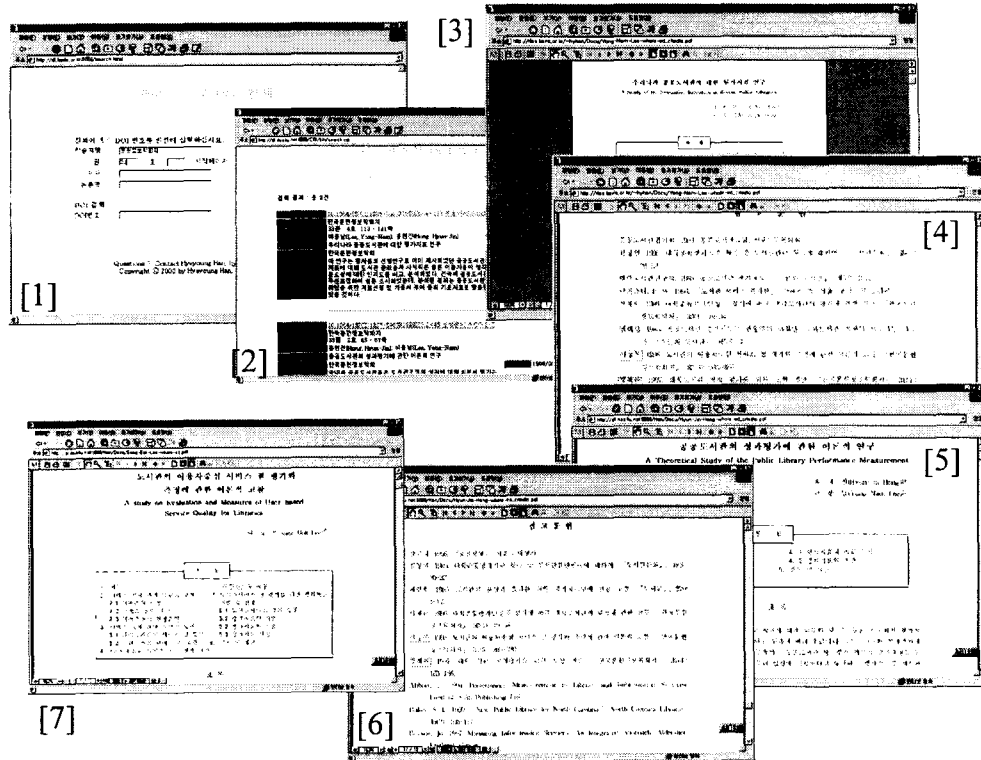
<그림 11>은 <그림 10>의 (1)~(4)까지의 과정을 E3R/LS를 이용하여 실연한 사례를 보여주고 있다.

## 5. 결론

본 연구는 국내 학회나 대학도서관 등에서 발간하는 학술지 논문의 디지털화 작업이 개별 기관별로 이루어지고 있는 상황에서 학술지 논문의 통합 검색과 한번의 검색으로 원문정보로의 접근까지도 가능하게 하는 방안을 제시하고 실제 시스템에 적용할 수 있는 프로토타입을 제안하였다.

프로토타입으로 개발한 E3R/LS은 중앙집중기관 시스템, CNRI DOI 서버, 원문제공 기관 시스템, 그리고 이용자 시스템으로 구성된다. E3R/LS는 시범적으로 분산된 시스템에 위치하고 있는 학술지 논문을 연계하고, 아울러 한글 PDF로 구축된 원문내 내장된 DOI 번호를 기반으로 참조연계가 가능한 기능까지 포함한 시스템으로 정의하였고, 구현된 시스템을 이용한 실연 사례를 보여 주었다.

인터넷상의 자원을 관리하기 위해 개발된 도구로서 URL과 URN은 인터넷상의 자원의 위치확인 및 식별 기능을 제공하지만 URL의 경우 해당 원문이 존재하는 네트워크상의 위치정보가 영구적이지 않다는 단점이 있다. 또한 자원의 이름으로 인터넷상



<그림 11> 참고문헌의 인용정보로부터의 원문연계 실현

의 자원을 확인하고 분산처리를 관리하는 URN은 최근 등장하는 지적재산권의 개념, 접근 권한 등과 같은 자원의 지적인 내용 관리의 식별자로서는 충분하지 않다. 1997년 등장한 DOI는 네트워크상에 존재하는 디지털 자원을 확인하는 기능과 디지털 자원의 지적인 내용을 전자상거래를 통하여 거래할 수 있는 기반 식별자이며, 위치 기반 방식을 사용하는 URL과는 달리 이름 기반 변환방식을 사용하므로 웹 상의 특정 위치정보에 의존하지 않는다.

DOI는 각 문서를 분리할 수 있는 기능을 수행하므로 본 연구에서는 주요 연구 대상인 원문의 식별과 원문의 위치확인을 위하여 DOI를 선정·활용하였다. DOI 접미부는 기존의 다양한 식별자를 포함하여 사용할 수 있으므로 연속간행물 식별용 국제 표준인 SICI를 국내 학술지 논문에 적합하도록 확장·개발하여 적용하였다. SICI는 학술지, 학술지 논문, 혹은 학술지 논문내의 하위 단위인 초록, 표, 그림, 절 등의 식별이 가능하고 DOI는 모든 지적저작물을 하나의 식별 단위로 인정하므로 본 연구에서 활

용한 학술지 논문의 상하 관계를 갖는 항목에 대한 식별 기능도 추가할 수 있다.

<표 1>는 URL과 DOI를 이용한 원문연계 및 참조연계를 비교한 것이다. URL을 이용한 연계방법은 URL을 문서 혹은 서지정보에 직접 삽입하므로 URL의 잦은 변경은 원하는 원문정보를 찾을 수 없게 할 뿐 아니라 문서와 서지정보에 삽입되어 있는 URL 정보를 찾아내어 갱신해야 하는 관리의 문제를 발생시킨다.

반면 영구적이고 고유한 식별자인 DOI 번호는 변환기를 통하여 위치정보로 변환되므로 위치정보가 변경될 경우 변환 서버내 데이터베이스의 URL을 갱신하면 되므로 위치정보의 일괄처리 및 중앙집중관리가 가능하다. \_

또한 인용정보에 URL을 삽입한 연계방식은 인용 문헌의 원문 위치가 변경될 경우 삽입한 URL을 매번 수정해야 하는 관리상의 어려움이 발생한다. 그러나 인용정보에 DOI 번호를 삽입한 연계 방식은 문헌에 삽입된 연계정보를 수정할 필요가 없고 단지 변환 서버내 저장된 URL 정보를 갱신하면 되므로 관리가 용이하다.

DOI는 디지털 저작물에 대한 소유기관의 정보를 식별자 내에 포함할 수도 있으며 기존의 ISRC, ISWC, ISBN, ISSN 등과 같은 식별자를 포괄하여 사용할 수 있으므로 학술지 논문에 대한 저작권 관리도 가능하다. 즉, 학술지 논문을 생산한 지적저작물의\_

<표 1> URL과 DOI를 이용한 연계 비교

구 분	URL 연계	DOI 원문 연계	DOI 인용 연계
MDDDB	서지정보(URL)	메타데이터(DOI 포함)	-
변환서버	-	DOI-->URL	인용정보(DOI) --> URL
원문서버	URL상의 원문제공	URL상의 원문제공	URL상의 원문제공
장단점	-특정서버, Path 변경으로 Dead End 발생 -문서에 삽입된 연계정보 수정의 어려움	-특정서버, Path 변경으로 인한 URL의 수정사항은 변환 서버의 URL을 갱신	-문서에 삽입된 연계정보의 수정사항 발생시 변환 서버의 인용문헌 URL을 갱신

권리보유자와 이를 이용하는 이용자를 연결하여 학술지 논문에 대한 전자상거래까지도 실현할 수 있다. 또한 본 연구에서의 메타데이터 요소에는 저작물의 권리 정보는 포함하지 않았으나 각종 인터넷 자원 혹은 학술지 논문에 대한 유료화 정책으로 저작권자의 권리 정보는 점차 중요한 요소로 강조되고 있다. 최근 국제적인 저작권 단체에서는 INDECS(INteroperability of Data in E-Commerce System) (<http://www.indecs.org/>)라는 전자상거래를 위한 메타데이터 설계 프로젝트를 진행하고 있으며 이 프로젝트에서는 모든 형태의 저작물과 관련되는 메타데이터를 개발하여 저작권자의 권리와 상거래를 가능하게 하는 기술적 요소를 포함하도록 하였다.

DOI는 최근 활발히 개발중인 전자상거래 모델에서 디지털 문서 뿐 아니라 지적저작물이라 할 수 있는 각종 문서의 하위 단위(예: 표, 그림, 장, 절 등), 오디오저작물, 멀티미디어 저작물 등의 거래시 단위 항목을 식별할 수 있는 기본 단위로 사용되고 있다. 본 시스템에서는 이용자의 이용 권한을 제어하고 이용자의 원문보기 회수를 계산하여 비용을 부과하는 전자상거래의 기능은 고려하지 않았으므로 이후에 학술지 논문 연계시스템 구축시 추가 기능으로 확장되어야 할 것이다.

## 참고문헌

- 한혜영. 1999. 국내 학술지 게재논문을 위한 SICI 기반 DOI 체계. 『이화여대 문헌정보학과 창립 40주년 기념논문집』, 157-186.
- Bush, Vannevar. 1945. "As we may think." *The Atlantic Monthly*, 176(1): 101-108. [online]. [cited 1999. 8. 30]. <<http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>>.
- Caplan, Priscilla and Arms, Willam Y. 1999. "Reference linking for journal articles." *D-Lib Magazine*, 5(7/8). [online]. [cited 1999. 7. 21]. <<http://www.dlib.org/dlib/july99/caplan.html>>.
- Davidson, L. and Douglas, K. 1998. "Promise and problems for scholarly publishing." *The Journal of Electronic Publishing*, 4(2). [online]. [cited 1999. 8. 21]. <<http://www.press.umich.edu/jep/04-02/davidson.html>>.
- Hitchcock, S., L. Carr, S. Harris, J. M. N. Hey, and W. Hall. 1997. "Citation linking: Improving access to online journals." *Proceedings of the 2nd ACM International Conference on Digital Libraries*, 115-122. [online]. [cited 1999. 10. 3]. <<http://journals.ecs.soton.ac.uk/acmdl97.htm>>.
- International DOI Foundation(IDF). 1999. *IDF Annual Review*. [online]. [cited 1999. 9. 4]. <<http://www.doi.org>(PDF file)>.
- Lynch, Clifford A. 1997. "Identifiers and their role in networked information applications." *ARL: A Bimonthly Newsletter of Research Library Issues and Actions*, 194 (October). Washington, DC: Association of Research Libraries. [online]. [cited 1999. 9. 4]. <<http://www.arl.org/newsltr/194/identifier.html>>.
- Martin, David and Mark Bide. 1997. *Descriptive Standards for Serials Metadata and Standards for Terms of Availability Metadata: Two Related eLib Supporting Studies* commissioned by UKOLN. [online]. [cited 1999. 9. 4].

<<http://www.ukoln.ac.uk/dlis/models/models1/serials-mtg>>.

Moats, R. 1999. "A URN namespace for IETF documents." *AT&T*. [online]. [cited 1999. 10. 3]. <<http://www.ietf.org/rfc/rfc2648.txt>>.

National Information Standards Organization(NISO). 1996. *Serial Item and Contribution Identifier(SICI): An American National Standard Developed by the National Information Standards Organization*. Approved August 14, 1996 by the American National Standards Institute.

———. 1999. *Meeting Report of the NISO/NFAIS/SSP Linking Workshop*. 1999. 2. 11. Washington. DC.: The International DOI Foundation. [online]. [cited 1999. 8. 21]. <<http://www.niso.org/linkrpt.html>>.

Paskin, Norman. 1999. "DOI: Current status and outlook." *D-Lib Magazine*, 5(5). [online]. [cited 1999. 6. 1]. <<http://www.dlib.org/dlib/may99/05paskin.html>>.

Sollins, K. 1994. "Functional Requirements for Uniform Resource Names." [online]. [cited 1999. 10. 3]. <<http://www.ietf.org/rfc/rfc1737.txt>>.

Sun, Sam X. and Lannom, Larry. 1999. "Handle System Overview." Corporation For National Research Initiatives. [online]. [cited 1999. 10. 3]. <<http://www.ietf.org/internet-drafts/draft-sun-handle-system-03.txt>>.

Xu, Amanda. 1998. "Metadata Conversion and the Library OPAC." [online]. [cited 1999. 6. 25]. <<http://web.mit.edu/wayenj/www.xu.htm>>.