

# 국내 석박사 학위논문의 식별과 검색을 위한 URN 이용 방안

## A Study on the use of URN for identifying and retrieving theses in Korea

이상철, 김태수  
연세대학교 문헌정보학과

Lee, Sang-Cheol, Kim, Tae-soo, Dept. of Lib. & Inf. Sci., Yonsei University

인터넷 환경에서는 정보자원을 식별하기 위한 기호로서 URL을 사용하고 있다. 그러나 URL은 자원이 위치한 장소를 지시하기 때문에 자원의 위치가 변경될 경우에는 접근하지 못하는 단점이 있다. 본 논문에서는 이러한 URL의 단점을 보완하기 위한 대안으로서 URN을 이용하여 국내 학위논문을 고유하게 식별하고 검색하기 위한 방안에 대해 연구하였다.

### 1. 서론

현재 인터넷 환경에서 정보자원을 식별하고 접근하는 기호로서 URL이 이용되고 있는데 URL은 정보자원이 위치한 장소를 지시하기 때문에 자원의 위치가 변경되거나, 관리 서버의 이름이 변경되었을 경우와 같이 정보자원을 둘러싼 환경이 변화되었을 때, 기존의 URL로는 접근하지 못한다는 단점이 있다. 따라서 정보자원을 위치가 아닌 이름으로서 식별하고 접근할 수 있는 새로운 메카니즘이 필요하게 되었는데 이러한 대안으로서 URN이 거론되고 있다. 본 연구는 IETF에서 제안하고 있는 URN을 바탕으로 하여 URN의 개념과 필요조건을 검토하고, CNRI(The Corporation for National Research Initiatives)의 한글 시스템을 대상으로 하여 국내 학위논문을 고유하게 식별하고 검색할 수 있는 시스템을 제안하는데 그 목적

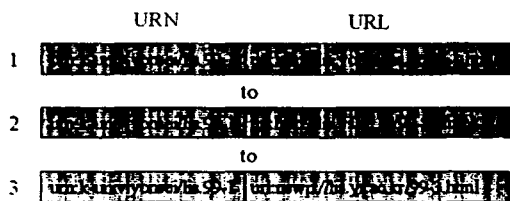
이 있다.

### 2. 본문

URN은 정보자원을 전송하는 프로토콜명과 정보자원을 등록한 호스트명, 그리고 정보자원의 파일명으로 구성되는데, 파일명이나 호스트명이 변경되거나, 정보자원의 소유권한이 변경되는 등의 이유로 자원을 관리하는 시스템이 변경될 경우 URL 자체도 완전히 바뀌게 된다. 이 때문에 기존의 이용자가 새 URL을 제공받지 못한 경우, 그 자원을 다시 이용할 수 없게 되는 상황에까지 이를 수 있다.

<그림 1>에서 [1-2]의 과정은 서버명이 변경된 경우이고 [2-3]은 새로운 프로토콜을 사용하여 URL이 변경된 경우이다. 이처럼 URL이 여러 번 변경될 때마다 이용자는 새로운 URL

을 알고 있어야 정보자원에 접근할 수 있다. 하지만 URN은 변하지 않기 때문에 이용자는 자신이 알고 있는 URN만을 가지고 원하는 정보자원에 언제든지 접근할 수 있는 것이다.



<그림 1> 변하지 않는 URN

### 2.1 URN 프레임워크

URN이 URN으로서 기능하기 위해서는 다음의 세 가지 조건을 갖추어야 한다. 첫째는 URN을 컴퓨터가 인식할 수 있는 구문이 있어야 한다. 다음으로는 URN을 URL로 변환할 수 있는 변환장치가 필요하다. 마지막으로 각각의 URN 제정기관에서 제정한 URN을 관리할 수 있는 관리 시스템이 필요하다.

#### 2.1.1 URN 구문

URN은 단순한 문자열이기 때문에 컴퓨터가 이를 인식하기 위해서는 일정한 규칙을 공유해야 한다. IETF의 URN 워킹그룹에서는 1997년 URN 구문을 RFC로 발표하여 모든 URN 제정기관에서 일정한 규칙에 맞게 URN을 제정하도록 하고 있다. URN 구문은 다음과 같다.

<URN> ::= "urn:" <NID> ":" <NSS>

모든 URN은 레이블로서 "urn"을 가지고, 이름부여기관을 의미하는 NID와 이름부여기관에서 정보자원에 고유하게 이름을 부여한 NSS로 구성된다. URN 구문은 예약어와 제외어를 갖는다.

#### 2.1.2 URN 변환장치

URN은 단지 정보자원의 고유한 식별명이므로 URN만으로는 정보자원에 직접 접근할 수 없다. 따라서 URN을 정보자원의 위치와 연결시켜 줄 수 있는 장치가 필요하다. URN 변환 시스템이란 URN을 URN이나 URC 혹은 또 다른 URN으로 변환시켜 주는 장치를 의미한다.

URN 변환방식의 가장 큰 특징은 이름부여방식과 변환방식을 독립시켰다는 것이다. 다시 말해, URN을 제정하는 곳에서는 URN 구문에만 맞으면 자유로이 URN을 제정할 수 있고, 변환 방식을 고안하는 곳에서도 다른 기관에서 어떠한 방식으로 URN을 제정한다고 해도 이에 상관없이 자유롭게 변환장치를 고안할 수 있는 것이다. 이처럼 이름부여방식과 변환방식을 독립시켰을 때의 문제점은 특정 URN을 어느 기관에서 변환시켜줄 수 있는가를 알아야 한다는 것이다. 현재 URN 워킹그룹에서는 이를 위해 NAPTR(Name Authority Pointer Resource) 변환방식을 제안하고 있다. NAPTR이란 일종의 데이터베이스로서 URN을 변환시킬 수 있는 변환기에 대한 정보(메타데이터)를 수록하고 있다. 웹브라우저가 특정한 URN을 만났을 때, 웹브라우저는 자동으로 NAPTR 레코드에 접근하여 해당 URN을 변환할 수 있는 서버를 검색한 후, 서버에 접근하여 URN을 URL로 변환하여 정보자원에 접근하는 것이다.

### 3. 핸들시스템

핸들 시스템은 URN 시스템의 하나로서 핸들(handle)이라는 디지털 자원의 식별기호를 제정하고, 관리, 저장하며 이를 변환하여 해당자원에 접근하게 하는 분산형 시스템이다. 핸들 시스템은 여러 로컬핸들서비스와 하나의 글로벌핸들서비스의 집합으로 이루어져 있는데 각각의 서비스는 각각의 핸들 네임스페이스를 갖

임진다. 각각의 서비스는 또한 여러 개의 핸들 서버로 구성되며 핸들서버에서는 각각의 핸들을 제정하고 관리한다. 핸들서비스 중에서 중심이 되는 서비스인 글로벌핸들서비스는 다른 모든 로컬서비스의 위치를 기억하고 로컬서비스의 네임스페이스를 책임진다. 각각의 로컬서비스는 글로벌핸들서비스에 접근하는 방법을 알고 있기 때문에 어떠한 핸들서버로 변환을 위한 질의가 들어오더라도, 이 질의를 해결할 수 있는 서버로 연결해 줄 수 있는 것이다. 핸들서버는 과부하를 방지하기 위해서 자신의 서버와 동일한 서버를 복제하여 운영할 수도 있다.

#### 4. 국내 학위논문 URN 시스템

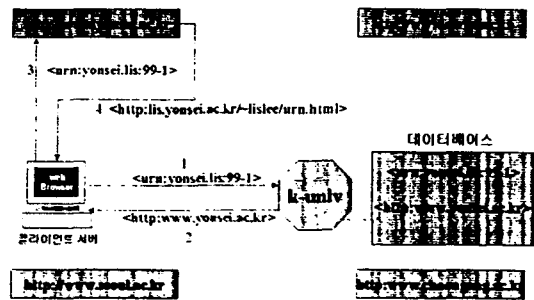
학위논문은 정보자원 가운데 학술정보로서 가장 가치가 있는 자원으로서 각 대학에서 고유하게 출판되는 출판물의 성격을 지니기 때문에 각 대학에서 URN을 부여하고 관리하기에 용이하다. 따라서 학위논문에 URN을 부여하고, 이 URN을 이용하여 정보자원에 접근하는 시스템을 설계하는 작업은 앞으로 다른 정보자원을 대상으로 URN을 부여하는 방안의 기초 연구가 될 수 있을 것이다.

국내 학위논문 URN 시스템은 다음의 원칙을 공유한다.

1. 각 대학에서 발행된 학위논문은 각 대학도서관에서 URN을 부여한다.
2. 각 대학에서 발행된 학위논문의 URC(메타데이터)는 각 대학도서관에서 작성한다.
3. 각 대학에서 발행된 학위논문 전문(全文)의 디지털화 작업은 각 대학도서관에서 담당한다.
4. 각 대학에서 발행된 학위논문의 URN 변환 작업은 각 대학도서관에서 담당한다.
5. 국내 대학도서관에서 작성된 모든 URN의 질기 데이터는 하나의 기관에서 일괄적으로 관

리, 운영되고 그 기관의 이름은 'k-univ'로 한다.

6. 특정 대학도서관 변환서버의 위치가 변경되었거나 정보자원의 위치가 변경되었을 경우, 이를 'k-univ'에 알려서 데이터베이스가 갱신될 수 있도록 한다.



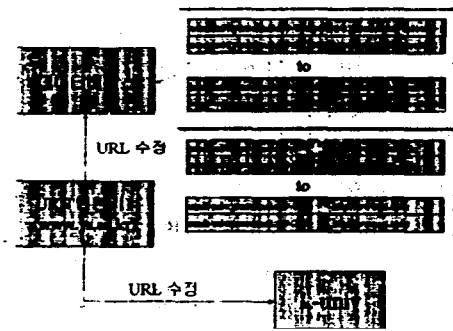
<그림 2> 국내 학위논문 URN 시스템

위의 원칙 중에서 개별 대학도서관에서 제정한 URN을 변환할 수 있는 변환서버의 위치에 대한 정보를 관리, 운영하는 'k-univ'의 역할은 국립중앙도서관이나 국회도서관에서 담당할 수 있을 것이다. 현재 두 도서관은 국내 학위논문에 대해 서지정보를 제공해 주고 있으므로, 별도의 검색시스템을 고안한 필요없이 기존의 검색시스템을 그대로 이용할 수 있다. 이용자가 URN을 클릭하였을 때 웹브라우저는 k-univ에 접근하여 이용자가 요청한 URN을 변환할 수 있는 서버를 검색한다. k-univ에서는 여러 기관 중에서 이용자가 요청한 URN을 변환할 수 있는 기관에 대한 정보를 이용자의 웹브라우저에 제공한다. 이용자의 웹브라우저가 변환서버에 접근하면 변환서버는 이용자가 요청한 URN에 해당하는 학위논문을 이용자에게 제공하는 것이다. 위의 과정은 <그림 2>와 같다. URN 식별기호는 변하지 않지만, URN을 변환하여 주는 변환기의 위치와 URN으로 식별되는 정보자원의 위치는 변환 수 있다. 이러한

URN의 변환기나 정보자원의 위치가 변경되었을 경우에는 그 변경사항을 각각의 데이터베이스에서 수정하여야 한다.

<그림 3>에서 <urn:k-univ:yonsei/m/lis/99-1>이라는 URN을 변환시켜 주는 변환서버의 주소가 <http://www.yonsei.ac.kr>에서 <http://www.ys.ac.kr>로 변경되었을 때에는 변경사항을 k-univ에 공지하여 변환서버의 주소를 갱신할 수 있도록 한다. 또한 정보자원의 소장위치가 변경되었을 경우에도 변경사항을 k-univ에 공지한다.

이처럼 이용자는 자신이 검색하고자 하는 학위논문의 소장위치가 서버명이 변경된다고 할지라도 새로운 URL을 알 필요없이 자신이 알고 있는 URN으로써 원하는 학위논문을 검색하고 접근할 수 있는 것이다.



<그림 3> 데이터베이스의 수정

#### 4. 결론

국내 학위논문 URN 시스템의 효과는 다음과 같다. 첫째, 이용자는 학위논문의 소장 위치가 변경되거나 이름이 변경되더라도 자신이 알고 있는 URN으로써 언제나 고유하게 자신이 원하는 학위논문에 안정적으로 접근할 수 있으므로, 정보검색과 정보자원으로의 접근을 일원화하여 URL의 단점을 보완할 수 있다. 둘째, 관

리자는 정보자원을 관리하던 기존의 이름부여 기관이 다른 이름부여기관과 통합되거나, 새로운 이름부여기관으로 변경되더라도 고유하게 학위논문의 소재지를 이용자에게 제시할 수 있다. 셋째, 학위논문 URN 시스템은 분산형 데이터베이스 구조를 유지하므로 개별 대학마다 동일한 학위논문을 중복해서 소장할 필요없이 각 대학에서는 자신의 대학에서 발행된 학위논문만을 소장하면 된다.

앞으로 URN에 관한 연구는 학위논문뿐만 아니라 단행본이나 연속간행물과 같은 도서관의 자료를 대상으로 연구 범위를 확대해야 할 것이다. 또한 도서관 자료 이외에도 특허자료, 정부 행정 자료, 고문서, 박물관 자료 등과 같이 앞으로 디지털 형태로 소장하고 관리할 수 있는 모든 자료를 대상으로 연구를 진행시켜 나가야 할 것이다.

#### <참고문헌>

- Daniel, Ron, and Mealling, Micheal. 1997. Resolution of Uniform Resource Identifiers using the Domain Name System. [1999.6.22] <ftp://ftp.isi.edu/in-notes/rfc2168.txt>
- Mealling, Micheal, and Daniel, Ron. 1999. The Naming Authority Pointer (NAPTR) DNS Resource Record. [1999.6.22] <http://www.ietf.org/internet-drafts/draft-ietf-urn-naptr-rr-01.txt>
- Moats, M. 1997. URN Syntax. [1999.6.22] <ftp://ftp.isi.edu/in-notes/rfc2141.txt>
- Sollins, K., and Masinter, L.. 1994. Functional Requirements for Uniform Resource Names. RFC1737. [1999.6.22] <http://info.internet.isi.edu:80/in-notes/rfc/files/rfc1737.txt>
- URN-WG. 1999. Uniform Resource Names (urn) homepage. [1999.6.22]. <http://www.ietf.org/html.charters/urn-charter.html>