식별체계 간 연계시스템 구축에 관한 연구

신동구°, 김재수*, 윤정모**, 권이남*, 전성진*, 정택영*
*한국과학기술정보연구원 표준화기술지원실
**서울산업대학교 컴퓨터공학과
*{lovesin°, jaesoo, ynkwon, sjjhun, tychung }@kisti.re.kr, **jmyoon@snut.ac.kr

A study on the construction of a linking system between Identification systems

donggu Shin°, J.S.Kim*, J.M.Yoon**, Y.N.Kwon*, S.J.Jhun*, T.Y.Chung*
*Dept. of Technical Support for Standardization, Korea Institute of Science and Technology Information
**Dept. of Computer Science & Engineering, Seoul National University of Technology

요 약

인터넷의 발달로 디지털콘텐츠의 생성과 활용이 급격히 증가하면서 디지털콘텐츠에 대한 서비스 방식이 URL 방식에서 URN 방식으로 변화하고 있다. 한국과학기술정보연구원의 KOI나 한국전산원의 UCI가 그 예라고 할 수 있다. 그러나 식별체계 간의 등록·변환서비스가 호환되지 않기 때문에 식별체계를 활용하여 디지털콘텐츠를 서비스하기 위해서는 커널메타데이터를 각각 등록하여야 하며, 디지털콘텐츠 변환서비스도 서로 연계가 이루어 지지 않기 때문에 개별적으로 서비스 되고 있다. 디지털콘텐츠를 보다 효율적으로 생성·활용하기 위하여 식별체계 간의 등록·변환서비스 연계가 가능한 식별체계 간 연계시스템 구축 방안을 제시하고자 한다.

1. 서 론

인터넷의 발달로 디지털콘텐츠의 생성과 활용이 급격히 증가하면서 디지털콘텐츠에 대한 서비스 방식이 사용자가 위치정보를 지속적으로 관리해야 하는 URL 방식에서[1] 사용자는 디지털콘텐츠에 대한 변하지 않는 유일한 식별번호만 관리하고, 서비스 시스템에서 위치정보를 관리하는 URN 방식으로 변화하고 있다[2]. 2005년 현재 디지털콘텐츠를 URN 방식으로 서비스 중인 식별체계를 살펴보면 해외에서는 IDF[1]의 DOI[2], JST[3]의 JOI[4], 국내에서는 한국과학기술정보연구원의 KOI[5], 한국전산원의 UCI[6] 등이 있다.

그러나 이들 식별체계들은 상호간에 등록·변환서비스가 호환되고 있지 않다. 그렇기 때문에 디지털콘텐츠 등록자나 이용자가 식별체계 서비스를 활용하기 위해서 디지털콘텐츠 등록자는 디지털콘텐츠를 각 식별체계별에서 제공하는 등록기능을 이용하여 개별적으로 등록하여야 하며, 이렇게 등록된 디지털콘텐츠는 각 식별체계에서 개별적으로 변환서비스 되고 있다. 즉 등록자나 이용자가 2개 이상의 식별체계에서 디지털콘텐츠를 활용하기 위해서는 비효율적인 등록·변환서비스를 이용해야하는 것이다.

본 논문에서는 이러한 문제점을 개선하고, 디지털콘텐츠를 보다 효율적이고 다양하게 생성 및 활용하기 위하여 식별체계 간 연계시스템 구축 방안을 제시하고자 한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 연구로 식별체계에 대하여 기술하고, 3장에서는 국내 식별체계 KOI, UCI 서비스 현황에 대하여 기술하고, 4장에서는 KOI, UCI 간 등록·변환 연계시스템 구축 방안에 대하여 기술하고, 5장에서는 결론과 향후 연구방안에 대하여 기술한다.

2. 관련연구

기존의 물리적인 저작물이 디지털화 되면서 기존의 식별체계가 디지털콘텐츠의 특성을 충족시키기에는 미흡하게 되었다. 디지털콘텐츠의 특성상 각각의 오브젝트 즉, 각 장르, 페이지, 그림이나 표 등을 개별적으로 식별할 필요성이 나타나게 되었다. 이러한 필요성에 의해 개발된 디지털콘텐츠 식별체계는 디지털콘텐츠에 변하지 않는 유일한 식별번호를 부여하여 디지털콘텐츠와 관련된 각종 메타데이터를 조직하고, 절차에 따라 운영·관리하는 시스템이다[6,7,9].

해외에서 서비스 중인 식별체계로 1997년 IDF에서 URN 원리를 이용한 식별체계로 DOI를 개발하였으며[3,4,5], 현재 학술지 장르에 대한 디지털콘텐츠를 CrossRef에서 활발하게 서비스 하고 있다. 이후 일본의 JST에서는 JOI를 개발하여 학술지 분야에서 활용하고 있다.

국내에서 서비스 중인 식별체계로 2002년 한국과학기술정보연구원에서는 과학기술 지식정보에 대한 디지털콘텐츠 서비스를 위하여 KOI를 개발하였으며, KISTI 종합정보서비스 시스템에서 KOI를 활용하여 디지털콘텐츠를 서비스 하고 있다[9,12]. 그리고 한국전산원에서는 2003년 디지털화가 가능한 모든 자원에 대한 디지털콘텐츠서비스를 위하여 UCI를 개발·운영하고 있다[8,10,11].

1) IDF : International DOI Foundation
2) DOI : Digital Object Identifier
3) JST : Japan Science and Technology Corporation
4) JOI : JST Object Identifier
5) KOI : Knowledge Object Identifier
6) UCI : Universal Content Identifier

3. 식별체계 KOI, UCI 서비스현황[7]

국내에서 서비스 중인 식별체계 KOI와 UCI의 식별대상으로 KOI는 과학기술 지식정보 중에서 디지털화가 가능한 콘텐츠를 대상으로 하고 있으나, UCI는 디지털화가 가능한 모든 콘텐츠를 대상으로 하고 있다. KOI와 UCI의 구문구조는 서로 다른 식별대상을 표현하기 위하여 서로 다른 구문구조를 사용하고 있다. KOI는 과학기술분야에 대한 디지털콘텐츠 서비스를 위하여 Classification을 포함하였으며, UCI는 디지털화가 가능한 모든 콘텐츠를 포함하기 위하여 Mode를 포함하였다. 그러나 커널메타데이터 전체적으로 식별번호, 기존식별번호, 제목, 저자와 저자역할 등을 포함하는 KOI 통합메타데이터와 UCI 식별메타데이터는 유사한 것으로 조사되었다.

식별체계 서비스를 위해 필요한 주요기능인 커널메타데이터 등록기능 처리프로세스와 디지털콘텐츠 변환기능 처리프로세스가 유사한 것으로 조사되었으며, 시스템 구성 측면에서는 KOI가 Root레벨, 등록관리레벨, 등록레벨의 3-tier 구조로 이루어 졌고, UCI는 Root레벨, 등록관리레벨, 등록레벨, 그리고 시스템 추가가 가능한 하위등록관리레벨로 이루어진 n-tier 구조로 이루어져 있어서 유사한 시스템 구성을 가진 것으로 조사되었다.

3.1 식별체계 구문구조 및 커널메타데이터

3.1.1 KOI 구문구조 및 통합메타데이터

KOI 구문구조는 prefix와 suffix로 나누어진다. prefix는 등록관리기관코드와 등록기관코드로 구성되어 있고, suffix는 학술지형태인 경우 자료유형, ISSN, Vol, Issue, Startpage의 종정보와 확장코드로 구성되고, 비학술지형태인 경우에는 자료유형과 등록기관에서 부여하는 관리번호, 그리고 확장코드로 구성된다.

표 1. KOI 구문구조

학술지 :등록관리기관.등록기관/자료유형.종정보.[확장코드]
비학술지 :등록관리기관.등록기관/자료유형.관리정보.[확장코드]

KOI 통합메타데이터는 다음과 같다.

표 2. KOI 통합메타데이터

영문요소	정의
KOI	Knowledge Object Identifier
Identifier	기존에 부여된 식별자
Title	알려져 있는 디지털콘텐츠의 이름
Year	디지털콘텐츠 제작년도
Copyright owner	디지털콘텐츠에 대한 책임을 가진 주체의 이름
Copyright Owner role	디지털콘텐츠에 대한 책임을 가진 주체의 역할
Type	디지털콘텐츠 자료유형(학술지, 연구보고서 등)
Classification	디지털콘텐츠 분류값
File format	디지털콘텐츠 파일형태
Description	디지털콘텐츠 설명

7) 식별체계의 커널메타데이터를 KOI는 통합메타데이타로 UCI는 식별메타데이터로 표현하고 있다. 또한 식별체계의 Root시스템은 KOI는 통합관리시스템으로 UCI는 총괄관리시스템으로 표현하고 있다

3.1.2 UCI 구문구조 및 식별메타데이터

UCI 구문구조는 접두코드, 개체코드, 한정코드로 나누어진다. 접두코드는 변환과 콘텐츠 등록 구조에 관련이 있는 코드이고, 개체코드는 콘텐츠 자체를 식별하기 위해 등록자가 콘텐츠에 부여하는 코드이며, 한정코드는 디지털콘텐츠의 의미상의 변화가 아닌 형태상의 변화 나타내는 코드이다.

표 3. UCI 구문구조

접두코드 - 개체코드 - 한정코드

UCI 식별메타데이터는 다음과 같다.

표 4. UCI 식별메타데이터

영문요소	정의
UCI	Universal Content Identifier
Identifier	기존에 부여된 식별자
Title	알려져 있는 자원의 이름
Mode	자원의 주된 감각적 혹은 지각적 표현 형태 4가지(시각, 청각, 시청각, 추상)
Format	자원의 데이터 표현형식
Contributor	자원의 내용에 주된 책임을 가진 주체와 역할의 의미를 포괄하는 명칭
ContributorEntity	자원의 내용에 주된 책임을 가진 자의 이름
ContributorRole	자원의 내용에 주된 책임을 가진 자의 역할

3.2 식별체계 시스템 구성 및 등록·변환서비스 현황

KOI 시스템구성은 3-tier 구조로 구성되어 있으며, UCI 시스템구성은 n-tier 구조로 구성되어 있다. Root레벨은 등록관리기관관리, 식별메타데이터관리기능 등을 수행하는 KOI 통합관리시스템과 UCI 총괄관리시스템, 등록관리레벨은 등록자관리, 응용메타데이터관리, URL유효성검사기능 등을 수행하는 KOI와 UCI 등록관리시스템, 그리고 등록레벨은 응용메타데이터를 등록하는 KOI와 UCI 등록시스템으로 구성되어 있으며, UCI의 경우 등록관리레벨의 UCI 등록관리시스템에 다수의 하위 UCI 등록관리기관을 추가 할 수 있는 시스템으로 구성된다.

KOI와 UCI 등록프로세스는 등록레벨에서 식별번호를 생성하고, 생성된 식별번호와 응용메타데이터 그리고 디지털콘텐츠 URL정보를 등록관리레벨에 등록한다. 등록관리레벨에서는 등록관리레벨에 저장된 응용메타데이터 중에서 커널메타데이터를 추출하여 Root레벨에 등록한다.

KOI와 UCI 변환프로세스는 이용자가 Root레벨에 디지털콘텐츠 1차 변환을 요청하게 되고, Root레벨에서는 1차 변환을 처리한 후 등록관리레벨로 2차 변환을 요청하고, 등록관리레벨에서는 2차 변환을 처리 한 후 등록시스템에서 등록한 URL 정보를 이용하여 이용자에게 디지털콘텐츠 서비스를 제공하는 것이다.

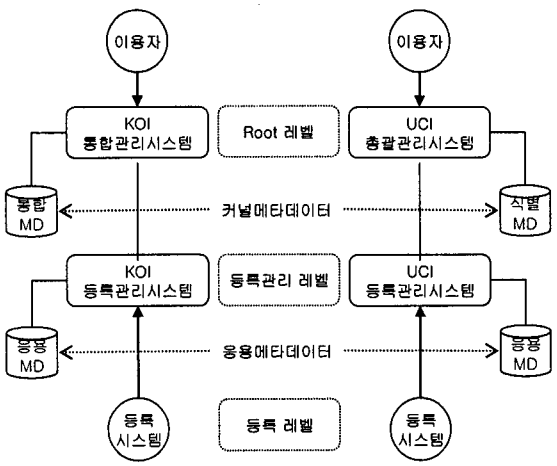

그림 1. KOI와 UCI 식별체계 시스템 구조도

4. 식별체계 간 등록·변환 서비스 연계

식별체계 KOI와 UCI의 커널메타데이터 등록서비스 연계 방안은 유사한 KOI 통합메타데이터와 UCI 식별메타데이터의 공통필드를 추출하여 메타데이터를 mapping함으로써 연계 가능하도록 하였다.

그리고 디지털콘텐츠 변환서비스 연계방안은 KOI와 UCI의 시스템구조가 Root레벨, 등록관리레벨, 등록레벨을 포함하는 서비스체계로 이루어져 있기 때문에 3-tier로 연계가 가능하다는 점과 각 식별체계의 변환프로세스가 유사하다는 점을 이용하였다.

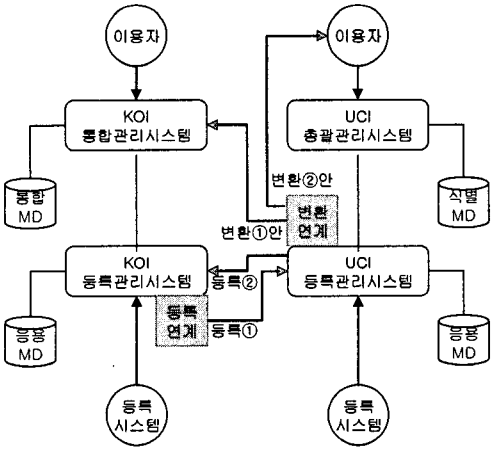

그림 2. KOI/UCI 식별체계 간 연계시스템 구성도

4.1 식별메타데이터 등록서비스 연계방안

KOI 등록시스템에서 KOI 등록관리시스템으로 KOI 응용메타데이터를 등록하면서 식별체계 UCI를 사용할 것이지 체크한다. KOI 등록관리시스템은 KOI 통합메타데이터를 KOI 통합관리시스템에 보내서 KOI 통합메타데이터 등록을 완료한다. 그리고 등록시스템에서 식별체계 UCI를 이용한다는 체크정보에 의해 UCI 응용메타데이터를 KOI 등록관리시스템에서 UCI 등록관리시스템에 등록하고(등록①), UCI 등록관리시스템에 등록된 UCI 식별메타데이터는 UCI 총괄관리시스템에 자동 전송된다. 이때 UCI 등록관리시스템에서는 메타데이터가 UCI에 정상적으로 등록되었음을 알려주는 등록완료메세지를 KOI 등록관리시스템에 보냄(등록②)으로써 식별메타데이터 연계 등록 프로세스가 완료된다.

4.2 디지털콘텐츠 변환서비스 연계 방안

디지털콘텐츠 변환서비스 연계방안은 KOI 등록관리시스템에서 연계를 위한 메타데이터 등록 시 URL필드의 값을 KOI 식별체계 변환서비스를 위한 KOI 변환요청 URL을 등록하느냐(변환①안), 실제 디지털콘텐츠URL을 등록하느냐(변환②안)에 따라 변환서비스 연계프로세스는 다르게 처리된다. 변환①안의 경우는 이용자가 UCI 변환요청을 하면, UCI 총괄관리시스템에 1차 변환을 하고, UCI 등록관리시스템에서 2차 변환을 수행하는데 2차 변환 결과물이 KOI 변환요청URL이므로 KOI에서 다시 변환서비스를 하게 되는 프로세스이다. 그러나 변환②안의 경우는 UCI 등록관리시스템에서 2차 변환의 결과물이 디지털콘텐츠URL이므로 이용자에게 직접 디지털콘텐츠 URL을 서비스 하게 되고, 변환로그는 KOI 시스템에 전송하는 프로세스이다.

5. 결론

본 논문에서는 식별체계 간 등록·변환 연계시스템 구축 방안에 대하여 논하였다. 식별체계 간의 연계시스템을 구축함으로서 디지털콘텐츠 등록자나 이용자는 보다 효율적이고 다양한 형태로 디지털콘텐츠를 등록하고 서비스를 받게 된다. 또한, 향후 개발되는 URN 식별체계에 대한 호환성과 확장성도 높아짐으로써 보다 많은 디지털콘텐츠가 식별체계를 활용하여 서비스 될 것이다.

향후 연구과제로는 URN 식별체계를 보다 활성화하기 위해서 디지털콘텐츠에 URN 식별자를 링크하는 참고문헌 링크방안 등에 대한 지속적인 연구가 필요하다.

6. 참고문헌

[1] Norman Paskin, "DOI:Current Staus and Outlook," D-Lib Magagine, May 1999.
[2] K. Sollins, L. Masinter, "Fuctional Requirments for Uniform Resourece Names," RFC1737, 1994.
[3] R. Moats, "URN Stntax," RFC2141, 1997.
[4] Brain Green & Mark Bide, "Unique Identifiers: a brief introduction," BIC and EDItEUR, 1997
[5] John Erickson, "The role if metadata supply chains in DOI-based, value-added services," ICSTI Forum, No.30, 1999.
[6] 안계성외, "디지털콘텐츠 식별자 구문구조 연구", 정보통신부, 2000.
[7] 조소연, "디지털콘텐츠 식별체계," TTA저널 제84호, 한국통신기술협회, 2002.
[8] 한국전산원, "국가 URN기반 식별체계의 식별메타데이터 연구",2003.
[9] 이상환, 신동구 외, "식별체계 기반의 과학기술분야 전자원문 연계 시스템 설계 및 구현" 한국정보과학회 춘계학술대회, 2004.
[10] 한국전산원, "UCI 명세서(ver 1.0)", 2004.
[11] http://www.uci.or.kr
[12] http://koi.kisti.re.kr