

# 건설사업보고서 전자납본을 위한 XML 스키마기반의 메타데이터 개발

## The Development of XML Schema based Metadata for the Electronic Specimen Copy of Construction Project Reports

정성윤\*

한국건설기술연구원 ICT융합연구소\*

Jeong Seong-Yun\*

Korea Institute of Civil Engineering and Building Technology\*

### 요약

건설사업 공공발주기관은 자체 발행하는 각종 건설기술관련 자료를 납본하도록 관련법이 마련되었다. 본 연구는 건설사업보고서를 전자적으로 납본하고, 납본된 보고서를 인터넷을 통해 신속하고 정확하게 검색할 수 있고, 향후 관련DB와의 연계성 등을 고려하여 선행사례를 토대로 17개의 데이터 요소들로 구성된 XML 스키마기반의 메타데이터 개발에 관한 방안을 제시하였다.

## I. 서론

### 1. 연구 필요성 및 목적

1990년대에 엘빈 토플러가 정보화 발전추세를 “제3의 물결”이라고 예고한 이래에 정보의 홍수라 일컬을 만큼 매년 방대한 분량의 정보가 생성하고 유통되고 있다. 정보량이 급속히 증가함에 따라 원하는 정보를 빠르고 정확하게 찾기 위한 요구도 증가하게 되었다. 이런 요구를 충족시키기 위해 시맨틱 웹(semantic web), 검색엔진(search engine), 검색로봇(search robot), 스파이더(spider), 빅데이터(big data) 등 다양한 정보검색기술이 개발되었다. 하지만 이들 검색기술 중 간단하면서 효과적으로 원하는 정보를 찾는 방법 중 하나가 메타데이터 기술이다.

한편 건설분야에서도 각종 건설제기준, 건설기술, 자재규격, 공사발주 등 다양한 정보들이 생성, 유통되고 있다. 이들 건설기술정보는 건설생산품의 품질향상, 건설업체의 경쟁력 강화 등에 영향을 미칠 수 있는 매우 중요한 정보들이라 할 수 있다. 정부는 이러한 건설기술정보에 대해 종합적인 유통체계를 갖추고, 그 보급과 확산을 위하여 건설기술진흥법과 시행규칙을 마련하여 건설기술자료와 정보를 수집·유통 등의 건설기술정보 유통체계를 구축하기 위한 법적근거를 마련하였다.

본 연구는 공공발주기관에서 자체 발행하는 건설사업보고서를 전자적으로 납본하고, 납본된 보고서를 인터넷을 통해 신속하고 정확하게 검색할 수 있고, 향후 전자납본 DB와의 연계성 등을 고려하여 선행사례를 토대로 XML 스키마기반의 메타데이터 개발에 대한 방안을 제시하였다.

## II. 건설기술관련 메타데이터 개발 사례

### 1. 더블린 코어(Dublin Core) 메타데이터

서론에서도 언급하였듯이 원하는 정보를 정확하게 접근하여 사용될 수 있는 여러 정보검색기술이 개발되었지만 간단하면서도 효과적으로 정보를 찾는 방법 중 하나가 메타데이터(Metadata) 방법이라 할 수 있다.

메타데이터는 사용분야에 따라 여러 가지의 의미로 사용되고 있지만 기본적으로 ‘필요한 정보자원을 접근할 수 있도록 정보자원을 대표하는 데이터 요소들을 하나의 집합체(metadata element set)로 묶은 정보’라고 말할 수 있다. 가장 대표적으로 사용되는 메타데이터로는 더블린 코어(Dublin Core) 메타데이터가 있다.

더블린 코어는 1995년에 미국 오하이오 주의 더블린에서 개최된 워크숍에 참가한 도서관의 관계자, 콘텐츠 제작자, 웹관련 전문가들이 정보자원을 효율적으로 검색하기 위해 합의한 메타데이터를 ‘더블린 코어’라고 부르기 시작하였다. 더블린 코어는 ‘Title’, ‘Creator’, ‘Subject’, ‘Description’, ‘Publisher’, ‘Contributor’, ‘Date’, ‘Type’, ‘Format’, ‘Identifier’, ‘Source’, ‘Language’, ‘Relation’, ‘Coverage’, ‘Rights’ 등 15개의 기본 데이터 요소로 구성되어 있다[1].

### 2. 건설교통기술지식정보 DB 사업의 메타데이터

건설교통기술지식정보 DB 사업은 한국정보사회진흥원에서 주관한 국가지식정보자원관리 사업 중 하나로 추진된 정보화 사업이다. 이 DB 사업은 건설공사 수행과정에서 빈번히 활용되는 자료들 중 시공계획서, 시공절차서, 원가절감 및 VE 사례, 건설공사지 등의 자료와 정보에 대한 메타데이터를 2001년에 개발하였다.

이 사업에서 개발한 메타데이터는 더블린 코어에서 정한 기본 데이터 요소 중 ‘Contributor’, ‘Source’, ‘Coverage’ 등 3개의 요소를 제외한 12개 요소와 자체적으로 정한 ‘ProjectOrganization’ 요소 등 총 13개의 데이터 요소로 구성되었다. 2013년에 더블린 코어의 데이터 요소에서 ‘Coverage’ 등 요소를 제외하고, ‘Collection’,

'KC subject' 'Recordunit' 'metaMetadata', 'Open' 등 요소를 추가한 건설사업보고서 메타데이터를 마련하였다.

### 3. 전자설계도서 성과품 메타데이터

지방국토관리청에서 발주하는 도로와 하천 건설사업에서 생성되는 설계 또는 준공 성과품 자료와 정보를 체계적으로 관리하고, 재활용할 수 있도록 "전자설계도서 작성·납품 지침"을 마련되었다. 이 지침에는 설계도서나 준공도서 성과품 전체를 대표할 수 있는 마스터정보와 함께 성과품에 포함된 개별도서의 주요사항을 뽑아, 찾아보기 쉽게 모아 놓은 색인데이터를 XML 스키마기반으로 정의하였다[3].

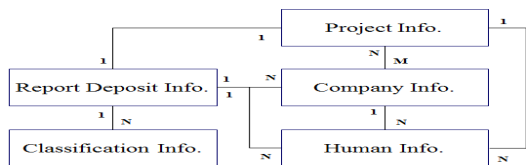
설계보고서의 색인데이터는 '문서관리번호', '공사목적', '설계구간정보', '공사내용', '설계속도', '도로폭원', '차로수', '공사량', '노선번호', '총공사비', '교통량정보', '독마루폭(하천)', '여유고(하천)', '교비탈경사(하천)', '호안형식(하천)', '최종작성날짜', '문서작성소프트웨어', '키워드', '원본파일정보' 등의 데이터 요소들로 구성되었다.

### III. 전자납본을 위한 XML 메타데이터 개발

과거에는 건설사업이 준공되면 성과품을 종이형태의 인쇄물로 납본하였으나, 현재는 대부분의 준공성과품을 전자파일로 작성하기 때문에 전자화된 준공성과품을 성과품 관리부서에 온라인으로 제출하고 있다.

본 연구는 건설기술정보 유통체계 구축에 일조할 수 있도록 전자화된 건설사업보고서를 온라인으로 납본하고, 납본된 보고서가 필요하면 언제, 어디서나 정확하게 접근하여 활용할 수 있도록 전자납본을 위한 XML 스키마기반의 메타데이터를 마련하였다. 여기서 XML 스키마는 자료나 정보가 갖고 있는 의미와 속성을 정의할 수 있다는 장점 때문에 웹기반의 정보검색분야에서 가장 많이 사용되는 구조적(markup) 언어를 말한다.

본 연구는 건설교통기술지식정보와 전자설계도서 성과품 중 보고서와 관련한 메타데이터의 데이터 요소와 전자납본 업무처리에 필요한 데이터 요소들을 조사, 분석하였다. 분석결과를 토대로 향후 전자납본 DB와의 연계성 확보를 고려하여 그림 1과 같이 데이터 요소 집합들 간의 논리 관계를 설정하였다.



▶▶ 그림 1. date element set 간의 관계

그림 1에서 Project Info.은 '사업식별코드', '사업명(사업, 공사, 공구)', '사업관리번호', '사업분야', '사업기간', '공사개요 등의 데이터 요소들로 구성하고, Company Info.은 '기관식별코드', '기관명', '주소', '기관유형'으로 구성하였다. Human Info.은 '인적식별코드', '성명', '부서명', '연락처', '참여유형'으로, Classification Info.은 '주제별분류코드', '분류명' 등의 데이터 요소로 구성하였다.

Report Deposit Info.로 표 1과 같이 더블린코어와 건설교통기술지식정보에서 사용한 11개의 데이터 요소와 전자납품과 관련한 4개 요소, 전자납본처리와 관련한 2개의 요소 등 17개의 데이터 요소를 정의하였다.

표 1. Report Deposit Info.의 데이터 요소 정의

Element	Sub Element	Skip	필수	검색	재사용
Title	Main Title, Alternative Title		Y	Y	Y
Author		Human Info.	N	N	Y
Publisher		Company Info.	Y	Y	Y
Date	Published, Registered, Modified	Date	Y	N	Y
Identifier		URI	Y	N	Y
Language		ISO-639-2	Y	Y	N
File Format			N	N	N
Rights			Y	N	N
Keyword			Y	Y	N
Relation	Is Part of, Is Required by, Is referred by	URI	N	N	Y
Project Name		Project Info.	Y	Y	Y
Abstract		Project Info.	N	N	Y
Classification Name		Classification Info.	Y	Y	Y
Pages			N	N	Y
Information Open			Y	N	N
Submitter		Human Info.	Y	N	N
Process Status		Deposit-process	Y	N	N

XML 스키마 언어규칙을 적용하여 정의된 데이터 요소들을 XML 메타데이터 스키마로 작성하였다. 그림 2는 Report Deposit Info.에 포함된 'Title', 'Date', 'Process Status' 데이터 요소를 XML 스키마로 정의한 예를 나타낸 것이다.

```

<xs:element name="Title" - Title element -
  <xs:complexType>
    <xs:sequence>
      <xs:element name="MainTitle" type="xs:string"/>
      <xs:element name="AlternativeTitle" type="string" minOccurs="0"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
<xs:element name="Date" - Date element -
  <xs:complexType>
    <xs:sequence>
      <xs:element name="Published" type="dateFormatType"/>
      <xs:element name="Registered" type="dateFormatType"/>
      <xs:element name="Modified" type="dateFormatType" minOccurs="0"/>
    </xs:sequence>
  </xs:complexType>
</xs:element>
- Process Status element -
<xs:element name="ProcessStatus" type="Deposit-processType"/>
  
```

▶▶ 그림 2. Report Deposit Info.에 대한 XML 스키마 정의의 예

### IV. 결론

전자설계보고서의 납본처리와 납본된 설계보고서를 인터넷을 통해 빠르게 정확하게 접근할 수 있도록 XML 스키마를 기반으로 한 건설사업보고서의 전자납본 메타데이터 개발방안을 제시하였다. 건설보고서외에 건설기준, 건설규격 등으로 확장하는데 본 연구를 준용한다면 기초자료로써 활용할 수 있을 것으로 사료된다.

### 참고 문헌

- [1] 한국 더블린코어, "http://www.dublincore.go.kr/metadata-basics/dces" (2015.03.03에 참조하였음)
- [2] 안순, 조정근, 김승관, "14 건설기술정보 DB 및 서비스시스템 운영", 한국건설기술연구원(KICT 2014-166), pp.62-65, 2014.
- [3] 국토교통부, "전자설계도서 작성·납품지침(도로·하천분야) [부속서 8]", pp.1-5, 2014.