

ISO 11179 기반 데이터 레지스트리에서의 XML 생성 인터페이스 설계 및 구현

양유승^o, 박혜숙, 나홍석, 백두권
고려대학교 컴퓨터학과 소프트웨어 시스템 연구실
e-mail : {[yys](mailto:yys@swsys2.korea.ac.kr), [hspark](mailto:hspark@swsys2.korea.ac.kr), [nhs](mailto:nhs@swsys2.korea.ac.kr), [baik](mailto:baik@swsys2.korea.ac.kr)}@swsys2.korea.ac.kr

Design and Implementation of a XML Generating Interface for an ISO 11179-based Data Registry

Yoo-seung Yang^o, Hea-sook Park, Hong-seok Na, Doo-kwon Baik
Software System Lab. Dept. of Computer Science & Engineering, Korea University

요 약

메타데이터를 표준화된 형태로 저장하기 위해서 제안된 데이터 레지스트리(DR)는 메타데이터의 유통을 위한 구체적인 방안을 제시하고 있지 않으며, 현재 데이터 레지스트리는 HTML 을 이용한 인터페이스만을 제공하고 있다. 본 연구에서는 데이터 레지스트리에 저장된 메타데이터를 효율적으로 유통할 수 있는 방법으로 XML 을 이용하였으며, 이를 위해 ISO/IEC 11179 에 기반한 데이터 레지스트리로부터 XML 문서를 생성할 수 있는 인터페이스를 설계 및 구현하였다. 구현한 XML 생성 인터페이스는 데이터 레지스트리에 저장된 메타데이터를 XML 형식으로 만들어 줌으로써, 응용 프로그램들이 효과적으로 데이터 레지스트리에 접근하여 메타데이터를 이용할 수 있는 환경을 제공한다.

1. 서론

지금까지는 데이터를 유통할 때 교환되는 데이터의 의미, 형식, 구문을 인지할 수 없기 때문에 프로토콜과 같은 상호약속에 의해서 데이터를 주고 받았다. 프로토콜과 같은 데이터 교환 방식에 기반한 응용은 해당 프로토콜에 종속이 된다. 그 이유는 데이터를 교환할 때 데이터의 의미, 형식, 구문을 해당 프로토콜에 적합하게 설계, 구현하여야 하기 때문이다. 또한 다른 프로토콜과는 호환성 문제를 발생시킨다. 이러한 문제는 데이터를 설명하는 메타데이터가 표준화되지 못하고 각각의 프로토콜이나, 데이터베이스의 스키마에 종속되어 있기 때문이다[1].

데이터 레지스트리는 이러한 문제를 해결하고자 제안된 것으로서, 데이터 레지스트리는 메타데이터를 데이터 요소라는 표준화된 형태로 저장하고 있다.

그러나, 현재 데이터 레지스트리는 데이터 요소를 유통하기 위한 구체적인 방법은 제시하고 있지 않으며, HTML 을 이용한 인터페이스만을 제공하고 있다[2].

따라서, 본 연구에서는 데이터 레지스트리에 저장된 데이터 요소를 효율적으로 유통할 수 있는 방법으로 XML 을 이용하였으며, 이를 위해서 ISO/IEC 11179 에 기반한 데이터 레지스트리로 접근할 수 있는 인터페이스를 설계 및 구현하였다.

구현한 XML 생성 인터페이스는 데이터 레지스트리에 저장된 메타데이터를 XML 형식으로 만들어 줌으로써, 응용 프로그램들이 효과적으로 데이터 레지스트리에 접근하여 메타데이터를 이용할 수 있는 환경을 제공한다.

2. 관련연구

2.1 XML

현재 웹에서 주요 매개체로 사용되고 있는 HTML 은 하이퍼텍스트, 하이퍼미디어의 기능을 지원하고 누구나 사용할 수 있을 만큼 간단하고 이식성이 높아서 많은 인기를 얻고 있지만, 문서의 논리적인 구조와 그 내용에 대하여 문서 자체가 자신을 표현하는 능력은 부족하다. 반면, XML 은 국제 표준 마크업 언

어인 SGML 의 일부를 계승한 언어로써, HTML 과 비교해서 다음과 같은 특징을 지니고 있다[3]. 첫째, 태그를 자체적으로 정의함으로써 태그에 의미를 부여할 수 있는 방법을 제시한다. 둘째, HTML 이 지원하지 않는 객체지향 구조 또는 데이터베이스 스키마 구성을 위해 필요한 증첩을 허용한다. 셋째, 문서 구조의 검증을 위해 문법적 정의가 DTD 를 사용하여 문서 안에서 제공된다.

이러한 특징들과 함께, SUN, IBM, ORACLE 등은 XML 을 지원해주는 파서와 XML 관련 응용 프로그램들을 제공한다.

2.2 ISO 11179 기반 데이터 레지스트리

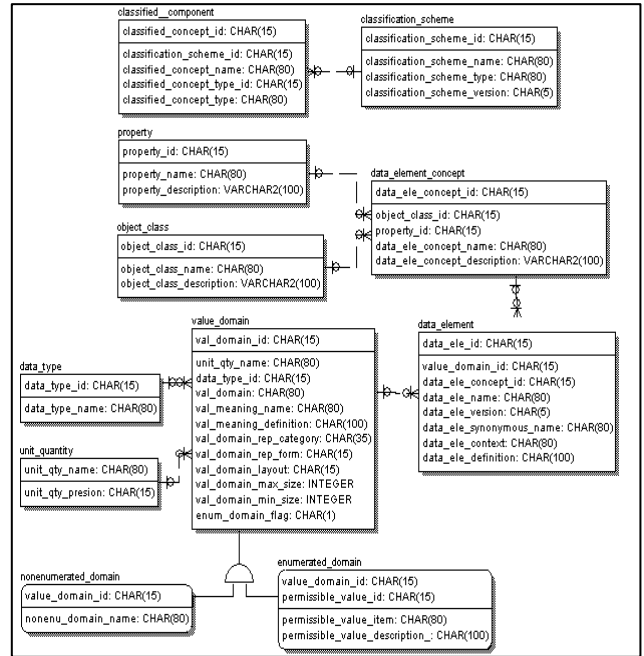
데이터 레지스트리는 데이터를 묘사, 검색, 분석, 분류하기 위해서 필요한 데이터의 특징들에 대한 사실을 보유하고 있는 장소이다[4]. 이러한 데이터 레지스트리의 구조를 X3.285 는 개념적인 메타 모델로서 6 개의 영역으로 나누어 표현하고 있다.

데이터 레지스트리에 저장되는 데이터를 기술하는 단위가 데이터 요소이다. 데이터 요소는 정의, 식별자, 표현 그리고 허용 가능한 값이 속성들의 집합에 의해서 명세 되어진 데이터의 단위이다[5]. 이러한 데이터 요소는 데이터 레지스트리에 저장되는 데이터를 기술하는 단위이다. 이 데이터 요소는 객체 클래스(Object Class), 특성(Property), 표현(Representation)의 3 부분으로 구성된다. 객체 클래스는 자동차, 사람, 고객, 주문, 사원 등과 같이 우리가 수집하거나 저장하려는 데이터를 지칭한다. 특성은 색, 모델, 성, 나이, 수입, 주소 등과 같이 사물을 구별하거나 설명하는데 사용되는 특징이다. 표현은 해당 데이터 요소가 가지는 값에 대한 도메인 및 표현 형태를 기술한다.

3. 데이터 레지스트리의 관계형 데이터베이스화

데이터 요소를 유통하기 위해 데이터 요소를 표준화된 형태로 저장하고 있는 데이터 레지스트리를 데이터베이스화 하여야 한다. 따라서, X3.285 가 제안하는 데이터 레지스트리 모델을 <그림 1>과 같이 데이터베이스 모델로 구현했다.

<그림 1>에서 표현된 것과 같이 데이터 요소 개념 관리(Data element concept administration) 영역은 Property Entity, Object_class Entity, Data_element_concept Entity 를 통합한 개념이다. 또한 분류(Classification) 영역은 Classified_componet Entity, Classification_scheme Entity 에 의해서 표현된다. Value_domain Entity, Data_type Entity, Unit_quantity Entity, Nonenumerated_domain Entity, Enumerated_domain Entity 는 개념 도메인과 값 도메인 관리(Conceptual and value domain administration) 영역을 표시한다. Data_element_concept Entity, Data_element Entity 는 데이터 요소의 관리(data element administration) 영역을 표시한다.



<그림 1> 데이터 레지스트리 메타모델의 관계형 모델 스키마 구조

4. XML 생성 인터페이스의 설계

4.1 메타데이터 기술을 위한 XML DTD 설계

데이터 요소를 표준화된 방법으로 유통시키기 위해서 XML 문서를 사용하는 것은 데이터 교환을 위한 상호작용을 완화 시키며, XML 을 처리하는 다양한 응용 프로그램을 이용할 수 있으며, XML 은 차세대 웹 표준 언어이므로 인터넷에 연결된 이질적인 시스템 사이에서도 정보를 교환 할 수 있다.

데이터 요소를 이러한 XML 의 장점을 활용해서 유통시키기 위해서는 XML 문서의 구조를 정의하는 DTD 를 정의해야 한다. 따라서 본 연구에서는 <표 1>의 표준화된 데이터 요소의 기본 속성을 이용하여 XML DTD 를 작성하였다[6].

식별(Identifying)	명칭(Name)
	식별자(Identifier)
	버전(Version)
	문맥(Context)
정의(Definitional)	정의(Definition)
표현(Representational)	표현범주 (Representation Category)
	표현양식 (Form of Representation)
	데이터 요소값의 데이터유형 (Datatype of data element values)
	데이터 요소값의 최대크기 (Maximum size data element value)

	데이터 요소값의 최소 크기 (Minimum size data element value)
	표현배치 (Layout of representation)
	허용가능값 (Permissible data element values)

<표 1> ISO/IEC 11179 의 데이터 요소의 일부 속성
 데이터 요소의 속성을 XML DTD 로 작성할 때는 데이터 요소가 가지는 5 개의 영역을 주 태그(Main Tag)로 정의하고 각각의 태그 영역 하부에 필요한 태그를 다시 정의하였다. 또한 속성의 필요성 여부와 기수(Cardinality)를 고려하여 태그가 XML 문서에 나타날 수 있는 횟수를 결정했다. <표 2> 는 생성된 XML DTD 의 일부를 표시한다

```

<?xml version = "1.0"?>

<!ELEMENT Data_element (Identifying, Definitional, Relational, Representational, Administrative)>

<!ELEMENT Identifying

(Name, Identifier*, Version*, Registration_Authority*, Synonymous_name*, Context*)>

<!ELEMENT Name (#PCDATA)>
<!ELEMENT Identifier (#PCDATA)>
<!ELEMENT Version (#PCDATA)>
<!ELEMENT Registration_Authority (#PCDATA)>
<!ELEMENT Synonymous_name (#PCDATA)>
<!ELEMENT Context (#PCDATA)>

```

<표 2> 데이터 요소 표현을 위한 XML DTD

4.2 XML 문서 생성을 위한 핵심함수 설계

일반적으로 개발자가 데이터 레지스트리 기반의 응용프로그램을 구현하고자 할 때, 데이터 레지스트리에 저장된 각 엔터티의 스키마 정보를 정확히 알고 있어야 하지만, 실제로 각각의 개발자가 데이터 레지스트리의 복잡한 스키마 정보를 정확히 파악하기는 어렵다. 또한, 데이터 요소의 속성 간에는 의존성 관계가 존재한다. 예를 들면, 이름(Name)과 식별자(Identifier)간에는 동일한 이름이 존재하는 조건하에서만 식별자를 필요로 하며, 문맥(Context)도 이름(Name)이 존재하는 경우에만 속성이 존재하게 된다. 버전(Version)과 같은 속성은 새로운 버전을 할당할 때 등록기관이 설정한 보수 규칙을 충족시키는 속성상 보완이 발생하는 경우에만 필수적인 속성이 된다. 이러한 속성에 관한 정보들은 XML DTD 에 정의하기 어려우므로 데이터 레지스트리의 제어 인터페이스가 필요하다.

따라서 본 연구에서는 핵심함수를 이용하여 데이터 레지스트리의 정보를 제어 할 수 있는

API(Application Programming Interface)를 구현했다. 정의된 핵심함수들은 데이터 레지스트리 제어함수와 XML Generator 로 구성된다. 제어함수는 데이터 요소와 데이터 요소 개념, 객체 클래스, 데이터 요소의 분류체계의 영역을 제어하는 함수이다. XML Generator 는 XML DTD 를 활용하여 XML 문서를 생성한다. <표 3>은 정의된 핵심함수들 기능에 따라서 5 개의 영역으로 표시한 것이다. 이러한 핵심함수 중 데이터 레지스트리 제어함수를 실질적으로 구현 할 때는 DBMS 에서 제공하는 스토어드 프로시저(Stored Procedure)를 사용했다.

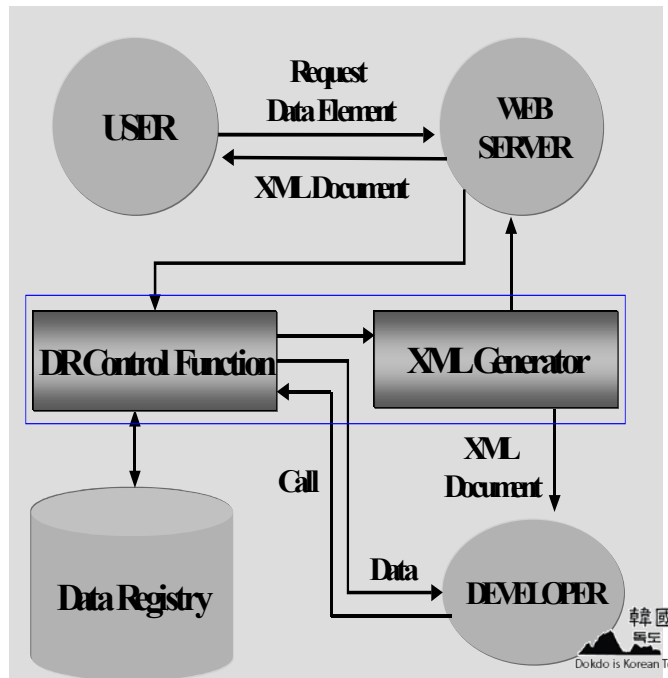
함수종류	기능
Manuplate_Data_Element	데이터 요소 관련 데이터들에 대한 조작기능을 수행한다.
Manuplate_Object_Class	객체와 관계된 데이터에 대한 조작기능을 수행한다.
Manuplate_Data_Element_Concept	데이터 요소 개념과 관계된 조작기능을 수행한다.
Manuplate_Classified_Scheme	데이터 요소의 분류체계 정보를 조작하는 기능을 수행한다.
XML_Generator	정의된 XML DTD 를 사용해서 메타데이터의 XML 문서를 생성한다.

<표 3> 핵심함수의 다섯 영역

5. XML 생성 인터페이스의 구현

5.1 XML 문서 생성 프로세스

핵심함수와 XML DTD 를 사용하여 데이터 요소를 XML 문서로 유통시키는 과정을 <그림 2>와 같이 표시했다.



<그림 2> 메타데이터의 XML 문서 생성과정

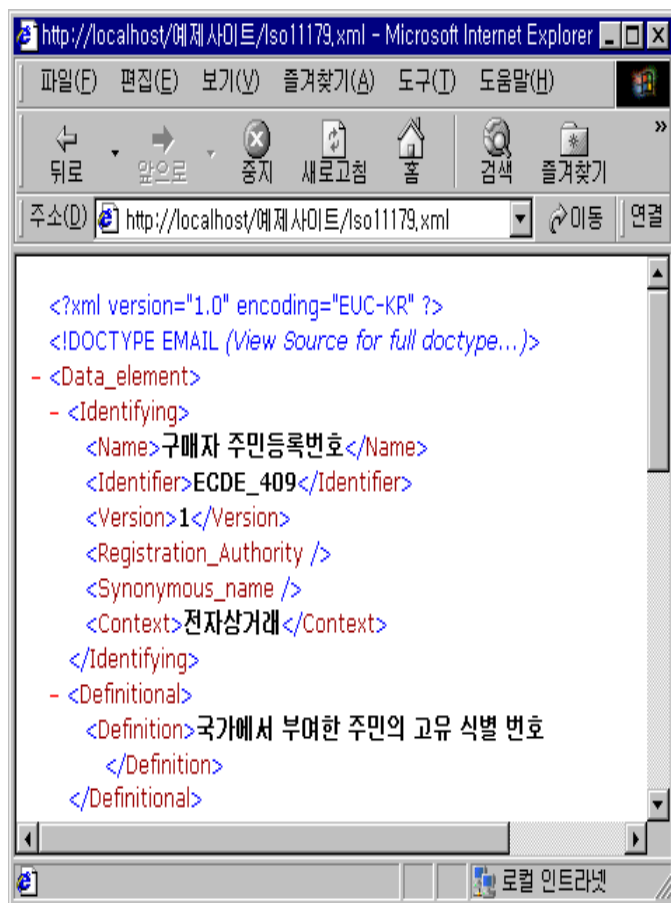
<그림 2>에서 DR Control Function 과 XML Generator 를 포함하고 있는 박스는 핵심함수를 표시한다. 핵심함수를 사용하여 데이터 레지스트리의 정보를 XML 문서로 생성하는 과정은 다음과 같다.

사용자가 웹을 통해 데이터 요소를 요구하면 웹서버의 처리 프로그램은 데이터 레지스트리의 제어함수를 이용하여 데이터 레지스트리의 검색결과를 XML DTD 와 함께 XML 생성기로 넘겨준다. XML 생성기는 정의된 XML DTD 형식으로 XML 문서를 생성해 준다.

개발자가 데이터 레지스트리에 접근하고자 하는 경우 제어함수를 통해서 직접 접근이 가능하다. 또는 XML 문서 형태로 작업하기를 원하는 경우에는 XML 생성기를 거쳐서 XML 문서로 결과를 얻게 된다.

5.2 구현 화면

<그림 3>은 마이크로소프트사의 DBMS 인 MS-SQL SERVER 과 웹서버 IIS 와 ASP 를 사용해서 데이터 요소의 속성을 웹을 통해서 XML 문서를 표현한 화면이다.



<그림 3> 데이터 요소의 XML 문서화 결과

6. 결론

메타데이터를 데이터 요소라는 표준화된 형태로 저장하고 있는 데이터 레지스트리는 메타데이터의 구체

적인 유통 방안을 제시하고 있지 않으며, 현재는 HTML 형식의 단순한 인터페이스만을 제공한다.

따라서 본 연구에서는 데이터 레지스트리의 메타데이터를 핵심함수와 XML DTD 를 사용하여 XML 문서로 유통 될 수 있는 인터페이스를 설계하고 구현했다.

향후에는 여러 개의 데이터 레지스트리들이 존재하는 경우 데이터 요소의 효과적으로 유통을 위한 인터페이스의 설계 및 구현방법에 대한 추가적인 연구가 필요하다.

참고문헌

- [1] <http://www.epa.org> 미국 환경청(EPA: Environmental Protection Agency)의 EDR(Environmental Data Registry)
- [2] Meta Messages in Electronic Data Interchange (EDI) C. Huemer & A.M. Tjoa
- [3] XML 소개, (주)언어기술, 1999.8.7 <http://ibase.co.kr/solution/xml>
- [4] "Metamodel for the Management of Shareable Data" ANSI X3.285 February, 1999
- [5] "Information technology - Specification and standardization of data element," ISO/IEC 1179, part1, FDIS, 1998
- [6] "Information technology - Specification and standardization of data element," ISO/IEC 1179, part3, IS, 1994