

메타데이터 스키마 관리를 지원하는 RDF 메타데이터 생성 시스템의 설계

성호상^o 강현철
중앙대학교 컴퓨터공학과

hssung@dblab.cse.cau.ac.k hckang@cau.ac.kr

Design of an RDF Metadata Generation System Supporting Metadata Schema Management

HoSang Sung^o Hyunchul Kang

Dept. of Computer Science and Engineering, Chung-Ang University

요 약

최근 인터넷 등 정보통신 기술의 발전으로 인해 정보량이 급증함으로써 이들 정보자원을 효과적으로 관리하고 검색하기 위한 메타데이터의 필요성이 증가하고 있다. 따라서 이러한 메타데이터를 생성하기 위한 효과적인 메타데이터 생성 시스템의 연구가 활발히 진행 중이다. 하지만 메타데이터를 생성하는 데 있어 정보자원의 종류에 따라 메타데이터 항목들에 대한 특성을 정의한 메타데이터 스키마의 활용은 미비한 실정이다. 본 논문에서는 여러 가지 자원기술 표준을 지원하는 메타데이터 스키마 관리 기법을 제안하고, 메타데이터 스키마를 이용한 RDF 표준을 따르는 메타데이터를 쉽고 빠르게 생성하는 메타데이터 생성 시스템의 설계를 제시한다.

1. 서론

최근 인터넷과 정보통신 기술의 발전으로 인해 정보자원의 양이 급증하고 다양화됨으로써, 정보자원을 정확하고 신속하게 검색하기 위한 메타데이터의 필요성이 증가하고 있다. 이러한 메타데이터는 자원에 대한 내용, 구조, 특징 등을 기술한 정보로서 자원에 대한 이해를 높이고 유용성을 판단할 수 있는 기준을 제공하여 자원의 이용을 활성화시킨다[1]. 이와 관련하여 Dublin Core[2], ANZLIC(Australia New Zealand Land Information Council), GILS(US Government Locator Service), EdNA(Education Network Australia), IMS(Instructional Management Systems), GEM(The Gateway to Educational Materials), vCard(vCard Electronic Business Card), AC(A-Core: Metadata about Content Metadata) 등과 같은 여러 가지 자원기술 표준[3]들과 HTML 메타데이터와 XML 을 이용한 RDF(Resource Description Framework)[4], XMI(XML Metadata Interchange), CDF(Channel Definition Format), MCF(Meta Content Framework) 등과 같은 메타데이터 표준들이

제안되었다. 이들 중에서 RDF 는 메타데이터의 인코딩, 교환, 재활용의 기반을 제공하기 위해 고안되었고, 메타데이터 클래스와 속성의 의미를 상호 충돌 없이 표현하는 작업에 필요한 구문구조를 갖추고 있다. 그리고 XML 네임스페이스 개념을 완벽하게 지원, 활용하는 RDF 는 상이한 메타데이터 클래스 및 속성 간의 호환성과 의미적 모듈화를 제공할 수 있는 기반이 된다[5]. 이러한 메타데이터를 활용하기 위한 메타데이터 생성기가 개발되어 있고, RDF 형식으로 기술하는 연구가 진행되고 있다[1][6][7][8]. 하지만 기존의 메타데이터 생성기들은 다양화되어 있는 정보자원을 메타데이터로 표현하는 데 있어, 메타데이터 스키마의 활용이 미비한 실정이며 메타데이터 스키마의 관리와 생성 및 편집에 어려움이 있다.

따라서 본 논문에서는 자원기술 표준 및 다양한 종류의 정보자원에 대한 메타데이터 속성들을 정의하기 쉽고, 정의된 메타데이터 스키마를 상호교환 편집하여 복합 메타데이터 스키마를 생성할 수 있는 관리 방법을 제안한다. 또한 메타데이터를 생성하는 데 있어 이들 메타데이터 스키마를 이용하여 메타데이터를 쉽고 빠르게 생성할 수 있도록 하는 RDF 메타데이터 생성

* 본 논문은 산학연 공동기술개발 컨소시엄 과제 지원에 의한 것임.

시스템을 설계한다.

본 논문의 구성은 다음과 같다. 2 절에서는 메타데이터 스키마 설계의 기본원칙과 기존 메타데이터 생성기에 대한 관련 연구를 기술한다. 3 절에서는 메타데이터 스키마 설계 기본원칙에 기반한 메타데이터 스키마 관리 방법을 기술하고, 4 절에서는 메타데이터 생성 시스템을 설계하고 메타데이터 생성 방법을 기술한다. 5 절에서는 결론을 맺고 향후 과제를 제시한다.

2. 관련연구

본 절에서는 메타데이터 스키마 설계의 기본원칙을 설명하고 기존 메타데이터 생성 시스템에 대해 기술한다.

2.1 메타데이터 스키마 설계의 원칙

Duval 등은 IEEE(Institute for Electrical and Electronics Engineers)의 LOM(Learnig Object Metadata) Working Group 과 DCMI(Dublin Core Metadata Initiative)가 오타와 메타데이터 회의를 통해 수립한 합의 내용에 모든 메타데이터 스키마 설계와 애플리케이션에 적용될 수 있는 중요한 원칙을 [9]에서 제시하였다[5]. [5]에서는 이를 다음과 같이 기술하였다. 첫째, 메타데이터의 모듈화는 메타데이터 스키마 설계자가 네임스페이스를 선언하고 일정 기관의 정책 또는 알고리즘에 따라 관리되는 어휘의 공식 집합으로 정의되는 것이다. 즉, 예로서 Dublin Core 메타데이터 구성요소 세트는 URI 로 웹상에 정의되고, 선언의 적용 범위내에 있는 모든 Dublin Core 요소는 접두어 "DC:"로 구분한다. 모듈화된 메타데이터는 구분적, 의미적인 상호운용성을 유지하면서 상이한 스키마에 속한 메타데이터 속성과 클래스들을 조합할 수 있는 바탕이 된다. 이는 메타데이터 모듈의 설계자가 특정 조합을 염두에 두지 않았더라도 다양한 모듈 혼합이 가능한 구조적 유연성을 제공하며, 새로운 복합 스키마로 통합될 수 있다는 것이다. 둘째, 확장성에서는 메타데이터의 상호운용성을 유지하면서 메타데이터의 새로운 구성 요소의 추가를 허용해야 한다는 것이다. 즉, 메타데이터 스키마에 적용되는 메타데이터 요소들이 모든 정보자원에 사용되는 일반적인 구성 요소도 있지만, 특정 정보자원에만 사용될 수 있는 구성 요소도 존재하기 때문에 특정 요구를 만족시키는 구성 요소의 확장이 필요하다. 셋째, 상세성은 기본 요소의 의미를 세분화하고 보다 분명히 하는 한정어를 추가하는 것이다. 또한, 요소가 취할 수 있는 값의 범위를 규정하는 기법 또는 가치 집합을 명시하는 것이다. 즉, 전자의 예로는 '작자', '작곡자', '감독' 등은 '생성자' 의 보다 세분화된 유형으로 사용하며 '생성자' 를 협의적인 용어로 쓰는 것이고, 후자의 예로는 날짜, 시간 등과 같은 구성 요소는 다양한 포맷을 가지므로 데이터의 표기 포맷을 명시해주어야 한다는 것이다. 넷째, 다국어 지원은 정보자원의 이용을 극대화하려는 정보화 시대의 요구에 부응하는 작업의 기본이 되어야 한다.

2.2 기존 메타데이터 생성 시스템

기존 메타데이터 생성 시스템으로는 메타데이터 작성자가 편집기를 통해 데이터 요소를 입력하고 특정 형식으로 메타데이터를 생성하거나 입력된 URL 의 페이지를 분석하고 데이터를 자동으로 추출하여 메타데이터를 자동으로 생성하는 도구들이 연구, 개발되고 있다. 먼저 Reggie[6]는 메타데이터 스키마와 기술 언어, URL 을 입력하여 메타데이터를 생성하고, Dublin Core 의 한정어를 데이터 요소에 사용할 수 있으며 각종 도움말을 제공한다. Dublin Core Metadata template[7]은 Dublin Core 의 모든 데이터 요소와 한정어를 사용하며, Dublin Core 자원기술 표준을 따르는 최소 수준의 메타데이터를 생성할 수 있다. DC-dot[8] 메타데이터 생성기는 자원의 URL 을 입력하면 해당 웹 페이지를 검색하여 Dublin Core 의 메타태그로 자동 생성하는 도구로서, Dublin Core 의 기본 데이터 요소만을 사용하고 한정어를 사용하지 않는다. 필요한 경우 데이터 요소에 관련 사항을 입력하여 메타데이터를 재 생성할 수 있다. 국내 관련 연구로는 SeriCore 메타데이터 편집기[10]가 있으며, SGML 기반의 메타데이터 문서를 저장, 관리, 삭제할 수 있는 메타데이터 관리 시스템이다. 그리고 [1]에서 제안된 시스템은 URL 을 입력하여 논문에서 제안된 데이터 요소 관련사항을 웹 사이트로부터 추출하여 RDF 메타데이터를 자동 생성한다.

본 연구에서는 메타데이터의 효율적인 생성을 위하여 메타데이터 스키마를 상호교환 편집하여 복합 메타데이터 스키마를 생성할 수 있는 관리 방법을 제안하고, 메타데이터 스키마를 활용하여 자원기술 표준 및 다양한 종류의 정보자원에 대한 메타데이터 구성 요소 확장이 가능한 RDF 메타데이터 생성 시스템을 설계한다.

3. 메타데이터 스키마 관리 기법

```

CREATE TABLE MetaSchema(
  ID NUMBER, // Metadata Schema ID
  MetaSchemaName VARCHAR2(256), // Metadata Schema 이름
  MetaSchemaLabel VARCHAR2(256), // Metadata Schema Label
  MetaSchemaType VARCHAR2(256), // System or User-Defined
  ObjectType VARCHAR2(256), // 웹 컨텐트 타입(jm, html, hwp, doc...)
  Description VARCHAR2(256), // Metadata Schema의 위치 명시
  Prefix VARCHAR2(256), //어휘
  AttributeSet VARRAY(100), // Metadata Schema 에 해당하는 메타데이터 항목
)

CREATE TABLE MetaSchemaItem(
  ID NUMBER, //메타 데이터 ID
  ItemName VARCHAR2(256), //메타 데이터 항목명
  ItemLabel VARCHAR2(256), //메타 데이터 라벨
  Description VARCHAR2(256), //구성 요소의 위치 명시
  Prefix VARCHAR2(256), //어휘
  ItemDataType VARCHAR2(256), //데이터 타입
  SystemFlag CHAR(1), //시스템 메타데이터 항목 인지 아닌지 체크
  RequestFlag CHAR(1), //반드시 요구되는 값인지 아닌지 체크
  BagFlag CHAR(1), //여러 번 나올 수 있는 값인지 체크
  DefaultValue Def_Value_Type, //기본값의 포인터
  SubItemSet VARRAY(50) //하위 데이터 항목 명시
)

CREATE TYPE Def_Value_Type AS OBJECT
(
  ID NUMBER, //기본값 ID
  Def_Value VARCHAR2(256) //기본값
)
    
```

(그림 1) 메타데이터 스키마 저장 구조

메타데이터 스키마는 정보자원의 종류에 따라 메타데이터 항목들에 대한 특성을 정의한 것이다. 이런 메타데이터 스키마는 각 정보자원의 특성에 맞게 정의되어 메타데이터 생성 시 메타데이터를 입력할 수 있는 폼을 자동으로 생성하도록 한다. (그림 1)은 메타데이터 스키마 저장을 위한 저장 구조를 나타낸 것이다. (그림 1)에서 MetaSchema 테이블은 정보자원의 종류에 따라 하나의 메타데이터 스키마를 분류 정의하도록 하며, MetaSchemaItem 테이블은 각 메타데이터 스키마에 포함된 메타데이터 구성 요소들의 특성을 정의한다. 이들 메타데이터 스키마 저장 테이블들은 2.1 절의 메타데이터 스키마 설계의 원칙을 반영한 것이다. 즉, MetaSchema 테이블의 Description 항목은 메타데이터 스키마 위치를 저장하고, Prefix 항목은 네임스페이스 약어를 명시한다. MetaSchemaItem 테이블에서도 구성 요소의 네임스페이스를 명시함으로써 메타데이터의 모듈화를 지원한다. 그리고 MetaSchema 테이블의 AttributeSet 항목을 배열을 통해 구성 요소 세트를 정의함으로써 메타데이터 스키마 구성 요소를 추가, 삭제할 수 있는 확장성을 지원하고, MetaSchemaItem 테이블의 SubItemSet 항목은 기본 요소의 의미를 세분화할 수 있는 원칙 등을 지원한다. (그림 2)는 (그림 1)의 메타데이터 스키마를 저장하기 위한 메타데이터 스키마 저장 구조에 Dublin Core 자원기술 표준 구성 요소를 가지는 메타데이터 스키마 저장을 나타낸 것이다. 저장된 메타데이터 스키마는 생성, 편집, 삭제가 가능하며, 로컬 사이트 또는 다른 사이트의 메타데이터 스키마를 호출하여 상호교환 편집함으로써 복합 메타데이터 스키마의 생성도 가능하다.

ID	MetaSchema Name	MetaSchema Label	MetaSchema Type	Objec tType	Description	Prefix	AttributeSet
1	Dublin Core	Dublin Core	System	html	http://dublin.csa.....	DC	1,2,3,4,5,6...
2	Person	Person	User	null	http://dlab.csa.oau.....	PS	19,20,21.....

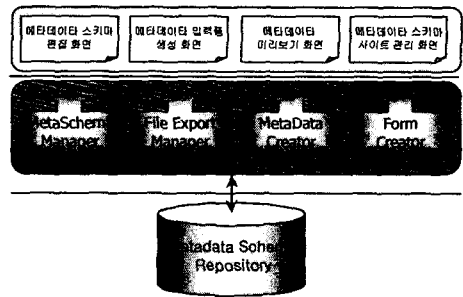
ID	ItemName	ItemLabel	Description	Prefix	ItemData Type	System Flag	Request Flag	Deflag	Default Value	SubItemSet
1	Title	Title	http://dublin.csa...	DC	text	true	true	true	null	null
2	Creator	Creator	http://dublin.csa...	DC	text	true	true	true	null	null
3	Date	Date	http://dublin.csa...	DC	Date	true	false	false	null	null
4	Type	Type	http://dublin.csa...	DC	text	true	false	false	null	null
5	Format	Format	http://dublin.csa...	DC	text	true	false	false	null	null
6	Language	Language	http://dublin.csa...	DC	text	true	false	false	null	null
19	Name	Name	http://dublin.csa...	PS	text	false	true	false	null	null
20	SSN	Social Security Number	http://dublin.csa...	PS	text	false	true	false	null	null
21	SEX	SEX	http://dublin.csa...	PS	text	false	true	false	null	null

ID	Def_value
1	Korean
2	French
3	English
4	German

(그림 2) Dublin Core 메타데이터 스키마 저장 예

4. 메타데이터 생성 시스템 설계 및 생성 방법

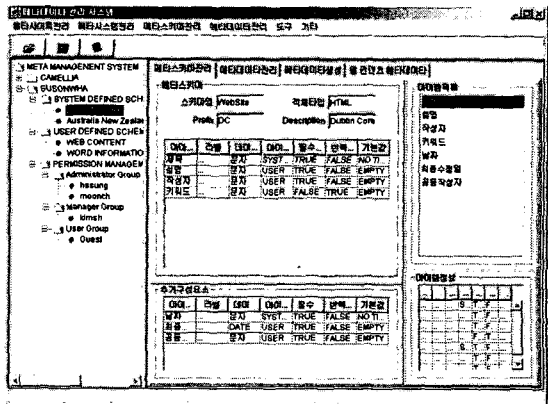
메타데이터 생성 시스템은 (그림 3)의 모듈들로 구성되어 있다. 먼저 MetaSchema Manager 모듈은 메타데이터 스키마를 저장, 편집, 삭제하는 기능을 수행하며,



(그림 3) 메타데이터 생성 시스템 구성도

메타데이터 스키마를 다른 사이트와 교환, 편집, 복합 생성하기 위한 사이트 정보를 관리한다. FileExport Manger 모듈은 저장된 메타데이터 스키마를 표준 RDF 스키마 형태로 파일 시스템에 변환 출력하거나 또는 생성된 메타데이터를 RDF 형식의 메타데이터 표준 형태로 출력한다. Form Creator 모듈은 저장된 메타데이터 스키마를 분석하여 메타데이터를 입력할 수 있는 폼을 자동 생성한다. Metadata Creator 모듈은 폼에 입력된 메타데이터 정보를 분석하여 RDF 표준을 따르는 메타데이터를 생성한다.

본 절에서 제시한 메타데이터 생성 시스템은 (그림 4)과 같은 형태의 화면을 제공한다. 이 시스템의 메타데이터 생성 방법은 다음과 같다. 먼저 메타데이터 스키마가 있는 사이트를 등록하고 정보자원의 종류에 따라 정의된 메타데이터 스키마를 선택한다. 만약 정보자원에 맞는 메타데이터 스키마가 없을 경우에는 로컬 사이트의 메타데이터 스키마를 편집하여 새로운 메타데이터 스키마를 생성하거나 다른 사이트의 메타데이터 스키마를 등록하여 상호교환 편집 등을 통해 새로운 메타데이터 스키마를 생성한다. (그림 4)에서는 이러한 메타데이터 스키마를 관리, 편집하는 것을 나타내고 있다. 정보자원에 맞는 메타데이터 스키마가 선택된 후 메타데이터 생성 메뉴를 통해 Form Creator 모듈은 저장된 메타데이터 스키마를 분석하여 메타데이터 입력 폼을 자동으로 생성한다. 메타데이터 생성



(그림 4) Metadata Schema 편집기

자는 정보자원의 메타데이터에 해당하는 값을 입력함으로써 메타데이터를 생성할 수 있다.

5. 결론 및 향후 연구과제

본 논문에서는 메타데이터 스키마 관리 기법을 제안하고 메타데이터 스키마를 이용하여 메타데이터를 쉽고 빠르게 생성하기 위한 메타데이터 생성 시스템의 설계를 제시하였다. 메타데이터 스키마 관리 기법에서는 2.1 절의 메타데이터 스키마 설계 원칙을 반영하였으며, Dublin Core 같은 자원기술 표준을 정의할 수 있도록 하였다. 또한, 다른 사이트와 로컬 사이트의 저장된 메타데이터 스키마를 호출함으로써 상호교환, 편집 가능하도록 하여, 새로운 복합 메타데이터 스키마를 생성할 수 있다. 저장 관리되는 메타데이터 스키마는 정보자원의 종류에 따라 선택되어 메타데이터를 입력할 수 있는 폼을 자동으로 생성하는 데 이용된다. 메타데이터 입력 폼을 통해 입력 받은 데이터로 생성된 메타데이터는 상호운용성을 갖는 표준 RDF 형식이다.

향후 연구과제는 본 논문에서 제시한 메타데이터 생성 시스템의 모듈을 확장하여 메타데이터 관리 시스템을 구현하는 것이다. 또한, 기계와 인간이 동시에 의미의 혼선 없이 전자정보를 교환할 수 있는 시맨틱 웹[11]의 연구가 활발해짐에 따라 이들 특징을 메타데이터 생성 시스템에 반영하는 것이 필요하다.

참고문헌

- [1] 이미경 외, "웹사이트 관리를 위한 RDF 메타데이터 생성시스템," 정보과학회논문지, 제 28 권, 4 호, 2001. 4, pp. 346-357.
- [2] Dublin Core Metadata Initiative., <http://dublincore.org/>.
- [3] Metadata.Net., <http://metadata.net/>.
- [4] O. Lassila et al., "Resource Description Framework(RDF) model and syntax specification," W3C, <http://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>, 1999.
- [5] 오삼균, "디지털도서관에서의 메타데이터 역할," 정보과학회지, 제 20 권, 8 호, 2002. 8, pp. 45-57.
- [6] Distributed System Technology Center, "Reggie: the metadata editor," <http://metadata.net/dstc/>.
- [7] Nordic Metadata Project, "Dublin Core Metadata Template (Creator)," <http://linnea.helsinki.fi/meta/>.
- [8] A. Powell, "DC-dot(Dublin Core Generator)," UKOLN, University of Bath, <http://www.ukoln.ac.uk/metadata/dcdot/>.
- [9] E. Duval et al., Metadata Principles and Practicalities., <http://www.dlib.org/dlib/april02/weibel/04weibl.html>.
- [10] 정효택 외, "Web 상의 전자문서를 위한 메타데이터 모델의 제안 및 관리 시스템의 개발," 정보처리학회 논문지, 제 5 권, 제 4 호, 1998. 4, pp. 924-940.
- [11] W3C, "Semantic Web," <http://www.w3c.org/2001/sw/>.