
방송 콘텐츠 메타데이터를 위한 온톨로지 개발

함종완 · 백승일 · 김남훈 · 정회경

배재대학교

The ontology development for broadcast content metadata

Jong-wan Ham · Seung-il Baek · Nam-hoon Kim · Hoe-Kyung Jung

Paichai University

E-mail : (jongwanham, bsi851, nhkim, hkjung)@pcu.ac.kr

요 약

온톨로지(Ontology)란 인간이 사물을 분류할 때 어떠한 개념적인 의미로 분류를 하는 것으로부터 출발하여, 현상이나 개념에 의미를 두어 객체화 하는 것을 말한다. 또한, 공학에서의 온톨로지는 공유하는 개념과 형식적이고 명확한 명세로 그 의미를 부여 하는 개념적인 것이 공학에서의 온톨로지이다. 방송 멀티미디어로 특화된 표준인 TV-Anytime과 멀티미디어 전반을 표현하는 MPEG-7의 표준을 바탕으로 방송용 멀티미디어 표현방법에 사용할 수 있는 온톨로지 개발이 요구된다.

이에 본 논문에서는 TV-Anytime의 멀티미디어 표준과 MPEG-7의 멀티미디어 표준을 기반으로 방송용 콘텐츠에 대한 온톨로지를 개발 하였으며, 온톨로지를 표현하기 위한 언어로 OWL(Web Ontology Language)과 RDF(Resource Description Framework)을 사용하여 방송 콘텐츠 메타데이터를 표현하는 온톨로지를 개발하였다.

ABSTRACT

The Ontology is the concept of human beings in any sense to classify things as departure from the classification based on the meaning at the present situation and concepts as the object of the Ontology. also, the technology is shared the conceptualization in a formal ontology as a meaning to give a clear description to the conceptual engineering in the ontology. Specialized multimedia broadcasting of TV-Anytime standard and multimedia representation of the overall standard of the MPEG-7 based on the concept of broadcasting multimedia presentations that can be used to develop an ontology who is required.

In this paper, in TV-Anytime and MPEG-7 multimedia standard for broadcasting multimedia contents based on standards developed for the ontology, an ontology language for expressing OWL (Web Ontology Language) and RDF (Resource Description Framework) and the broadcast contents using To represent metadata of the system who has been developed.

키워드

Ontology, TV-Anytime, MPEG-7, 멀티미디어 표준 융합

I. 서 론

최근 IPTV나 모바일 시장에서 각종 방송 및 멀티미디어 콘텐츠를 소비 하려는 경향이 급속도로 발전하고 있다. 그리고 수 많은 멀티미디어 콘텐츠가 각종 단말에서 그 소비가 증가하고 있으며, 이에 대한 콘텐츠들의 표현에 대한 연구가 활발히 진행되고 있다. 본 논문에서는 멀티미디어

콘텐츠를 표현하는 메타데이터인 MPEG-7[1]과 방송용 멀티미디어 메타데이터를 표현하는 TV-Anytime[2]이 있다. 두 표준은 멀티미디어 메타데이터를 표현하는 방법이 서로 달라 하나의 방송 콘텐츠 메타데이터를 표현하기위한 어려움이 있다. 이에 본 논문에서는 두 표준의 메타데이터 표현의 유사점을 찾아 방송 콘텐츠 메타데이터를 위한 온톨로지를 개발하였다.

II. 관련 연구

2.1 MPEG-7

MPEG-7은 MPEG(Moving Picture Experts Group)에서 그 표준을 제정하고, 멀티미디어 콘텐츠를 정의하는 인터페이스이며, 동화상, 정지영상, 음성, 텍스트 등의 각종 멀티미디어 데이터의 메타데이터를 표현하기 위한 표준이다. 본 논문에서는 MPEG-7에서 표현할 수 있는 메타데이터 중 동영상 데이터와 음성데이터를 표현에 관하여 온톨로지를 개발하기 위한 연구를 진행하였다.

2.2 TV-Anytime

TV-Anytime은 방송 매체에 대한 규정 및 표준이며, TV-Anytime은 Phase-1, Phase-2 로 구분한다. Phase-1은 단방향 환경에서 방송 프로그램 및 메타데이터를 전송하고, 추가적으로 양방향 네트워크를 통해 메타데이터를 획득할 수 있는 환경에서 메타데이터를 이용하여 저장된 데이터를 효율적으로 탐색, 선택, 획득 및 소비 할 수 있는 서비스를 제공한다. Phase-2는 TVA-1(TV-Anytime의 Phase-1)의 표준을 확장하여 홈 네트워크 환경에서 단말간의 콘텐츠 공유(Sharing) 및 재분배, 다양한 사용자 환경에 맞는 콘텐츠를 제공하는 타겟팅(Targeting) 등의 서비스를 제공한다.

2.3 Ontology

온톨로지(Ontology)란 인간이 사물을 분류함에 있어 사물에 의미를 두는 것으로 시작하여 그 사물에 대한 개념을 확립하고 그 의미를 부여하는 과정으로 이다. 공학에서의 온톨로지는 의미를 구체화하여 표현을 하는 것이며, 의미를 구체화 하는 메커니즘으로는 비형식적(Informal)인 방법과 구조화된 비형식적(Structured-Informal)방식, 반 형식적(Semi-formal)인 방식, 형식적(Formal) 방식이 존재한다. 다른 방법으로는 적용 범위에 따라 상위 하위 개념으로 분류할 수 있다. 상위 온톨로지로는 다양한 영역에 적용될 수 있는 기본적인 보편적인 개념 이며, 하위 온톨로지는 보편성을 추구하지 않는 경계가 확실한 특정 영역의 현실 세계를 구체화 하는 것을 말한다.

2.4 OWL과 RDF

대표적인 온톨로지 언어로는 OWL(Web Ontology

Language)[3]과 RDF(Resource Description Framework)[4]가 있으며, OWL과 RDF는 XML[5] 형식으로 정의한다. RDF는 웹 자원을 기술하는 XML 기반의 언어로 W3C에 의해 개발되었다. RDF의 특징으로는 웹 자원에 대한 메타데이터들을 어휘 및 구문을 공통적인 규칙에 따라 컴퓨터가 이해할 수 있는 정보로 표현하고, 상호 운용을 가능하게 하는 언어이다. OWL은 RDF와 RDF Schema의 문제점을 보완하기 위해 W3C에서 개발된 온톨로지 마크업 언어이다. OWL은 OWL Lite, OWL DL, OWL Full의 세 종류로 나뉘고 후자로 갈수록 표현력이 우수하지만 현재까지는 이를 지원하는 완벽한 추론 시스템이 현재는 없고, 이를 극복하기 위하여 OWL DL이 각광 받고 있다.

III. 온톨로지 개발

3.1 온톨로지 설계

본 논문에서는 온톨로지를 설계하기 위해서는 대응되는 개념을 지식베이스로부터 만들었다. 온톨로지를 지식베이스로부터 개발한 방법으로는 첫 번째 기존 지식베이스를 온톨로지의 형식으로 변환하는 방법과, 두 번째 기존의 지식베이스로부터 유용한 정보를 추출하여 온톨로지 지식으로 변환하는 방법, 세 번째 각 대응되는 관계 정보를 확장과 동시에 유의어(Thesaurus)를 온톨로지 리 모델링하는 방법과, 네 번째 유의어의 BT(Broader Term)-NT(Narrower Term)/RT(Related Term) 관계를 사람에 의하여 정의한 규칙이나 패턴을 이용하여 그 의미관계를 분류 하는 방법과, 관계 되는 의미를 규칙기반의 is-a 관계 분류와 패턴에 대응되는 의미 관계부여, 지도식 의미관계로 각 대응되는 의미를 연결하여 온톨로지를 구성하였다.

3.2 온톨로지 구성

온톨로지를 구성하기 위하여 두 표준의 관계를 이해하여야 한다. 두 개의 멀티미디어 표준을 하나로 융합하기 위해서는 대응되는 개념과 그 개념을 포용할 수 있는 상위 개념이 존재하여야 한다. 그림 1의 MPEG-7과 TV-Anytime 관계도에서 첫 번째 개념 단계는 가장 상위의 개념으로 두 표준의 개념을 하나로 융합하기 위한 하나의 상위 개념으로 Multimedia Ontology라는 이름으

로 구성을 하였으며, 그 하위 온톨로지를 구성하기 위하여 상위 개념의 하위 온톨로지로 MPEG-7과 TV-Anytime 표준의 이름을 그대로 가져와 개념적 엘리먼트로 만들고, 온톨로지의 그 다음 하위 온톨로지로 두 표준에 대응되는 메타데이터 표현에 대한 이름으로 온톨로지를 구성하였다.

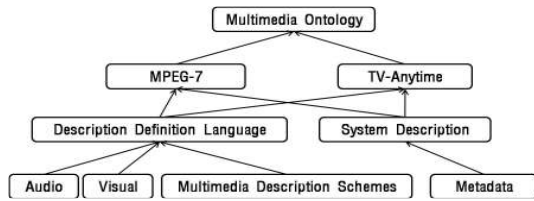


그림 1. MPEG-7과 TV-Anytime의 관계도

위의 관계도를 바탕으로 각각의 엘리먼트들을 RDF의 트리플관계로 표현 할 수 있다. 그림 2는 MPEG-7과 TV-Anytime의 엘리먼트들간의 관계를 트리플로 표현하였다.

Subject	Predicate	Object
Audio	subclassOf	Description Definition Language
Visual	subclassOf	Description Definition Language
Multimedia Definition language	subclassOf	Description Definition Language
Metadata	subclassOf	System Description
Description Definition Language	subclassOf	MPEG-7
System Description	subclassOf	TV-Anytime
MPEG-7	subclassOf	Multimedia Ontology
TV-Anytime	subclassOf	Multimedia Ontology

그림 2. MPEG-7과 TV-Anytime의 트리플 구성

트리플로 구성한 엘리먼트들을 RDF Schema로 구성할 수 있으며, 이 스키마를 시작으로 하위 온톨로지를 구성하였으며, 전체적인 온톨로지를 설계 할 수 있는 바탕을 만들었다. 그림 3은 그림 2의 트리플을 바탕으로 구성된 온톨로지의 스키마의 일부이며, MPEG-7과 TV-Anytime의 상위 하위 엘리먼트들을 가지고 전체적인 온톨로지를 설계하였다.

```

<xs:element name="RDF">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="Description" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute ref="resource" use="required"/>
  </xs:complexType>
</xs:element>
<xs:element name="Description">
  <xs:complexType>
    <xs:sequence>
      <xs:element ref="type"/>
      <xs:element ref="rdfs:subClassOf" minOccurs="0" maxOccurs="unbounded"/>
    </xs:sequence>
    <xs:attribute name="ID" use="required"/>
  </xs:complexType>
  <xs:restriction base="xs:string">
    <xs:enumeration value="Description Definition Language"/>
    <xs:enumeration value="System Description"/>
    <xs:enumeration value="MPEG-7"/>
    <xs:enumeration value="Visual"/>
    <xs:enumeration value="TV-Anytime"/>
    <xs:enumeration value="Multimedia Ontology"/>
  </xs:restriction>
  <xs:simpleType>
    <xs:attribute/>
  </xs:simpleType>
</xs:complexType>
</xs:element>
  
```

그림 3. 트리플을 바탕으로 구성된 스키마의 일부

IV. 결 론

최근 방송용 콘텐츠가 각종 단말에서 소비됨에 따라 이를 표현하기 위한 표준이 활발히 연구되고 있으며, 다양한 방법으로 콘텐츠를 표현하기 위한 메타데이터 또한 활발히 연구가 진행 중이다. 본 논문에서는 방송용 메타데이터를 표현하기 위하여 MPEG-7과 TV-Anytime의 동영상 메타데이터와 음성 데이터를 표현하는 메타데이터를 융합하기 위하여 방송 콘텐츠 메타데이터를 위한 온톨로지를 개발을 하였다.

방송 메타데이터 온톨로지는 서로 상반된 개념이 하나로 융합되기 때문에 두 메타데이터를 하나처럼 사용 할 수 있다는 것이 특징이다. 이와 같은 상반된 개념을 하나로 융합 하는 것은 MPEG-7과 TV-Anytime표준이 아닌 다른 개념들도 이와 같은 방법으로 융합하여 또 다른 새로운 개념을 만들 수 있기 때문에 그 가치는 온톨로지를 개발하는 것으로도 충분한 가치가 있을 것이다. 향후 연구 과제로는 방송 콘텐츠 메타데이터 온톨로지를 표현하는 프레임 워크를 구현 하는 것을 목표로 해야 할 것이다.

참고문헌

- [1] MPEG-7 표준 <http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm>
- [2] TV-Anytime 포럼 <http://www.tv-anytime.org/>
- [3] OWL 표준 <http://www.w3.org/TR/owl-features/>
- [4] RDF 표준 <http://www.w3.org/RDF/>
- [5] 정회경, 김창수, "알기 쉽게 해설한 XML", 이 한출판사, 2007