

## Linked Data 연계를 위한 SKOS 기반 용어 온톨로지 모델링 Term Ontology Modeling for Linked Data using SKOS

김평, 이승우, 서동민, 정한민, 성원경  
한국과학기술정보연구원

Pyung Kim, Seungwoo Lee, Dongmin Seo,  
Hanmin Jung, Won-Kyung Sung  
Korea Institute of Science and Technology Institute

### 요약

시맨틱 웹은 인간 중심의 데이터 표현을 위한 HTML 기반의 기존 웹과는 달리, 웹에서 데이터의 의미를 표현함으로써 다양한 어플리케이션 간의 데이터 상호 교환을 통한 데이터 통합, 재사용성 증대, 기계에 의한 자동화된 처리를 가능하게 해준다. 온톨로지는 데이터의 의미를 표현하기 위한 방법으로 식별자(URI) 기반의 리소스 명명을 통해 데이터의 의미를 표현하며, Linked Data는 RDF 형식의 데이터 간 링크를 통해 웹 데이터 간의 연계 및 활용할 수 있는 환경을 제공해 준다. 본 연구에서는 용어 정보의 효과적인 공유 및 연계를 위한 방법으로, SKOS 기반 용어 온톨로지 모델링을 통해 용어 정보가 Linked Data에 연계되기 위한 방법을 제시한다.

## I. 서론

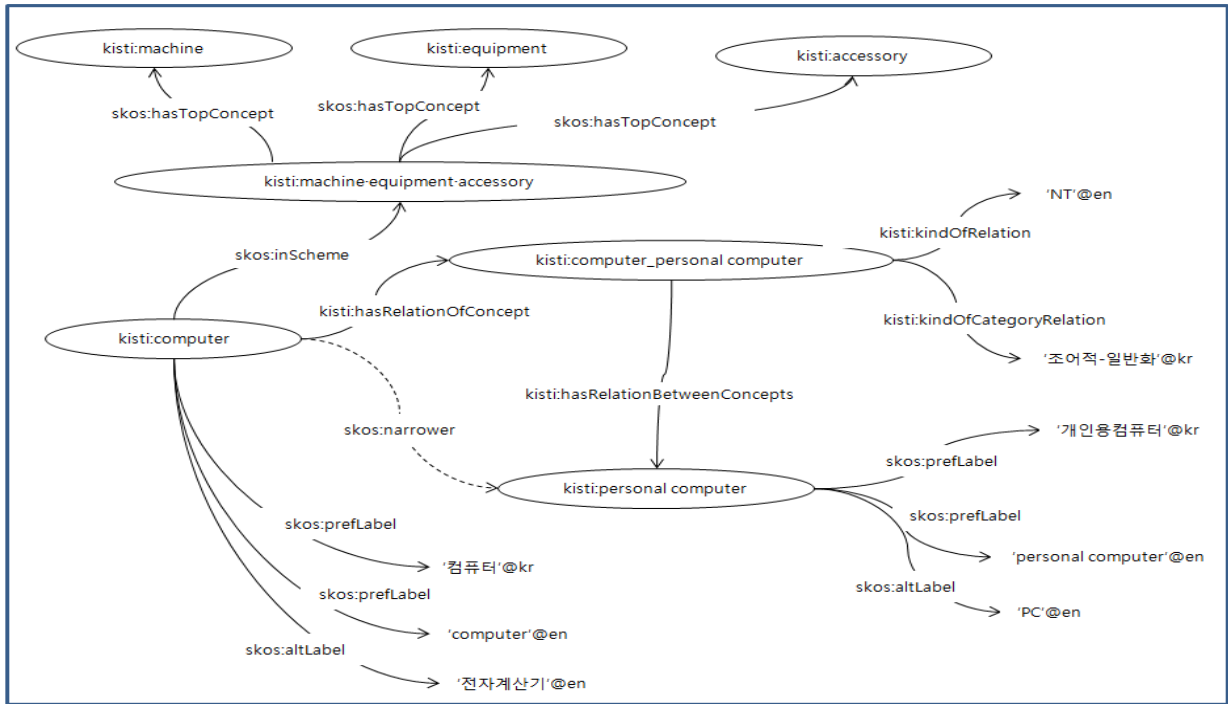
시맨틱 웹은 웹 환경에서 지식을 표현하고 공유하기 위한 기술로서, 온톨로지를 통해서 지식의 의미를 표현하고 식별자와 규칙을 통해서 정보를 연계하고 통합하는데 효과적으로 사용되고 있다. 최근 미국과 유럽을 중심으로 Linked Data 프로젝트를 통해 보유 자원 간 연계를 통해 부가 정보를 생성하고 서비스하기 위한 연구가 수행되고 있다[1]. 용어는 학분 분야와 사용 주체에 따라 서로 다르게 정의되고 활용될 수 있으며, 활용 목적에 따라 용어를 중심으로 다양한 관계가 생성될 수 있다. 최근 용어의 중요성이 부각되면서 동시 다발적으로 용어를 구축하고 활용하기 위한 연구들이 수행되고 있으며, 용어 구축하고 관리하기 위한 표준화 연구[2]도 진행되었다.

하나의 용어를 중심으로 용어의 의미와 대역 정보(영어, 기타 외국어), 용어 간 의미의 상·하 관계어, 관련어 및 유사어 관계 등이 생성될 수 있으며, 관계가 다양할수록 용어 정보를 구축하는데 막대한 비용이 소요

되기 때문에, 용어 정보를 공유하고 재활용·연계하기 위한 연구의 필요성이 점점 증대되고 있다. SKOS (Simple Knowledge Organization System)는 기계가 웹을 통해 지식 자원에 접근, 연계 및 공유를 위한 데이터 모델로, 보유 자원을 적은 비용으로 시맨틱 웹에 연계하기 위한 표준화된 방법을 제공한다. 본 연구에서는 KISTI가 보유한 과학기술 시소러스[3] 및 전문 용어를 Linked Data에 연계하기 위해서 SKOS를 기반으로 데이터를 모델링하고 연계, 활용하기 위한 방법을 제시한다.

## II. 관련연구

시소러스는 용어의 의미와 상·하 관계, 유사어와 관련어 관계를 표현하기 위해 사용되었지만, 시소러스로 구축된 용어를 웹을 통해 공유하고 연계하는 것은 쉽지 않기 때문에 시소러스를 표준화된 지식 표현 모델을 통해 체계화하는 연구들[4,5]이 수행되었다. 시소러스 구



▶▶ 그림 1. 개념 표현과 관계 정의를 위한 온톨로지 모델링

조를 토픽맵을 기반으로 구조 체계 매핑과 구조 체계 병합이라는 방법을 사용해서 표현한 연구[4]는 시소러스 이외의 관계들을 표현하기 위한 범용성과 외부 용어 연계 측면에서 한계를 가지고 있다. 또한 용어와 용어 관계를 범용적으로 표현하기 위한 주제어 공통 표현 모델에 관한 연구[5]도 수행되었지만 외부 용어와의 연계를 위한 방안은 제시되지 못하였다.

### III. SKOS 기반 용어 온톨로지 모델링

온톨로지는 지식의 의미를 명확하게 표현하고 공유하기 위한 지식 표현 모델로서 시맨틱 웹에서 보유 자원을 표현하기 위해서 사용된다. 하지만 온톨로지로 표현된 데이터가 웹을 통해 연계·활용되기 위해서는 클래스와 속성이 공유되고, 식별자를 통해 용어 간 연결점이 형성되어야한다. 본 연구에서는 온톨로지를 이용해서 용어 관계를 범용적으로 표현하고, Linked Data와 연계성을 확보하기 위해서 SKOS 기반 용어 온톨로지를 모델링하였다.

#### 1. 데이터 타입 속성(Data type property)

용어는 “skos:concept”를 사용해서 개념화 되며, 개념은 표1은 개념을 표현하기 위한 데이터 타입 속성 중 일부이다. SKOS의 3개 속성은 개념을 표현하고, KISTI의 2개 속성은 사용해서 개념 간 관계를 표현한다. 표 1의 SKOS의 3개 속성인 skos:prefLabel, skos:altLabel, skos:hiddenLabel 속성들은 서로 disjoint한 관계를 가지며, 같은 속성을 2개 이상 부여해서 개념을 표현할 수 있다. KISTI의 추가 속성들은 새로운 관계를 자유롭게 추가 정의할 수 있도록 만들어졌다.

표 1. 용어 속성

속성	의미
skos:prefLabel	개념의 대표어
skos:altLabel	개념의 대체 표현어
skos:hiddenLabel	의도된 표현어
kisti:kindOfRelation	속성 관계
kisti:kindOfCategoryRelation	관계 종류

kisti:kindOfRelation과 kisti:kindOfCategoryRelation은 개념 간 관계(BT/NT/RT 등)와 관계의 종류 정보(조어적-일반화/의미적-일반화 등)를 표현할 수 있다. 개념들은 Linked Data를 통해 다른 부가 데이터에 접근할 수 있도록 하기 위해서, DBPedia에 정의된 용어는 DBPedia의 식별자(URI)를 사용하고, 정의되지 않은 경우는 KISTI에서 부여하는 식별자를 사용하였다.

## 2. 객체 타입 속성 (Object type property)

개념 간 자유로운 관계 속성을 생성하기 위해서 관계 종류를 지원하는 속성을 추가하였다. 표 2는 개념 관계를 기술하기 위해 사용되는 속성 중 일부이다.

표 2. 관계 속성

속성	의미
skos:narrower	NT 표현
skos:broader	BT 표현
kisti:hasRelationOfConcept	개념 관계
kisti:hasRelationBetweenConcept	연관 개념

그림 1은 'kisti:computer' 개념을 중심으로 개념의 속성 정보와 다른 개념과의 관계를 보여준다.

'kisti:computer'은 기계(machine)-장치(equipment)-부속(accessory)의 하위 개념으로 'kisti:personal computer'와 '조어적-일반화' 관점에서 NT 관계를 형성하며, 추론 규칙을 통해 'skos:narrower' 관계로 확장할 수 있다. 관계 정보는 개념 관계와 연관 개념을 모델링을 통해서 개념의 속성과 개념 간 관계를 자유롭게 추가할 수 있으며, 추가된 개념 관계는 추론 규칙은 BT/NT/RT 관계를 SKOS 관계 정보로 표현해준다. SKOS를 기반으로 표현된 용어의 개념과 개념 간 관계는 Linked Data에 바로 연계·활용 될 수 있다.

## IV. 결론

용어 정보는 활용 목적에 따라 매우 다양한 관계들이 형성되며, 용어가 서비스에 유용하게 활용됨에 따라 구축에 소요되는 비용도 점점 커지고 있다. 시맨틱 웹은 다양한 분야에서 구축된 용어 정보를 공유하고 상호 연

계하는데 효과적인 기술이며, SKOS는 적은 비용으로 보유 자원을 구조화하고 표준화하여 시맨틱 웹 기술을 적용하는데 활용될 수 있다. 또한 Linked Data는 다양한 자원들이 상호 연계를 통해 고부가 정보를 생성하기 위한 데이터 환경을 제공하기 때문에 Linked Data 연계를 통해 용어의 활용성을 극대화하는 것은 물론 외부 자원 연계를 통한 부가 정보 획득이 용이하다. 본 연구에서는 KISTI가 보유한 시소러스와 전문용어를 Linked Data에 연계하기 위한 방법으로 SKOS 기반 온톨로지 모델링을 설계하였고, 본 모델링을 통해 다양한 용어가 공유·활용될 수 있는 기반을 제공하였다. 향후 연구에서는 용어와 유관 데이터(인력, 성과, 기관, 등)들이 Linked Data에 연계된 통합 온톨로지를 구축하고 서비스할 예정이다.

## ■ 참고 문헌 ■

- [1] Linked Data, "http://linkeddata.org/"
- [2] 김평, 성원경, 류범중, "전문 용어 표준화 지원 시스템 개발", 한국콘텐츠학회 학술대회, 제7권, 제1호, 2009.
- [3] 최석두 외 12, "과학기술 시소러스 구축 연구", ISBN 89-5884-633-1 93500, KISTI, 2007.
- [4] 남영준, "토픽맵을 이용한 시소러스의 구조화 연구", 정보관리학회지, 제22권, 제3호, 2005.
- [5] 정한민, 김평, 이미경, 성원경, "시맨틱 웹 서비스를 위한 온톨로지 기반 주제어 공통 표현 모델", 한글 및 한국어 정보처리 학술대회, 2008.