

외교부 LOD

SPARQL 도움말

목 차

1. 주요 용어.....	3
2. Sparql 도움말.....	10

1. 주요 용어

- Ontology
- OWL
- Data Property
- Linked Data
- FOAF
- RDF
- Class
- Object Property
- SPARQL
- SKOS
- RDFS
- Instance
- URI
- DC

Ontology(온톨로지)

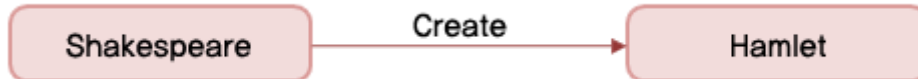
“T.Gruber : 온톨로지란 공유된 개념화(shared conceptualization)에 대한 정형화되고 명시적인 명세(formal and explicit specification)”

- 도메인 내에서 공유되는 데이터들의 개념화한 형식적이고 명세화한 규정
- 특정분야에서 사용되는 표준 어휘들의 모음
- 특정 분야에 대한 개념을 명확히 정의하고, 개념간의 상하구조와 기타 다른 관계 및 제약이 표현되어 있는 데이터베이스
- 시맨틱웹의 가장 핵심적인 개념으로서, 이를 표현하기 위하여 RDF, RDFS, OWL, TopicMaps 등의 언어가 있음

RDF

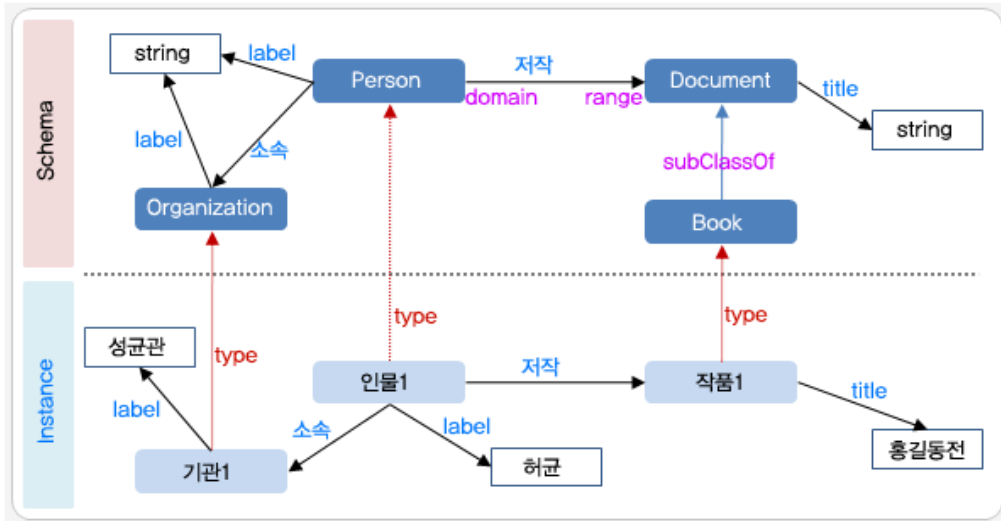
- RDF(Resource Description Framework)는 자원(Resources)을 속성(property)과 속성 값(property value)으로 표현하는 모델
- 'rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>'에서 정의
- RDF 는 주어(subject), 술어(predicate), 목적어(object) 형태의 Triple 로 구성됨
- 예시 :

Subject(Resource) - 주어(자원)	Shakespeare
Predicate(Property) - 술어(속성)	Create
Object(Value) - 목적어(속성값)	Hamlet



RDFS

- RDFS(RDF Schema)는 RDF 를 'Schema Vocabulary'로 확장한 것
- 클래스와 속성을 정의
- 연관된 자원의 그룹과 자원 사이의 관계를 묘사하기 위한 방법을 제공
- 주요 Vocabulary : Class, subclassOf, range, domain...
- 'rdfs:<http://www.w3.org/2000/01/rdf-schema#>' 에서 정의



OWL

- OWL (Web Ontology Language)은 RDF, RDFS 에 형식적인 의미를 가진 Vocabulary 를 추가하여 XML, RDF, RDFS 에서 지원하는 것보다 웹 콘텐츠의 기계적 해석을 더 용이하게 한 것
- Class 와 property 의 상속적 계층구조의 개념과 그 개념들 사이의 관계가 보다 명료하게 정의
- 관계들 간의 계층구조, 관계 인스턴스 내에서의 논리적 제약조건 등을 포함
- 주요 Vocabulary : Class, DatatypeProperty, ObjectProperty, inverseOf, sameAs ...
- 'owl:<http://www.w3.org/2002/07/owl#>'에서 정의

Class(클래스)

- 동일한 속성을 지니고 있어 하나로 묶을 수 있는 Individual 집합

Instance(인스턴스)

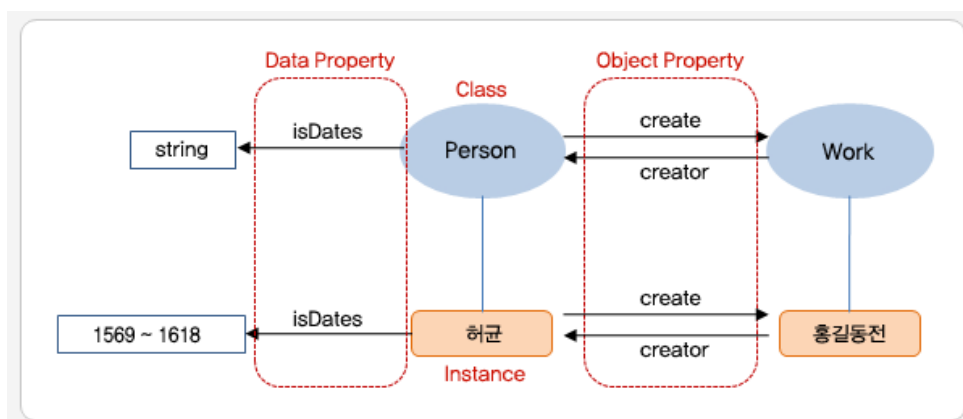
- Class 에 속한 개별적인 객체

Data Property

- Data Property(Datatype Property)는 데이터 속성으로서, Class 가 취해야 하는 데이터의 형식과 값을 정의

Object Property

- Object Property 는 관계에 대한 속성으로서, Class 요소간의 관계를 정의

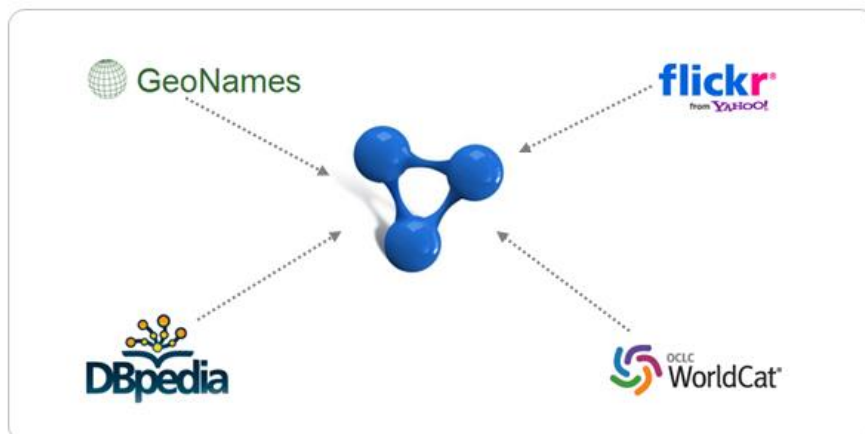


URI

- URI(Uniform Resource Identifier)는 웹상의 다양한 자원을 식별할 수 있는 표준 식별체계
- 데이터에 대한 정보를 기계가 처리하기 위해서는 데이터를 분명하게 구별할 수 있어야 하는데, 이 구별을 위하여 각 자원을 URI로 생성하여 사용함
- URI 는 자원의 Identity 를 유일하게 지칭할 수 있는 식별자로서, 네임스페이스 (Namespace), 접두어(Prefix), 식별자(identifier)로 구성
- 예) (황석영) <http://data.riss.kr/resource/Agent/000000023354>

Linked Data

- Linked Data 는 URI 를 통해 웹 데이터를 공개, 연결, 공유하는 방법
- 모든 데이터를 리소스 단위로 제작하여 가공되지 않은 raw data 를 이용자에게 공개, 사용자가 개별 요구에 맞게 데이터를 이용할 수 있도록 함
- URI 기반의 HTTP 역참조(dereference)를 통해 외부의 Linked Data 에 접근 가능
- RDF 를 통해 기계가독형으로 보유 정보를 웹으로 발행하며, 내외부의 정보를 질의하기 위한 질의 언어로서 SPARQL 을 이용



SPARQL

- SPARQL(Simple Protocol and RDF Query Language)은 온톨로지 질의(Query) 언어
- SPARQL 은 크게 PREFIX, {SELECT, ASK, DESCRIBE, CONSTRUCT}, WHERE 로 구성
- PREFIX 는 일반 데이터베이스에서 데이터베이스 선택과 비슷한 의미로써 사용할 데이터 셋을 지정할 때 사용
- Linked Data 는 SPARQL 을 Query 할 수 있는 SPARQL EndPoint 를 구축하여 사용자로 하여금 직접 Query 를 작성하여 데이터를 활용할 수 있도록 함
- 예) '김대중'과 연관된 외교문서는?

```
SELECT ?doc ?docname
WHERE {
    ?uri <http://www.w3.org/2004/02/skos/core#altLabel> ?uriname.
    filter( ?uriname = "김대중"^^xsd:string )
    ?uri <http://opendata.mofa.go.kr/mofadocu/relatedDoc> ?doc.
    ?doc <http://www.w3.org/2004/02/skos/core#prefLabel> ?docname .
}
ORDER BY ?docname
LIMIT 10
```

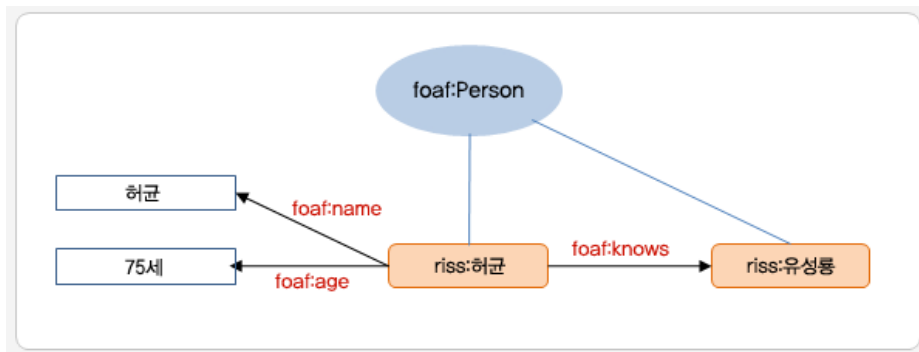
DC(Dublin Core)

- ISO 15836 으로 표준화된 메타데이터 요소 집합.
- 'dc: <http://purl.org/dc/elements/1.1/>'에서 정의

- 주요 Vocabulary : Title(표제), Creator(제작자), Type(유형), Contributor(기타제작자), Publisher(발행처), Date(날짜), Language(언어), Format(형식), Description(설명), Subject(주제), Relation(관계), Rights(권한), Source(출처), Coverage(범위)...

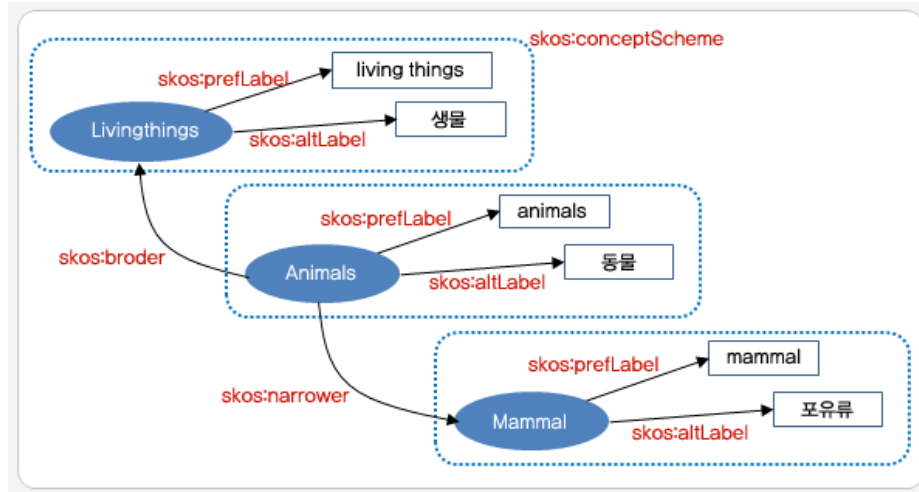
FOAF

- 사람과 사람간의 관계(Relationship)를 의미적으로 표현하기 위한 RDF Vocabulary.
- Class 와 Property 를 정의할 수 있음
- 'foaf: <http://xmlns.com/foaf/0.1>'에서 정의



SKOS

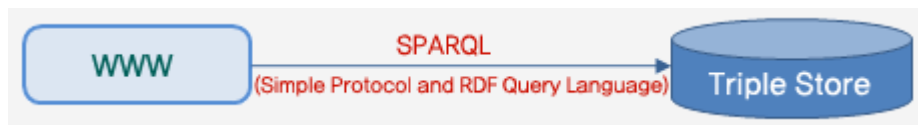
- SKOS (Simple Knowledge Organization System)는 시소러스, 계층구조, 분류체계, 주제어 표목 등을 표현하기 위해 제안된 지식분류체계 Vocabulary
- 다른 시맨틱웹 어휘와 통합, 연계 및 공유가 용이
- 개념(Concepts), 이름(Label), 개념간 의미관계 표현(Semantic Relationship)을 정의
- 'skos: <http://www.w3.org/2004/02/skos/core#>'에서 정의



2. Sparql 도움말

SPARQL 이란?

- SPARQL(Simple Protocol and RDF Query Language)은 온톨로지 질의(Query) 언어
- SPARQL 은 크게 PREFIX, {SELECT, ASK, DESCRIBE, CONSTRUCT}, WHERE 로 구성
- PREFIX 는 일반 데이터베이스에서 데이터베이스 선택과 비슷한 의미로써 사용할 데이터 셋을 지정할 때 사용



SPARQL 질의 구성

- BASE 정의부(선택사항), PREFIX 정의부, SELECT 파트, WHERE 절, LIMIT 절, OFFSET 절
- 변수 : '?'을 변수명 앞에 붙여 선언 예) ?x
- PREFIX : URL 을 별칭으로 지정

- SELECT : SPARQL Query 결과를 표현, 변수명이 위치
 - DISTINCT : 중복되는 결과는 제거하고 표현
 - WHERE : Query 하고자 하는 조건이 위치
 - '{' 로 시작해서 '}' 으로 조건절의 끝을 표현
 - Triple 구조로 표현 주어 술어 목적어로 표현하며 공백으로 분리
 - Triple 끝은 '.'으로 표시
 - Triple 에는 URL 이나 PREFIX:영문명, 변수가 위치
- ORDER BY : Query 결과의 순서를 지정
- LIMIT : Query 결과를 출력할때 몇 건을 보여줄 것 인지 지정
- OFFSET : Query 결과를 몇 번째부터 출력할 것인지 지정.(결과가 많을 경우)

SPARQL 예제

질의) '김대중'과 관련된 외교문서?

Query)

```
SELECT ?doc ?docname
WHERE {
    ?uri <http://www.w3.org/2004/02/skos/core#altLabel> ?uriname.
    filter( ?uriname = "김대중"^^xsd:string )
    ?uri <http://opendata.mofa.go.kr/mofadocu/relatedDoc> ?doc.
    ?doc <http://www.w3.org/2004/02/skos/core#prefLabel> ?docname .
}
ORDER BY ?docname
LIMIT 10
```

Query 결과)

doc	docname
http://opendata.mofa.go.kr/mofadocu/resource/Document/21445	A.I.(국제사면위원회) 대응 활동, 1985
http://opendata.mofa.go.kr/mofadocu/resource/Document/21055	Armacost, Michael 미국 국무부 정무차관 방한, 1985.3.20-22
http://opendata.mofa.go.kr/mofadocu/resource/Document/17211	Brandt, Willy 독일 사민당 총재 (전수상) 및 Dingels, Hans E. 사민당 국제부장 방한 초청계획, 1981-83
http://opendata.mofa.go.kr/mofadocu/resource/Document/52257	Graham, Katharine 미국 `Washington Post` 지 회장일행 방한, 1986.1.24-28
http://opendata.mofa.go.kr/mofadocu/resource/Document/52267	Jackson, Jesse 미국 목사 방한, 1986.12.12-14
http://opendata.mofa.go.kr/mofadocu/resource/Document/21545	Kalfoglou, Marios 그리스 `Ethnos`지 기자 방한, 1985.5.13-19
http://opendata.mofa.go.kr/mofadocu/resource/Document/52813	Lee Lam Thye 말레이시아 민주행동당(DAP) 사무차장 방한, 1981.4.25-5.7
http://opendata.mofa.go.kr/mofadocu/resource/Document/56673	Scalapino, Robert A. 미국 Berkeley 대학 동아시아연구소 소장 방한, 1987-89
http://opendata.mofa.go.kr/mofadocu/resource/Document/53245	Sigur, Gaston J. 미국 국무부 동아시아.태평양담당차관보 방한, 1987
http://opendata.mofa.go.kr/mofadocu/resource/Document/19874	UN 인권소위원회, 제 37 차. Geneva, 1984.8.6-31