

# 구조적 학술용어사전 “STNet”의 추론규칙 생성에 의한 의미 검색에 관한 연구\*

## A Study on the Semantic Search using Inference Rules of the Structured Terminology Glossary “STNet”

고영만 (Young Man Ko)\*\* , 송민선 (Min-Sun Song)\*\*\*  
이승준 (Seung-Jun Lee)\*\*\*\* , 김비연 (Bee-Yeon Kim)\*\*\*\*\*  
민혜령 (Hye-Ryoung Min)\*\*\*\*\*

### 목 차

- |  |                 |
|--|-----------------|
| 1. 서 론                                 | 4. RDF 온톨로지 변환  |
| 2. 구조적 학술용어사전 “STNet”의<br>텍사노미 클래스와 속성 | 5. 온톨로지 검증 및 평가 |
| 3. 온톨로지 구조 생성                          | 6. 결 론          |

### 초 록

본 연구의 목적은 이미 구축되어 있는 RDB 형식의 학술용어사전에 온톨로지 구조와 추론 규칙을 형성시킨 후 이를 의미 검색에 적용하는 상향(Bottom-up) 방식의 방법론을 제안하고 검증하는 것이다. 이를 위해 구조적 학술용어사전 “STNet”을 테스트베드로 삼아 Protege를 이용해 온톨로지 구조를 생성하고, 온톨로지 구조의 오류를 검증하였으며, STNet에 입력된 실제 데이터로 실험을 위한 테스트 데이터를 구축하였다. 그리고 추론 제한 규칙과 검증에 필요한 시나리오를 설정한 후, TBox 검증과 SPARQL 질의에 의한 결과 값을 평가하였다. TBox 검증 결과 본 연구에서 생성한 추론규칙이 모두 참으로 나타났으며, SPARQL 질의를 통한 결과값의 평가 결과 기존의 키워드 검색 수행에서는 파악하기 힘든 복잡한 검색 시나리오에 대해, 의미적으로 연관되는 용어를 효율적으로 조합해 검색 결과로 보여주는 것으로 나타났다.

### ABSTRACT

This study describes the Bottom-up method for implementation of an ontology system from the RDB. The STNet, a structured terminology glossary based on RDB, was served as a test bed for converting to RDF ontology, for generating the inference rules, and for evaluating the results of the semantic search. We have used protege editor of the ontology developing tool to design ontologies with test data. We also tested the designed ontology with the Inference Engine (Pellet) of protege editor. The generated reference rules were tested by TBox and SPARQL queries through STNet ontology. The results of test show that the generated reference rules were verified as true and STNet ontology were also evaluated to be useful for searching the complex combination of semantic relation.

키워드: 한국학술지인용색인, KCI, 구조적 학술용어사전, STNet, 텍사노미, 속성, 의미검색, 추론규칙, 온톨로지  
Korea Citation Index, KCI, Structured Terminology Glossary, STNet, Taxonomy, Property,  
Semantic Search, Inference Rule, Ontology

\* 이 논문은 2012년 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임  
(NRF-2012S1A5B4A01035131).

\*\* 성균관대학교 문과대학 문헌정보학과 교수(ymko@skku.edu) (제1저자)

\*\*\* 성균관대학교 정보관리연구소 연구원(songseory@skku.edu) (교신저자)

\*\*\*\* 성균관대학교 정보관리연구소 연구원(dino@skku.edu) (공동저자)

\*\*\*\*\* 성균관대학교 정보관리연구소 연구원(korkby@skku.edu) (공동저자)

\*\*\*\*\* 성균관대학교 정보관리연구소 연구원(miniplam@skku.edu) (공동저자)

논문접수일자: 2015년 8월 4일 최초심사일자: 2015년 8월 4일 게재확정일자: 2015년 8월 13일  
한국문헌정보학회지, 49(3): 81-107, 2015. [http://dx.doi.org/10.4275/KSLIS.2015.49.3.081]

## 1. 서론

### 1.1 연구 목적

지금까지 구축된 대부분의 학술용어사전은 해당 학문 분야의 전문용어를 단순하게 정의하는 형식을 벗어나지 못하고 있다. 용어간의 관계를 보다 상세하게 정의하기 위해 구축한 시소러스 역시 상위개념, 하위개념, 연관개념만을 정의하는 조직체계 상의 한계를 지니고 있어, 전문용어가 지니는 풍부한 내포적, 외연적 의미를 수용하는 것이 불가능하다. 오늘날의 연구자들이 해결해야 할 정보문제인 해당 연구 분야의 학술적 이해형성을 지원하기 위해서는 단순한 구조를 가지고 있는 용어사전이나 시소러스와 같은 기존 지식조직체계의 한계를 극복하는 새로운 형식의 지식조직체계가 필요하다.

본 연구에서 말하는 “구조적 학술용어사전”이란 새로운 형식의 지식조직체계의 하나로서, 학술 용어를 유사한 속성을 가진 개념범주(이하 클래스)에 따라 분류하고, 그 클래스가 가지는 속성을 체계화 한 다음, 그 속성에 따라 해당 학술 용어 하나하나의 의미를 정의하는 학술용어사전을 말한다. 구조적 학술용어사전에서 동일한 클래스에 속하는 학술용어는 해당 클래스의 개념속성과 관계속성을 통해 정의된다. 따라서 구조적 학술용어사전을 토대로 클래스별 속성 간에 형성되는 연관 관계를 부여하면 속성과 관련된 용어가 속해 있는 연구성과물 간의 연관 관계를 네트워크 형식으로 조직하는 것이 가능하게 된다. 그리고 이렇게 규정된 여러 연관 관계로부터 추론 규칙을 생성시킬 경

우 의미적 연관 검색이 가능한 지식베이스가 만들어진다.

입력한 쿼리의 나열에 의한 단순한 검색이 아닌 의미 기반 검색을 위한 시스템은 온톨로지를 필요로 한다. 온톨로지는 개념의 계층분류와 개념 간 관계 및 이를 토대로 한 공리와 추론규칙을 기반으로 구성된다. 추론 규칙이란 어떤 명제나 논리식으로부터 다른 명제나 논리식을 끌어내는 규칙을 말하며, 의미 기반 검색시스템에서는 새로운 사실을 자동으로 추출하거나 제약 조건에 맞지 않는 오류를 찾아내는 데 이용된다. 본 연구는 한국학술지인용색인(Korea Citation Index, 이하 KCI)의 키워드를 기반으로 관계형 데이터베이스(Relational Database, 이하 RDB)로 구축된 인문학, 사회과학, 예술체육 분야의 구조적 학술용어사전인 ‘STNet(<http://www.stnet.re.kr>)’을 의미 검색이 가능한 지식베이스로 발전시키기 위해, STNet의 RDB 구조에서 추론 가능한 의미 규칙을 도출하여 온톨로지를 구축하고, 이를 STNet의 데이터 검색에 적용해 봄으로써 의미 검색 방식의 유용성을 평가하는 것을 목적으로 한다.

### 1.2 선행 연구

의미 검색을 위한 추론규칙에 대한 이론적 연구는 주로 컴퓨터공학 분야에서 추론 시스템 개발의 토대 연구로 수행되었다. 온톨로지의 지식 표현과 추론에 따른 단계를 정의하고 정의된 단계에 따라 TBox와 ABox의 지식표현 구조와 SWRL 기반의 추론 규칙을 바탕으로 웹 온톨로지 모델링 방법을 제안한 김수경과 안

기흥(2008)의 연구, Open World Assumption (OWA) 기반의 온톨로지와 Closed World Assumption(CWA) 기반의 비단조 추론을 지원하는 규칙의 지식베이스를 통합하여 Open World와 Closed World 추론을 모두 지원하는 실질적인 추론 시스템을 제안하는 최정화와 박영택(2010)의 연구, Drexel 대학에서 개발한 OWLJessKB 추론규칙을 바탕으로 subClassOf 추론에 대한 그래프 레이블링의 효율성을 다시 검토한 김재훈과 박석(2011)의 연구, 구축된 온톨로지의 타당성 검증을 위해 공개되어 있는 12개 온톨로지에 추론기 HermiT, Pellet, FaCT++를 사용해 각 온톨로지 공리 검증을 진행한 Horridge, Parsia, Sattler(2009)의 연구, Pellet을 중심으로 OWL-DL의 추론 범위와 성능을 측정된 Evren Sirin 외(2009)의 연구 등을 들 수 있다.

또한 특정 도메인을 중심으로 온톨로지를 이용해 추론 시스템을 개발하기 위한 연구도 다수 진행되었다. 와인 온톨로지를 기반으로 Jena를 사용하여 추론 모형을 생성한 송우종과 김유성(2008)의 연구, 전자우편으로부터 텍스트 의미 분석처리 과정을 통해 얻어진 정보들을 전자우편 온톨로지에 기술하고 의미추론규칙을 적용하여 해당 전자우편의 범주 및 종류를 분류하도록 한 허정환, 정진우, 손진현, 이동호(2008)의 연구, 한의학 약재와 처방에 대하여 OWL 모델링을 기반으로 추론 규칙을 적용한 김상균 외(2009)의 연구, 모바일 기기 사용자의 개인화 서비스를 위해 OWL 모델링을 기반으로 추론규칙을 적용한 장창복, 김만재, 최의인(2012)의 연구, 자동차 정비를 위한 자동차 몸체, 엔진, 정비도구에 대한 온톨로지를 구축하고 추론한 길식, 박성철, 김준태(2012)의 연구가 여

기에 해당된다.

문헌정보학을 비롯한 사회과학 분야에서는 온톨로지의 의미 정보 표현에 초점을 맞춘 연구가 주로 진행되었다. 국제기구 온톨로지를 DAML+ OIL을 이용해 설계하고 검색시스템을 구현하여 검색 시간과 적합성 측면에서의 효율성을 일반 인터넷 검색엔진과 비교한 김현희와 안태경(2003)의 연구, 과학기술 컨퍼런스의 학술적 의미 추론 시스템에 관한 정상원(2008)의 연구, 이야기 쓰기를 돕는 본문 및 문장 검색 시스템의 구축을 위해 이야기와 단락 및 문장의 구조를 분석하고 색인 작성과 탐색 질문에 적용되는 언어 추론을 연구한 이태영(2009)의 연구, RDF/OWL 개체속성을 활용하여 청록과 관계 온톨로지를 구축하고 SPARQL을 이용한 질의 결과를 시각화하여 효율성과 편의성 측면에서 전통적 전거제어 업무와 비교한 강현민(2010)의 연구가 수행되었으며, 고영만과 송인석(2011)이 연구문헌의 지식구조를 반영하는 공리와 의미관계 추론 규칙으로 구성된 실험적 연구문헌 온톨로지 모형을 제시한 바 있다.

지금까지 수행된 연구들은 거의 대부분 구체적인 도메인의 온톨로지와 추론 규칙을 도출한 후 해당 구조에 맞춰 하향(Top-down) 방식으로 시스템을 구성하고 평가하는 방법론을 사용하였다. 본 연구는 이와 달리 이미 구축되어 있는 RDB 형식의 STNet을 테스트베드로 삼아 온톨로지 구조와 추론 규칙을 형성시킨 후 이를 STNet의 의미 검색에 적용해 보는 상황(Bottom-up) 방식의 방법론을 제안하고 검증하고자 한다.

### 1.3 연구 방법

본 연구는 STNet의 용어 간 관계 중에서 'X 용어 ↔ Y 용어' 형태로 연결되는 Y 용어가 개념속성 데이터로 가장 많이 연결되어 있는 '실존인물(코드명 y01-01)' 클래스를 연구 대상으로 삼았다. 따라서 본 연구의 실험데이터가 되는 온톨로지 구축 대상은 'y01-01 실존인물' 클래스에 할당된 용어 및 'y01-01 실존인물' 클래스의 속성과 연결되어 있는 9개 클래스에 해당하는 용어이다. Protege를 이용해 온톨로지 구축 대상에 대한 RDF 온톨로지 구조를 정의하고 Protege의 Pellet 추론기를 통해 온톨로지 구조의 오류를 검증하였으며, 해당 온톨로지 구조를 STNet에 적용한 후 TBox 검증을 통해 추론 규칙의 오류를 검증하고 SPARQL 질의를 통해 검색 결과 값에 대한 평가를 실시하였다.

본 연구의 절차는 크게 온톨로지 구조 생성과 검증, 실험 데이터의 RDF 온톨로지 변환 및 검증, 추론 제한규칙이 적용된 온톨로지에 대한 의미 검색 결과 평가의 세 단계로 구성된다.

첫 단계는 STNet 온톨로지를 구성하기 위해 주요 클래스 및 클래스 구조를 생성하고, 각

클래스의 속성에 대해 ObjectType Property와 DataType Property를 설정하여 각 속성의 Domain과 Range를 정의한 후, 추론기를 통해 각 클래스와 속성에 대한 기본적인 설정의 오류를 검증하는 단계이다. 추론기는 Tableaux 알고리즘을 기반으로 DIG 인터페이스를 지원하는 Description Logic 추론기인 'Pellet 추론기'를 사용하였다.

두 번째 단계는 구성된 온톨로지 구조의 논리적 오류가 없을 경우 STNet RDB 실험데이터를 RDF 온톨로지 변환기를 통해 RDF 데이터로 변환하는 단계이다. RDF 온톨로지 변환기로는 데이터의 변형이 빈번하고 새로운 데이터의 추가나 데이터들 간의 연결관계가 지속적으로 변화하는 동적인 RDB에 적합한 것으로 평가된 D2RQ를 사용하였다(고영만, 이승준, 송민선 2015). 먼저 실험 데이터로부터 일부 샘플 데이터를 추출하여 RDF 온톨로지 변환 후 동일한 추론기를 통해 온톨로지 구조에 적용된 샘플 데이터의 오류를 검증하였다. 샘플 데이터에 대한 오류 검증과 수정 후, 실험 데이터인 STNet 데이터베이스의 'y01-01 실존인물' 클래스에 할당된 용어 및 'y01-01 실존인물' 클래스의 속성과 연결되어 있는 9개 클



〈그림 1〉 STNet의 추론규칙 생성 및 의미검색 검증 절차

래스에 해당하는 용어, 9개 클래스의 속성과 연결되어 있는 용어를 RDF 온톨로지로 변환하여 반입하였다.

세 번째 단계는 반입된 데이터를 대상으로 추론 제한 규칙을 생성하고 STNet 온톨로지에 대한 검증과 평가를 하는 단계이다. TBox 검증을 통해 온톨로지 구조와 인스턴스를 갖춘 전체 데이터에 대한 추론 규칙의 오류를 검증하였으며, 시나리오에 의한 SPARQL 질의를 통해 결과값에 대한 의미 검색의 유용성을 평가하였다(〈그림 1〉 참조).

## 2. 구조적 학술용어사전 “STNet”의 텍사노미 클래스와 속성

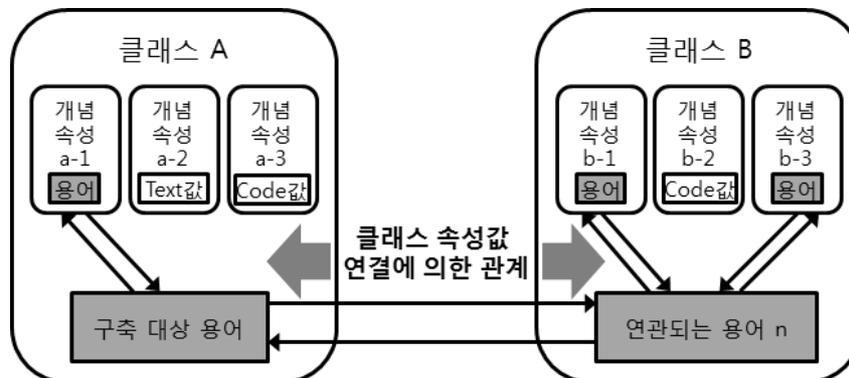
### 2.1 STNet의 텍사노미 구성

구조적 학술용어사전 “STNet”의 텍사노미는 STNet 전체 구조의 근간이 되는 것으로 인문, 사회 및 예술, 체육 분야 학술용어의 개념을 계층적으로 체계화한 것이다. STNet이 유사한

시스템과 구별되는 특징 중의 하나는 석사과정 이상의 전문인력이 직접 구축 작업을 수행하였으며, 특히 전문지식과 실무경험을 갖춘 분류 전문가가 구축 과정에서 나타나는 문제점의 반복적인 검토와 수정 작업을 통해 텍사노미를 상향(bottom-up)식으로 구축했다는 점이다(고영만, 김비연, 민혜령 2014).

STNet에서 정의되는 모든 용어는 텍사노미를 통해 클래스가 부여되고 해당 클래스의 속성에 따라 그 용어의 의미목록이 구축된다. 2015년 7월 6일 현재 STNet의 텍사노미는 최상위에 7개의 기본 클래스, 그 아래에 27개의 중위 클래스, 그리고 145개의 하위 클래스(1단계 하위클래스와 2단계 하위클래스로 세분)로 구성되어 있다. 클래스마다 각각의 코드와 클래스 명칭을 가지고 있으며, 각 클래스는 그 클래스를 대표하는 (개념)속성으로 구조화되고, 각 속성은 객체(object) 값, 텍스트(text) 값, 코드(code) 값으로 구분되는 속성값을 갖는다. 객체값은 STNet 데이터베이스 내의 용어에 해당한다.

STNet의 구축 대상 용어는 클래스의 속성값과 연결관계를 형성한다(〈그림 2〉 참조). 예



〈그림 2〉 STNet의 클래스(개념범주)와 개념속성의 연결 관계

를 들어 『문헌명』이라는 클래스에 속하는 용어와 클래스의 속성인 ‘저작자’, ‘저작년’ 등의 속성값 사이에 관계가 형성되는 것이다. 사례를 들어 설명하면, 『문헌명』에 속하는 용어 ‘안네의 일기’는 속성값인 저작자 ‘안네 프랑크’, 저작년 ‘1947년’과 연결 관계를 갖게 된다는 것이다.

STNet의 정의 대상이 되는 모든 학술용어는 텍사노미에 의해 클래스가 부여되고, 해당 클래스의 속성에 따라 정의되며, 관계명을 통해 클래스와 클래스, 클래스와 속성, 속성과 다른 속성, 용어와 용어 사이의 의미적 연결관계가 서술된다. 2015년 7월 6일 현재 STNet에는 55,589개의 용어가 구축되어 있으며 이중 개념 속성이 부여된 용어는 26,979개이고 다른 용어의 속성값으로 참조되어 연관관계를 형성한 용어는 28,452개이다.

## 2.2 실험 데이터의 범위

STNet의 용어 간 의미 관계는 크게 두 중

류로 나누어진다. 하나는 클래스와 클래스 속성 간의 관계로서 하나의 용어와 그 용어가 속한 클래스의 속성값이 연결되는 관계이며, 다른 하나는 용어와 다른 용어 간에 맺어지는 관계이다. 본 연구의 범위는 클래스와 클래스 속성 간의 관계이며, 클래스 속성은 온톨로지에서 ObjectType Property로 사용하는 용어형과 DataType Property로 사용하는 문자형, 숫자형, 날짜형 값으로 표현된다. 본 연구에서는 STNet의 클래스 속성 유형에 해당하는 용어, 문자, 숫자, 날짜형 모두를 온톨로지로 표현하여 의미검색을 위한 추론규칙 검증에 적용하였다.

상향(bottom-up) 방식의 온톨로지 구조와 추론 규칙을 형성시킨 후 이를 의미 검색에 적용해 보고자 하는 연구의 목적에 따라 본 연구에서는, 2015년 7월 6일 현재 STNet에서 검수 작업이 완료된 전체 데이터의 객체값으로 정의된 개념속성 Y용어가 입력된 X용어의 클래스를, Y용어가 많이 입력된 상위 10개 순으로 <표 1>과 같이 정리하였다. 그리고 가장 많은 개념

<표 1> 개념속성 Y용어가 많이 입력된 X용어 클래스 순위

순위	X용어 클래스	클래스별 개념속성(object)	Y용어 입력수
1	y01-01 실존인물	시대, 직업, 출생지, 국적, 이론/사상, 소속, 저작, 출처, 일반주제어	25,681
2	d01-01 이론/사상	관련자, 해당국가, 주창자, 대상주제, 시대, 반대이론	4,542
3	y02-02 예술작품명	장르/유형, 저작자, 시대, 이론/사상, 원서명, 수상명, 일반주제어	2,785
4	y02-01 문헌명	장르/유형, 저작자, 시대, 이론/사상, 원서명, 수상명, 일반주제어, 판본, 시대	2,581
5	y06-01 조직/단체명	소재지, 시대, 설립자, 설립목적, 근거법/제도, 이전기관명, 이후기관명	2,554
6	x01-03 도시/구/동명	위치, 원지명, 일반주제어	2,464
7	b01-01 행위/활동	행위자, 행위대상자	2,143
8	y03-01 사건명/회담명	주관자, 관련자, 장소, 시대, 대상자, 원인, 결과	1,881
9	d01-04 개념(정의)	주창자, 관련자, 해당국가, 시대, 대상주제, 반대이론	1,755
10	Y04 기념물명(문화재)	제작시대, 소재지, 유형/장르, 구분, 제작자, 관리자(처)	1,658

〈표 2〉 ‘y01-01 실존인물’과 연결된 클래스 및 용어 수

클래스			용어 수
코드	국문 명	영문 명	
a01-06-01	인간(직업)	Human(Job)	691
d01-01	이론/사상	Theory/Ideology	1,966
x01-02	국가명	Country Name	271
x01-03	도시/구/동명	Name of City(Town/State/Street/Avenue)	2,397
x02-01	시대	x02-01 Era	121
y01-01	실존인물	y01-01 Real Person	4,574
y02-01	문헌명	y02-01 Title of Literature	1,017
y02-02	예술작품명	y02-02 Title of Creative Works	928
y06-01	조직/단체명	y06-01 Organization Name/Group Name	1,040
합계			13,005

속성 Y용어가 입력된 X용어 클래스 ‘y01-01 실존인물’을 온톨로지 구조의 가장 기본이 되는 중심 클래스로 삼았다.

따라서 본 연구의 실험대상 용어는 온톨로지 구조의 중심 클래스로 설정한 ‘y01-01 실존인물’ 클래스 및 이 클래스의 속성과 연결되어 있는 9개 클래스에 해당하는 용어 총 13,005개이다(〈표 2〉 참조).

### 3. 온톨로지 구조 생성

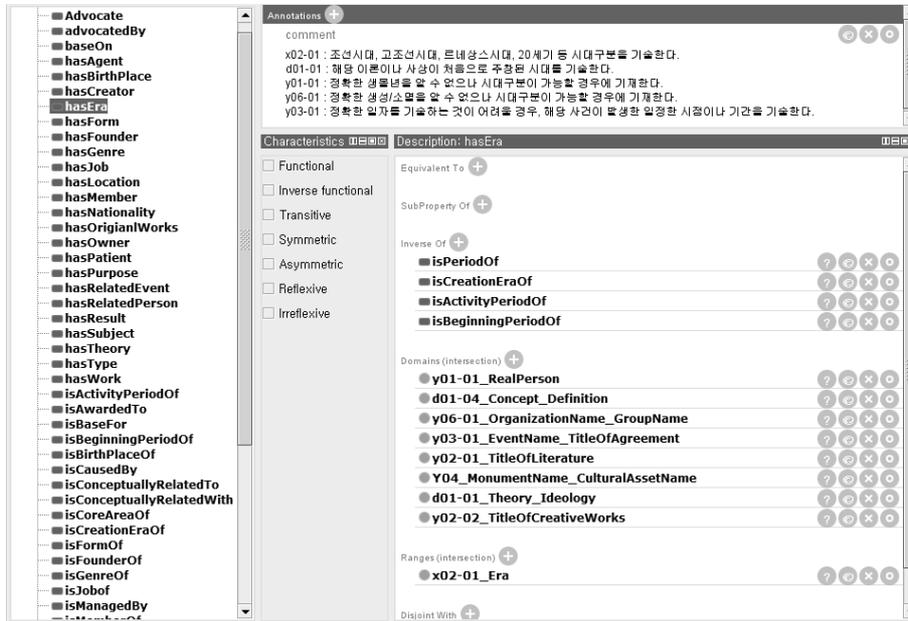
#### 3.1 온톨로지 클래스 생성

온톨로지 클래스는 STNet의 텍사노미 구조를 참조하여 구성하였다. 실험을 위해 ‘y01-01 실존인물’ 및 이 클래스에 속하는 속성값과 연결되는 9개 클래스를 1단계 연결 관계로 설정하고, 이후 9개 클래스의 개념 속성 중 추론 검증을 위해 의미적으로 반드시 포함되어야 할 2단

계 연결고리 생성을 위해 연관 클래스들을 추가하였다(부록 1 참조). 또한 클래스 상관관계를 고려하여 의미적으로 연관이 없거나 동일한 속성을 공유하지 않는 클래스에 대해서는 Disjoint를 통해 ‘서로소’를 구성하였다.

#### 3.2 OWL Properties 설정

STNet 온톨로지를 OWL-DL의 형태로 구성하였으며, ‘y01-01 실존인물’을 대상으로 관련 클래스들의 속성을 참조하여 ObjectType Property 56개와 DataType Property 15개를 정의하였다(〈그림 3〉 참조). ObjectType Property의 경우 역관계(InverseOf)를 설정하고 재귀호출구조(Reflexive)를 생성하였으며 클래스별 속성구조에 따라 Domain과 Range를 설정하였다. DataType Property의 경우 STNet의 텍스트와 코드 속성값을 참조하여 각 유형에 따라 String, DateTime, Integer의 Range를 부여하였다.



〈그림 3〉 ObjectType Property 설정 사례(예: hasEra)

### 3.3 온톨로지 구조 검증

STNet 온톨로지가 OWL-DL의 형태로 구성되어 있으므로 Pellet 추론기를 통해 DL(Description Language, [부록 2] 참조)의 표현력 표현 규칙 *ALCHIF(D)*를 만족하는 클래스와 속성이 생성된 온톨로지 구조의 오류를 검증하였다. 검증 결과 〈그림 4〉와 같이 오류가 없는 것으로 확인되

었다.

Pellet 추론기는 온톨로지 구조나 데이터의 문제가 발생할 경우 경고와 함께 수정대상 객체의 URI와 Report를 출력해주는 기능이 있으며 클래스 추론, 객체속성 추론, 데이터속성 추론, 예시 추론에 대한 구조적 오류를 검증할 수 있다(〈표 3〉 참조).

```
Setting active ontology to OntologyID(OntologyIRI<<http://www.semanticweb.org/administrator/ontologies/2015/6/untitled-ontology-15>>)
Rebuilding entity indices...
... rebuilt in 15 ms
... active ontology changed
Using OWL API version 3.4.2
Rebuilding entity indices...
... rebuilt in 0 ms
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  data property assertions
  same individuals
Pellet classified in 203ms
```

〈그림 4〉 Pellet 추론기에 의한 검증 결과

〈표 3〉 Pellet 추론기의 검증 대상 추론 유형과 검증 항목

검증 대상 추론 유형	검증 항목
Displayed Class Inferences	Unsatisfiability
	Equivalent classes
	Superclasses
	Class Member
	Disjoint Classes
Displayed Object Property Inferences	Unsatisfiability
	Domain
	Range
	Equivalent Properties
	Super Properties
Displayed Data Property Inferences	Inverses
	Domain
	Equivalent Properties
Displayed Individual Inferences	Super Properties
	Types
	Object Property Assertion
	Data Property Assertion
	Same Individual

## 4. RDF 온톨로지 변환

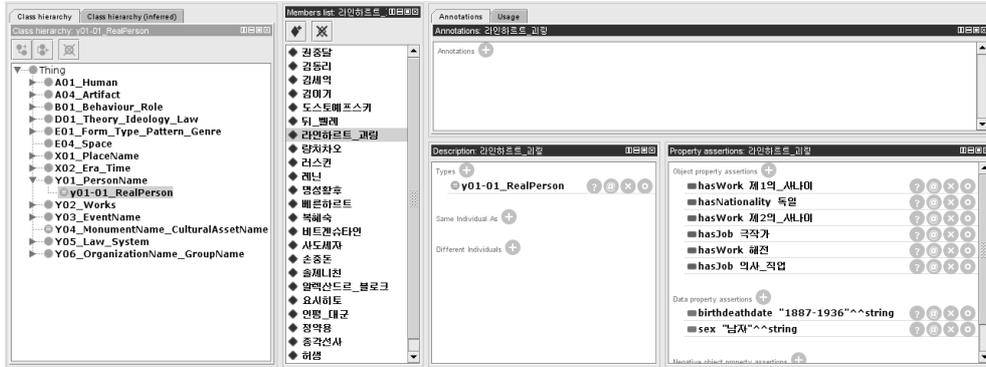
### 4.1 샘플데이터 생성 및 검증

온톨로지 구조 검증을 수행한 온톨로지 스키마에 실제 STNet의 인스턴스 데이터 중 ‘y01-01 실존인물’ 데이터와 ‘시대, 직업, 출생지, 국적, 저작’의 속성으로 입력된 데이터 중 사전 테스트를 위해 샘플 데이터를 추출하였다. 추출된 데이터를 토대로 테스트 인스턴스를 생성하였으며, Pellet 추론기를 통해 데이터의 오류 여부를 검증하였다(〈그림 5〉 참조).

추론기를 통한 검증 결과 동일한 상위 클래스를 지니는 ‘y02-01 문헌명’과 ‘y02-02 예술작품명’ 클래스가 동일한 속성과 구조를 지니고 있어 두 클래스에 속한 데이터가 상호 호환될

수 있는 문제점이 발견되었다(〈그림 6〉 참조).

개념적으로 문헌과 예술작품이 중복될 수 없으므로 상호간의 disjoint를 생성하여 구조적 오류를 수정하였으며, 이러한 과정을 거쳐 ‘y01-01 실존인물’ 클래스의 공리셋을 구축하였다(〈그림 7〉 참조). 〈그림 7〉의 상단 부분은 주어부인 ‘y01-01 실존인물(RealPerson)’ 클래스가 개념속성인 국적(hasNationality), 저작(hasWork), 출생지(hasBirthPlace), 시대(hasEra), 직업(hasJob), 소속(isMemberOf), 이론/사상(hasTheory)을 술어부로 가지며, 각 속성 값에 연결되는 목적어부에 해당하는 대표 클래스들을 설정한 것이다. 하단 부분은 상단 부분에서 설정한 역관계 형태에 관한 것이다. 실존인물의 속성값으로 연결되는 각 클래스를 주어부로 할 때 실존인물 클래스가 목적어부로 연결되는 것을 보여준다.



〈그림 5〉 온톨로지 인스턴스 예시(예: 라인하르트 괴링)



〈그림 6〉 'y02-01 문헌명' 과 'y02-02 예술작품명'의 추론결과 비교



〈그림 7〉 제약조건이 포함된 'y01-01 실존인물' 클래스의 공리 예시

유형구분	예시화면																																												
RDB 데이터	<table border="1"> <thead> <tr> <th>CLASS_NAME</th> <th>TERM_NAME</th> <th>PREDICATE_NAME</th> <th>VAL</th> </tr> </thead> <tbody> <tr><td>1 y01-01 실존인물</td><td>라인하르트 괴링</td><td>국적</td><td>독일</td></tr> <tr><td>2 y01-01 실존인물</td><td>라인하르트 괴링</td><td>생몰년_t</td><td>1887-1936</td></tr> <tr><td>3 y01-01 실존인물</td><td>라인하르트 괴링</td><td>성별</td><td>male</td></tr> <tr><td>4 y01-01 실존인물</td><td>라인하르트 괴링</td><td>시대</td><td>서양 근대</td></tr> <tr><td>5 y01-01 실존인물</td><td>라인하르트 괴링</td><td>저작</td><td>제1의 사나이</td></tr> <tr><td>6 y01-01 실존인물</td><td>라인하르트 괴링</td><td>저작</td><td>제2의 사나이</td></tr> <tr><td>7 y01-01 실존인물</td><td>라인하르트 괴링</td><td>저작</td><td>해전</td></tr> <tr><td>8 y01-01 실존인물</td><td>라인하르트 괴링</td><td>직업</td><td>극작가</td></tr> <tr><td>9 y01-01 실존인물</td><td>라인하르트 괴링</td><td>직업</td><td>의사</td></tr> <tr><td>10 y01-01 실존인물</td><td>라인하르트 괴링</td><td>출생지</td><td>헤센</td></tr> </tbody> </table>	CLASS_NAME	TERM_NAME	PREDICATE_NAME	VAL	1 y01-01 실존인물	라인하르트 괴링	국적	독일	2 y01-01 실존인물	라인하르트 괴링	생몰년_t	1887-1936	3 y01-01 실존인물	라인하르트 괴링	성별	male	4 y01-01 실존인물	라인하르트 괴링	시대	서양 근대	5 y01-01 실존인물	라인하르트 괴링	저작	제1의 사나이	6 y01-01 실존인물	라인하르트 괴링	저작	제2의 사나이	7 y01-01 실존인물	라인하르트 괴링	저작	해전	8 y01-01 실존인물	라인하르트 괴링	직업	극작가	9 y01-01 실존인물	라인하르트 괴링	직업	의사	10 y01-01 실존인물	라인하르트 괴링	출생지	헤센
CLASS_NAME	TERM_NAME	PREDICATE_NAME	VAL																																										
1 y01-01 실존인물	라인하르트 괴링	국적	독일																																										
2 y01-01 실존인물	라인하르트 괴링	생몰년_t	1887-1936																																										
3 y01-01 실존인물	라인하르트 괴링	성별	male																																										
4 y01-01 실존인물	라인하르트 괴링	시대	서양 근대																																										
5 y01-01 실존인물	라인하르트 괴링	저작	제1의 사나이																																										
6 y01-01 실존인물	라인하르트 괴링	저작	제2의 사나이																																										
7 y01-01 실존인물	라인하르트 괴링	저작	해전																																										
8 y01-01 실존인물	라인하르트 괴링	직업	극작가																																										
9 y01-01 실존인물	라인하르트 괴링	직업	의사																																										
10 y01-01 실존인물	라인하르트 괴링	출생지	헤센																																										
RDF 온톨로지로 변환된 데이터																																													

〈그림 8〉 RDB데이터의 RDF 온톨로지 변환결과

#### 4.2 전체 실험 데이터의 RDF 온톨로지 변환

D2R서버(<http://d2rq.org/>)를 사용하여 RDB에 저장되어 있는 STNet의 실험 데이터를 RDF 온톨로지로 변환하였다. D2R 서버는 자체의 D2RQ 매핑 언어를 사용하여 응용 프로그램의 특정 데이터베이스 스키마와 RDFS, 또는 OWL 온톨로지 간에 매핑을 할 수 있도록 지원한다. STNet의 RDB에 접속하여 변환하고자 하는 데이터를 정의하고 이에 대한 속성값을 설정한 후 D2RQ 형태로 변환된 스크립트를 통해 STNet의 RDB 데이터를 RDF 데이터로 변환하였다(고영만, 이승준, 송민선 2015; 〈그림 8〉 참조). 또한

D2RQ 매핑언어를 생성한 후 D2R 서버에서 제공하는 d2r-query를 통해 해당 내용을 검증하여 오류를 확인하고 수정하였다.

### 5. 온톨로지 검증 및 평가

#### 5.1 실험 데이터에 대한 추론 제한 규칙 정의

실제 실험 데이터를 위한 일반화된 추론 제한 규칙 정의를 위해 〈표 1〉에서 정리하였던 상위 10개 X용어 클래스를 대상으로 ‘X용어 클래스의 속성값으로 입력된 Y용어’를 ‘구축된

모든 X용어 클래스의 Null 값을 포함한 전체 속성 용어 수로 나누어 “X용어 클래스의 속성별 입력 비율”을 산출하였다. 이때 X용어 클래스의 전체 속성 용어 수는 아무런 값도 입력되지 않는 경우(Null)와 Y용어가 입력된 경우를 모두 합산한 것이다. 계산 식 적용 결과 상위 10위의 X용어 클래스 중 속성값이 가장 많이 입력된 비율은 92.5%, 가장 적게 입력된 비율은 0.04%로 나타났으며 전체 입력 비율의 평균값은 21.8%로 나타났다.

$$\text{X용어 클래스의 속성별 입력 비율} = \frac{\text{X용어 클래스의 속성값으로 입력된 Y용어 수}}{\text{구축된 모든 X용어 클래스의 Null 값을 포함한 전체 속성 용어 수}}$$

또한 각 속성에 대해 Y용어가 입력된 데이터를 대상으로, 각각의 Y용어가 어떤 클래스에

연결되는지의 구성 비율을 계산하였다. 이러한 작업을 수행한 이유는 각 개념속성이 갖는 Y용어 클래스의 종류가 1종에만 국한되지 않기 때문에 가장 많이 입력되는 Y용어 클래스를 중심으로 ‘X용어 클래스 ↔ Y용어 클래스’의 관계로 정리되는 추론 제한 규칙을 도출하기 위한 것이다.

STNet의 실제 구축 데이터를 중심으로 입력된 값이 상대적으로 많은 비율을 계산하여, 상위 10위 X용어 클래스를 중심으로 ‘주어부(X용어 클래스) ↔ 술어부(속성) ↔ 목적어부(Y용어 클래스)’의 형식으로 추론 제한 규칙을 정리하였다. 또한 속성에 40% 이상 데이터가 입력되고 그 중 90% 이상의 입력 비율을 차지하는 Y용어 클래스를 기준으로 삼아 상대적으로 많은 비율이 입력된 경우를 대상으로 추론 제한 규칙을 정의하였다(〈표 4〉 참조).

〈표 4〉 추론 제한 규칙 정의 사례

주어부(X용어 클래스)	술어부(속성)	목적어부(Y용어 클래스)	
y01-01 실존인물	시대	x02-01 시대	
	직업	a01-06-01 인간(직업)	
	출생지		x01-03 도시/구/동명
			x01-02 국가명
	국적	x01-02 국가명	
	이론/사상	d01-01 이론/사상	
	소속	y06-01 조직/단체명	
저작		y02-02 예술작품명	
		y02-01 문헌명	

- 1-1 실존인물 X는 특정 Y 시대에 활동했다.
- 1-2 실존인물 X는 Y 직업을 갖고 있다.
- 1-3 실존인물 X는 Y 도시/구/동(국가)에서 출생했다.
- 1-4 실존인물 X는 Y 국가의 국적을 갖고 있다.
- 1-5 실존인물 X는 Y 이론/사상을 주장했다(이론/사상과 관련되어 있다).
- 1-6 실존인물 X는 Y 조직/단체에 소속되었다.
- 1-7 실존인물 X는 Y 작품을 생산했다(저술했다. 창작했다).

## 5.2 TBox 검증

온톨로지에서의 추론은 명제들을 기반으로 결론을 얻기 위한 일련의 프로세스를 구성하는 것으로, 결정가능성에 따라 Description Logic, Horn Logic, Frame Logic 등이 사용된다. STNet 온톨로지가 OWL-DL의 형태로 구성되어 있으므로 본 연구에서는 OWL-DL 기반 추론에 적합한 Description Logic을 사용하였다.

DL 기반의 OWL 추론은 TBox(Terminologies Box)와 ABox(Assertion Box)로 구분할 수 있다. TBox는 데이터의 스키마 박스라 할 수 있으며, 클래스를 정의하고 ObjectType Property와 DataType Property를 통해 클래스와 클래스 사이에 존재하는 관계 유형을 정의한다. ABox는 데이터의 인스턴스 박스에 해당하며, 정해진 규칙에 따라 TBox에서 정의된 각각의 클래스에 속하는 인스턴스를 생성한다. STNet 데이터베이스는 지금도 구축이 진행 중인 상태이며, 본 연구를 위해 추출한 데이터도 STNet 전체 데이터의 일부분이므로 본 연구에서는 TBox 검

증을 실시하였다.

TBox는 임의의 개념이 주어졌을 때 상하위 관계를 구축하는 분류(classification) 추론, 주어진 임의의 개념에 대한 포함 여부를 결정하는 포함(subsumption) 추론, 개념 정의에 대한 일관성을 체크하는 일관성(consistency) 추론 검증을 통해 클래스 구조에 대한 하위(subclass), 동기(sibling), 서로소(disjointness) 등의 공리를 검증한다. 본 연구에서 생성한 ‘y01-01 실존인물’의 추론 제한규칙(〈표 4〉 참조)을 포함하는 서술 논리(〈그림 9〉 참조)에 대하여, FaCT++와 Pellet 추론기에 의한 TBox 검증을 실시한 결과 모두 참으로 나타났다(〈그림 10〉 참조).

## 5.3 SPARQL 질의에 의한 평가

본 연구에서 도출한 추론 규칙을 토대로 Protege에서 온톨로지 구조를 정리하고, 단순 키워드 검색만으로는 수행하기 힘든 검색 시나리오 여섯 개를 〈표 5〉와 같이 설정하여 SPARQL 쿼리([부록 3] 참조)의 결과 값을 추출하였다.

```

y01-01_RealPerson ∩ Y01_PersonName
    ⊆ ∃hasNationality.x01-02_CountryName
    ⊆ ∃hasBirthDate
    ⊆ ∃hasSex
    ⊆ isMemberOf.y06-01_OrganizationName_GroupName
    ⊆ hasEra.x02-01_Era
    ⊆ hasTheory.d01-01_Theory_Ideology
    ⊆ hasWorks.{y02-01_TitleOfLiterature, y02-02_TitleOfWorks}
    ⊆ hasJob.a01-06-01_Human_Job
    ⊆ ∃hasBirthplace.x01-03_NameOfCity_Town_State_Street_Avenue
    ⊆ ¬ {D01_Theory_Ideology_Law, Y02_Works, X01PlaceName,
        Y06_OrganizationName_GroupName, Y03_EventName,
        Y04MonumentName_CulturalAssetName}
    
```

〈그림 9〉 ‘y01-01 실존인물’에 대한 서술논리

```

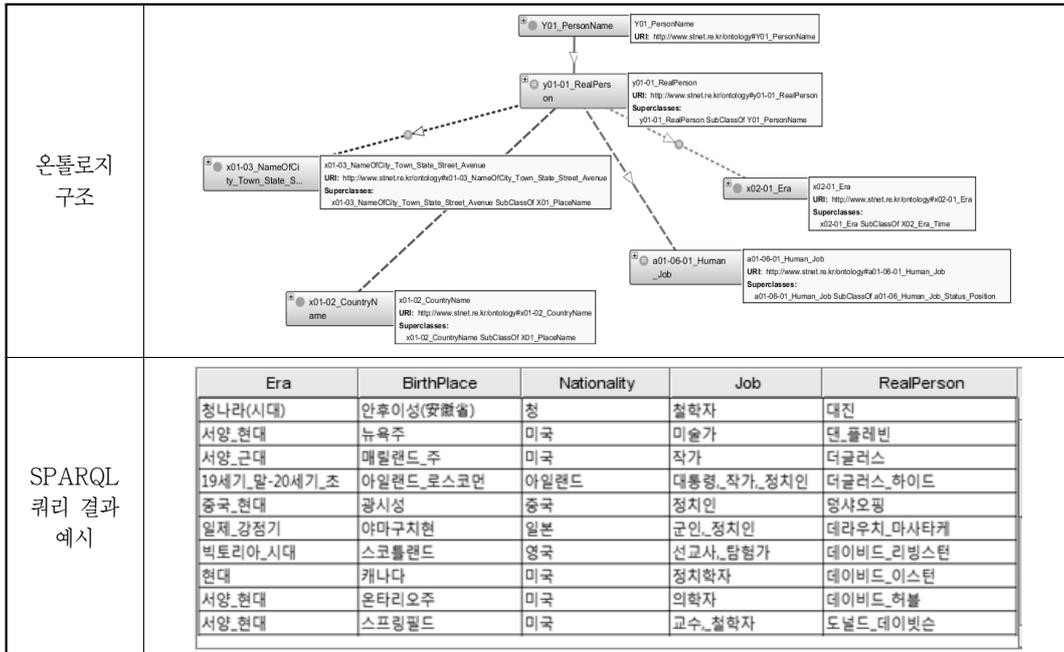
FaCT++.Kernel: Reasoner for the SROIQ(D) Description Logic, 64-bit
Copyright (C) Dmitry Tsarkov, 2002-2013. Version 1.6.2 (19 February 2013)
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  data property assertions
  same individuals
FaCT++ classified in 2389ms
Initializing the reasoner by performing the following steps:
  class hierarchy
  object property hierarchy
  data property hierarchy
  class assertions
  object property assertions
  data property assertions
  same individuals
Pellet classified in 6381ms
    
```

〈그림 10〉 FaCT++ 와 Pellet 추론기를 사용한 TBox 검증 결과

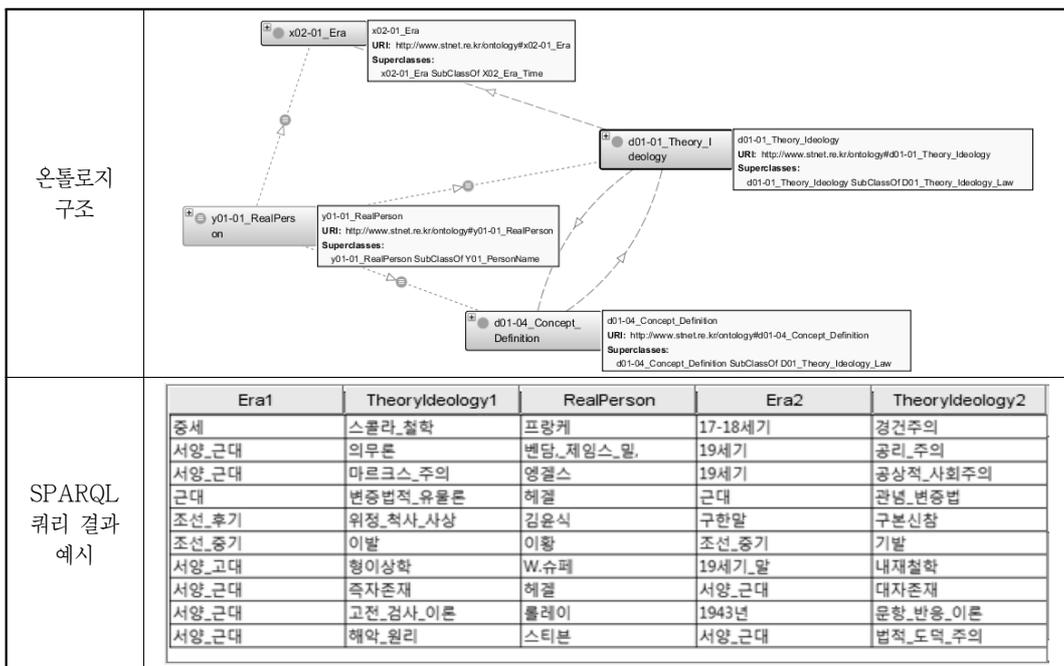
〈표 5〉 의미 검색 시나리오

번호	검색 시나리오
1	[시대 ee]에 [도시/구/동명 또는 국가 bp]에서 출생하여 [직업 jj]을 가진 [실존인물 pp]
2	[시대2 e2]의 [이론/사상 ti]에 반대하여 [실존인물 pp]은 [시대3 e3]에 [이론/사상 ti]을 주창하였다.
3	[실존인물 pp]은 [실존인물2 p2]가 설립한 [조직/단체 og]에 소속되어 활동하였다.
4	[실존인물 pp]은 [시대4 e4]에 [문헌 l]과 [문헌2 l2] 등을 저술하였다.
5	[실존인물 pp]은 [시대5 e5]에 [국가2 p2]에서 발생한 [사건명/회담명 ea]을 주도하였다.
6	[사건명/회담명2 ea2] 때문에 발발한 [사건명/회담명 ea]의 관련자로 [실존인물2 p2]와 [실존인물3 p3]이 있다.

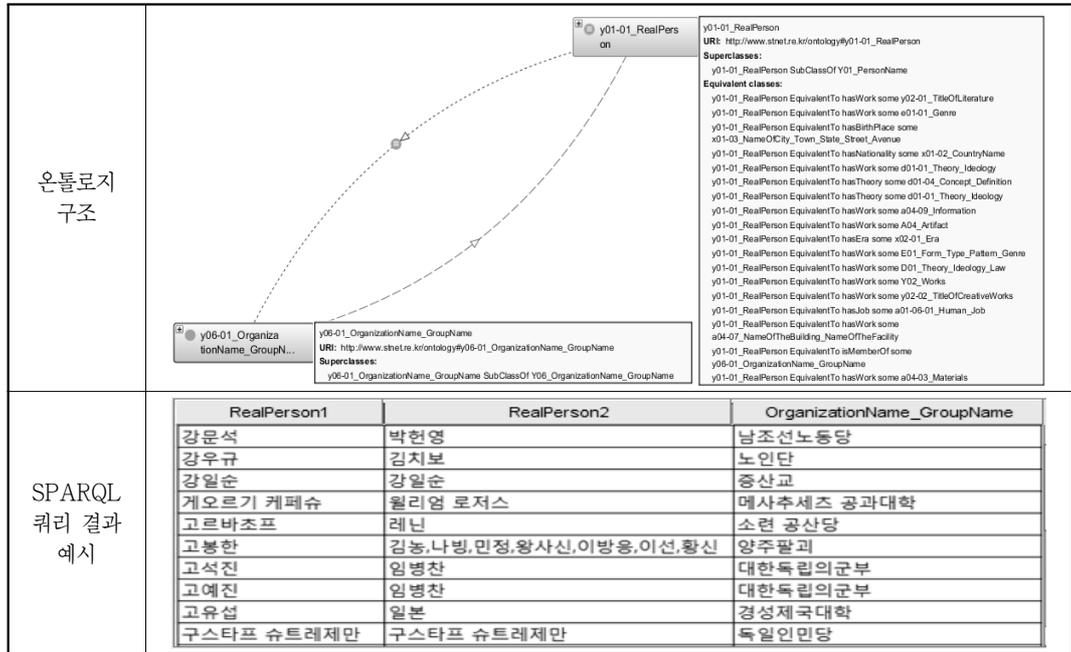
- 시나리오 1: “[시대 ee]에 [도시/구/동명 또는 국가 bp]에서 출생하여 [직업 jj]을 가진 [실존인물 pp]”에 대한 온톨로지 구조와 SPARQL 쿼리 결과는 〈그림 11〉과 같다.
- 시나리오 2: “[시대2 e2]의 [이론/사상 ti]에 반대하여 [실존인물 pp]은 [시대3 e3]에 [이론/사상 ti]을 주창하였다”에 대한 온톨로지 구조와 SPARQL 쿼리 결과는 〈그림 12〉와 같다.
- 시나리오 3: “[실존인물 pp]은 [실존인물2 p2]가 설립한 [조직/단체 og]에 소속되어 활동하였다”에 대한 온톨로지 구조와 SPARQL 쿼리 결과는 〈그림 13〉과 같다.
- 시나리오 4: “[실존인물 pp]은 [시대4 e4]에 [문헌 l]과 [문헌2 l2] 등을 저술하였다.”에 대한 온톨로지 구조와 SPARQL 쿼리 결과는 〈그림 14〉와 같다.
- 시나리오 5: “[실존인물 pp]은 [시대5 e5]에 [국가2 bp2]에서 발생한 [사건명/회담명 ea]을 주도하였다.”에 대한 온톨로지 구조와 SPARQL 쿼리 결과는 〈그림 15〉와 같다.
- 시나리오 6: “[사건명/회담명2 ea2] 때문에 발발한 [사건명/회담명 ea]의 관련자로 [실존인물2 p2]와 [실존인물3 p3]이 있다.”에 대한 온톨로지 구조와 SPARQL 쿼리 결과는 〈그림 16〉과 같다.



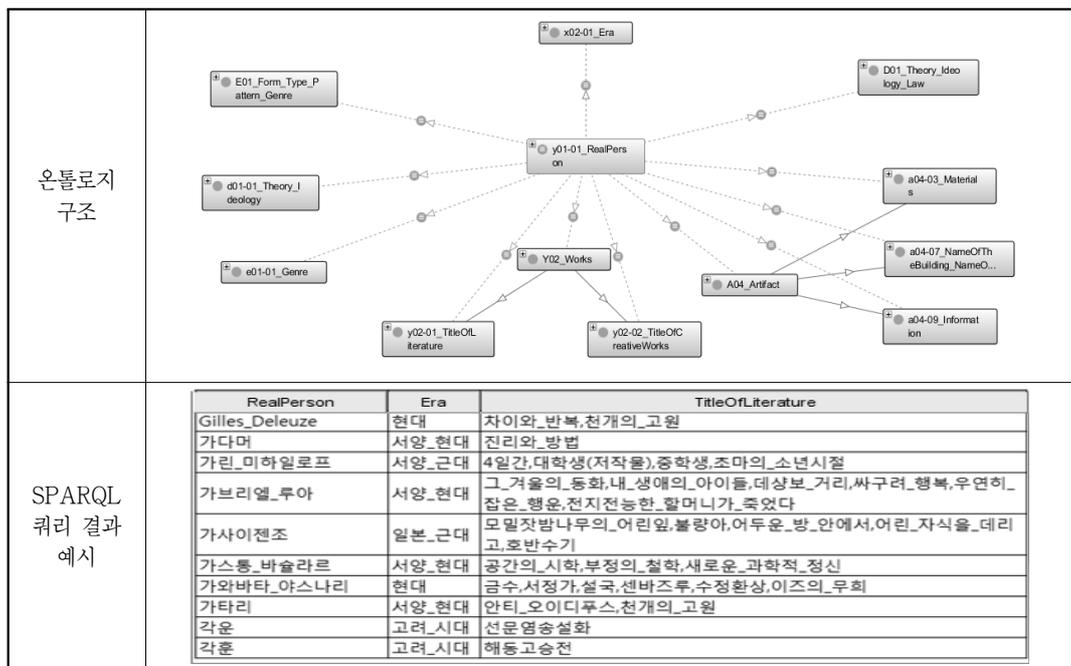
<그림 11> 시나리오 1의 온톨로지 구조와 SPARQL 쿼리 결과 예시



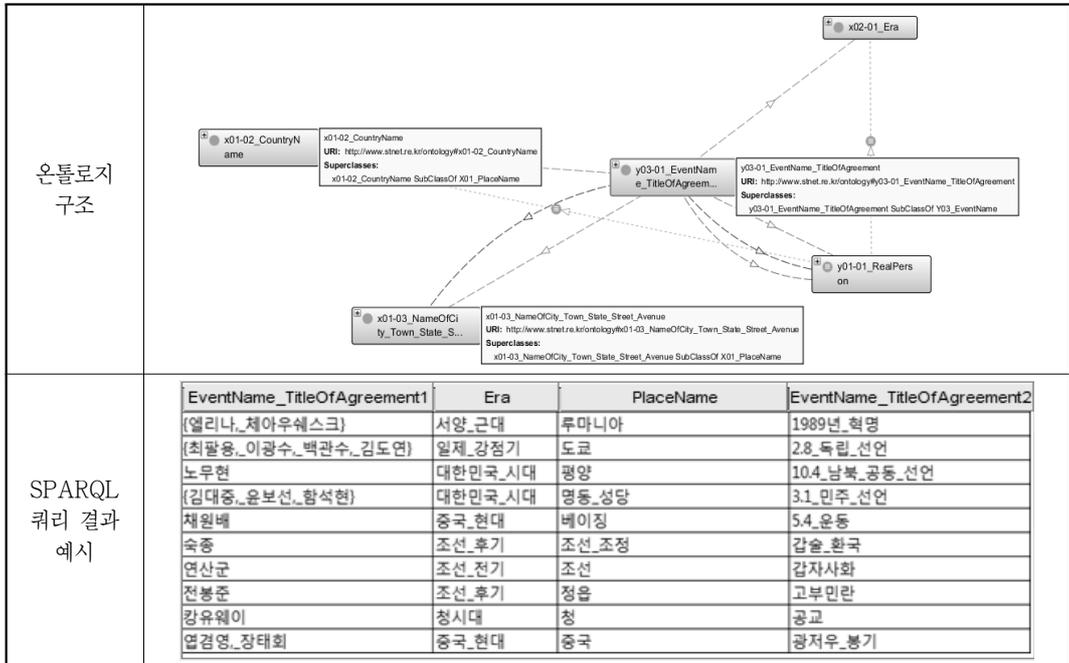
<그림 12> 시나리오 2의 온톨로지 구조와 SPARQL 쿼리 결과 예시



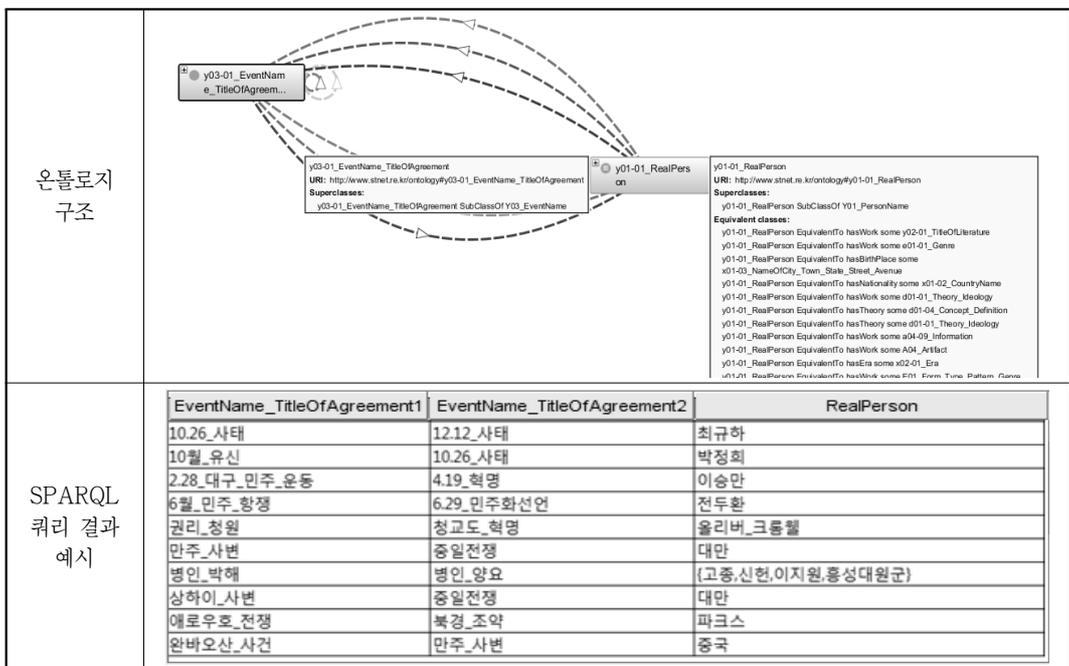
<그림 13> 시나리오 3의 온톨로지 구조와 SPARQL 쿼리 결과 예시



<그림 14> 시나리오 4의 온톨로지 구조와 SPARQL 쿼리 결과 예시



<그림 15> 시나리오 5의 온톨로지 구조와 SPARQL 쿼리 결과 예시



<그림 16> 시나리오 6의 온톨로지 구조와 SPARQL 쿼리 결과 예시

## 6. 결론

본 연구의 목적은 KCI의 인문학, 사회과학, 예술체육 분야 학술 논문에 수록된 키워드를 대상으로 구축된 RDB 기반의 구조적 학술용어사전 “STNet”을 의미적 연관 검색에 활용이 가능한 지식베이스로 발전시키기 위해, RDB 구조에서 추론 가능한 규칙들을 도출하고 RDF로 변환하여 온톨로지를 구축한 후 해당 구조를 실제 STNet 데이터 검색에 적용함으로써 의미 검색의 유용성을 평가하는 것이다. 이를 위해 먼저 STNet의 ‘y01-01 실존인물’ 클래스 및 해당 클래스와 연관된 클래스들을 대상으로 Protege를 이용해 온톨로지 구조를 생성하고, STNet 시스템에 입력된 관계 구조를 토대로 각각의 연관 속성들을 정의하였다. Pellet 추론기를 사용해 구성된 온톨로지 구조의 오류를 검증하였으며, STNet에 입력된 실제 데이터를 가지고 실험을 위한 테스트 데이터를 구축하였다. 그리고 의미적 연관 검색의 효용성을 평가해 보기 위해 STNet의 ‘y01-01 실존인물’ 클래스에서 속성으로 가장 많은 입력값들을 가진 상위 9개 클래스 및 해당 클래스들과 연결된 연관 클래스들을 분석하여 추론 제한 규칙과 검증에 필요한 시나리오를 설정한 후, TBox 검증과 SPARQL 질의에 의한 결과 값을 평가하였다. TBox 검증 결과 본 연구에서 생성한 ‘y01-01 실존인물’에 대한 추론 규칙의 서술 논리가 모두 참으로 나타났다. 기존 단순 키워드 검색만으로는 수행하기 힘든 검색 시나리오에 대한 SPARQL 질의에 대해서는 의미적으로 연관되는 값들이 조합된 검색 결과를 보여주는 것으로 평가되었다.

본 연구의 차별성은 의미 검색을 위한 추론

시스템에 관한 다른 연구들이 대부분 하향식 방법론을 취하고 있는 것과 달리 이미 구축되어 있는 RDB 기반 시스템인 STNet 구조를 바탕으로 온톨로지를 구성하고 추론 규칙을 형성하여 의미 검색에 적용하는 상향식 방법론을 제안했다는 점이다. 따라서 본 연구에서 제안한 방법론은 기존에 구축되어 있는 다양한 분야의 RDB 시스템들에 대해 의미 있는 추론 기능을 탑재할 수 있게 하는 하나의 이론적 토대가 될 수 있을 것이다. 다만 본 연구는 3단계 계층 구조로 나열된 150여개의 STNet 전체 텍사노미 클래스 중 입력 데이터 값이 가장 많이 몰려 있는 ‘y01-01 실존인물’과 연결된 주요 클래스들만을 대상으로 실험을 진행했기 때문에, 전체 STNet 구조에 대한 온톨로지를 구축하여 추론 규칙을 도출한 것은 아니라는 제한점을 가진다.

본 연구는 RDB 기반의 구조적학술용어사전 STNet을 의미 검색이 가능한 지식베이스로 발전시키는 과정에 있어서 ‘용어선정에 관한 연구(고영만 외 2013)’, ‘텍사노미 구축에 관한 연구(고영만, 김비연, 민혜령 2014)’, ‘관계유형에 관한 연구(고영만, 송민선, 이승준 2015)’, ‘RDF 온톨로지 변환 방식 연구(고영만, 이승준, 송민선 2015)’에 이어지는 다섯 번째 단계의 연구이다. 따라서 첫 번째 연구부터 본 연구까지의 연구 내용과 결과는 RDB 기반의 학술정보 데이터베이스를 의미검색이 가능한 지식베이스로 발전시키는 프로세스와 기술적 고려에 대한 지침이 될 수 있을 것이다. 또한 STNet에 구축되어 있는 수많은 학술용어 간 의미 구조는 특정 학술정보 시스템의 도메인과 특성에 맞추어 응용해서 사용할 수 있으며, 해당 시스템에 적합한

의미 추론 시스템을 개발하는 데 활용될 수 있을 것이다. 특히 STNet은 KCI의 저자키워드를 토대로 구축한 인문, 사회 및 체육, 예술 분야의 구조적 학술용어사전이므로, 본 연구 결과를 KCI

에 적용시킬 경우 추론규칙에 의한 의미검색 결과값으로 제시된 용어들이 포함된 논문들을 검색하는 시스템 구축에 직접 활용될 수 있을 것이다.

## 참 고 문 헌

- [1] 강현민. 2010. RDF/OWL의 객체속성을 이용한 관계온톨로지 시스템 구축과 활용에 관한 연구. 『정보관리학회지』, 27(4): 219-237.
- [2] 고영만. 2006. 시소러스 기반 온톨로지에 관한 연구. 『정보관리학회지』, 5: 5-22.
- [3] 고영만, 송인석. 2011. 연구문헌의 지식구조를 반영하는 의미기반의 지식조직체계에 관한 연구. 『정보관리학회지』, 28(1): 145-170.
- [4] 고영만 외. 2013. 인문학 및 사회과학 분야 국내 학술논문의 저자키워드 출현빈도와 피인용횟수의 상관관계 연구. 『정보관리학회지』, 30(2): 227-243.
- [5] 고영만, 김비연, 민혜령. 2014. 한국학술지인용색인(KCI)의 인문학, 사회과학, 예술체육 분야 저자 키워드의 의미적, 형태적 분석에 의한 개념범주 텍사노미 연구. 『한국문헌정보학회지』, 48(4): 297-322.
- [6] 고영만, 송민선, 이승준. 2015. 한국학술지인용색인(KCI)의 인문학, 사회과학, 예술체육 분야 저자 키워드의 의미적 관계 유형 최적화 연구. 『한국문헌정보학회지』, 49(1): 45-67.
- [7] 고영만, 이승준, 송민선. 2015. 관계형 데이터베이스 기반 구조적학술용어사전(STNet)의 RDF 온톨로지 변환 방식 연구. 『정보관리학회지』, 32(2): 131-151.
- [8] 김상균 외. 2008. 한국한의학연구원 소셜 네트워크 온톨로지 구축. 『한국콘텐츠학회논문지』, 9(12): 485-495.
- [9] 김수경, 안기홍. 2007. 지능형 이미지 검색 시스템을 위한 추론 기반의 웹 온톨로지 구축. 『정보관리학회지』, 24(3): 119-147.
- [10] 김재훈, 박석. 2011. OWLJessKB 추론 규칙을 바탕으로 한 subClassOf 추론에서의 그래프 레이블링 효율성 재평가. 『정보과학회논문지: 데이터베이스』, 38(4): 238-247.
- [11] 김진성. 2003. 데이터마이닝과 사례기반추론 기법에 기반한 인터넷 구매지원 시스템 구축에 관한 연구. 『한국경영과학회지』, 28(3): 135-148.
- [12] 김현희, 안태경. 2003. 온톨로지를 이용한 인터넷웹 검색에 관한 실험적 연구. 『정보관리학회지』, 20(1): 417-455.

- [13] 박길식, 박성철, 김준태. 2012. 온톨로지 기반 Jess 추론 규칙을 이용한 자동차 정비 이-트레이닝 시스템에 대한 연구. 『한국컴퓨터종합학술대회 논문집』, 39(1(B)): 417-419.
- [14] 송우중, 김유성. 2008. Wine 온톨로지와 추론엔진 Jena를 활용한 의미추론 기반의 Wine정보검색 시스템. [online] [cited 2015. 7. 10.]  
<[http://ice.inha.ac.kr/xe/?module=file&act=procFileDownload&file\\_srl=66002&sid=e70eb0fdbf226e8715f5027200a941f2](http://ice.inha.ac.kr/xe/?module=file&act=procFileDownload&file_srl=66002&sid=e70eb0fdbf226e8715f5027200a941f2)>
- [15] 송인석. 2008. 연구·학술정보 효율적 검색을 위한 온톨로지 기반의 주제 색인어 구조화 방안 연구. 『정보관리연구』, 39(4): 121-154.
- [16] 이태영. 2009. 시맨틱 웹 환경에서 적합한 문장을 제공하는 이야기 쓰기 도우미에 관한 연구. 『정보관리학회지』, 26(4): 7-33.
- [17] 장창복, 김만재, 최의인. 2012. 상황 인식 추천 서비스를 위한 온톨로지 이용 OWL 모델링. 『한국인터넷방송통신학회 논문지』, 12(1): 265-273.
- [18] 정상원. 2009. 『과학기술 컨퍼런스 정보의 반자동 수집 및 학술적 의미추론 시스템 연구개발』. 대전: 한국연구재단, H00001(100917).
- [19] 최정화, 박영택. 2010. 온톨로지 Open World 추론과 규칙 Closed World 추론의 통합. 『정보과학회논문지: 소프트웨어 및 응용』, 37(4): 282-295.
- [20] 허정환 외. 2008. 의미추론규칙을 이용한 온톨로지 기반의 전자우편 자동 분류 시스템. 『한국정보과학회 2008 가을 학술발표논문집』, 35(2(C)): 234-239.
- [21] Horridge, M., Parsia, B. and Sattler, U. 2009. "Explaining Inconsistencies in OWL Ontologies." *Lecture Notes in Computer Science*, 5785: 124-137.
- [22] Sirin, E. et al. 2007. "Pellet: A Practical OWL-DL Reasoner." *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 5(2): 51-53.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Kang, Hyen-Min. 2010. "A Study on Implementation and Applying Relationship Ontology System Using RDF/OWL Object Property." *Journal of the Korean Society for Information Management*, 27(4): 219-237.
- [2] Ko, Young Man. 2006. "A Study on the Ontology based on Thesaurus." *Information Management*, 5: 5-22.
- [3] Ko, Young Man and Song, Inseok. 2011. "A Study on the Knowledge Organizing System of Research Papers Based on Semantic Relation of the Knowledge Structure." *Journal of*

- the Korean Society for Information Management*, 28(1): 145-170.
- [4] Ko, Young Man et al. 2013. “A Study on the Correlation between the Appearance Frequency of Author Keyword and the Number of Citation in the Humanities and Social Science Journal Articles of the Korea Citation Index (KCI).” *Journal of the Korean Society for Information Management*, 30(2): 227-243.
- [5] Ko, Young Man, Kim, Bee-Yeon and Min, Hye-Ryoung. 2014. “A Study on a Conceptual Taxonomy of Author Keywords of Humanities, Social Sciences, and Art and Sport in the Korea Citation Index (KCI) by Analysis of its Meaning and Lexical Morpheme.” *Journal of the Korean Society for Library and Information Science*, 48(4): 297-322.
- [6] Ko, Young Man, Song, Min-Sun and Lee, Seung-Jun. 2015. “A Study on the Optimization of Semantic Relation of Author Keywords in Humanities, Social Sciences, and Art and Sport of the Korea Citation Index (KCI).” *Journal of the Korean Society for Library and Information Science*, 49(1): 45-67.
- [7] Ko, Young Man, Lee, Seung-Jun and Song, Min-Sun. 2015. “A Study on Conversion Methods for Generating RDF Ontology from Structural Terminology Net (STNet) based on RDB.” *Journal of the Korean Society for Information Management*, 32(2): 131-151.
- [8] Kim, Sang-Kyun et al. 2008. “Construction of Social Network Ontology in Korea Institute of Oriental Medicine.” *Journal of the Korea Contents Association*, 9(12): 485-495.
- [9] Kim, Su-Kyung and Ahn, Kee-Hong. 2007. “Web Ontology Modeling Based on Description Logic and SWRL.” *Journal of the Korean Society for Information Management*, 24(3): 119-147.
- [10] Kim, Jaehoon and Park, Seog. 2011. “An Efficiency Revaluation of Graph Labeling for subClassOf Inference based on OWLJessKB Inference Rules.” *Journal of Korea Institute of Information Scientists and Engineers: Database*, 38(4): 238-247.
- [11] Kim, Jin Sung. 2003. “A Study on the Development of Internet Purchase Support Systems Based on Data Mining and Case-Based Reasoning.” *Journal of the Korean Operations Research and Management Science Society*, 28(3): 135-148.
- [12] Kim, Hyun-Hee and Ahn, Tae-Kyung. 2003. “An Experimental Study on the Internet Web Retrieval Using Ontologies.” *Journal of the Korean Society for Information Management*, 20(1): 417-455.
- [13] Park, GilSik, Park, SungChul and Kim, JunTae. 2012. “The Study about Ontology based e-training System for Automobile Maintenance Education using Jess Inference Rule.” *Proceedings of Korea Computer Congress 2012*, 39(1(B)): 417-419.

- [14] Song, Woojong and Kim, Yusung. 2008. "A Reasoning based Wine Information Searching System with Wine Ontology and Reasoning Engine Jena." [online] [cited 2015. 7. 10.] <[http://ice.inha.ac.kr/xe/?module=file&act=procFileDownload&file\\_srl=66002&sid=e70eb0fdbf226e8715f5027200a941f2](http://ice.inha.ac.kr/xe/?module=file&act=procFileDownload&file_srl=66002&sid=e70eb0fdbf226e8715f5027200a941f2)>
- [15] Song, In-Seok. 2008. "A Study on Ontology-based Keywords Structuring for Efficient Information Retrieval." *Journal of Information Management*, 39(4): 121-154.
- [16] Lee, Tae-young. 2009. "A Study on Retrieval Model Providing Relevant Sentences in Storytelling on Semantic Web." *Journal of the Korean Society for Information Management*, 26(4): 7-33.
- [17] Chang, Changbok, Kim, Manjae and Choi, Euiin. 2012. "OWL Modeling using Ontology for Context Aware Recommendation Service." *The Journal of the Institute of Internet, Broadcasting and Communication*, 12(1): 265-273.
- [18] Jeong, Senator. 2009. *Semi-Automatic crawling of Conference data and Inferring of Scholarly meaing*. Daejeon: National Research Foundation of Korea. H00001(100917).
- [19] Choi, Jung-Hwa and Park, Young-Taek. 2010. "Integration of Ontology Open-World and Rule Closed-World Reasoning." *Journal of Korea Institute of Information Scientists and Engineers: Software and Applications*, 37(4): 282-295.
- [20] Heu, Chung-Hwan et al. 2008. "Ontology-based E-mail Classification System using Semantic Inference Rules." *Proceedings of Korea Information Science Society 2008 Fall Conference*, 35(2(C)): 234-239.

[부록 1] 본 연구에서 구축한 STNet 온톨로지 클래스와 속성

클래스 코드	클래스 명	영문명	속성 타입	속성 명	속성 영문명 (↔inverseOf)
a01-06-01	인간(직업)	Human Job	DataType	직업분류(군)	jobClassifications
a04-03	교재/자료	Materials			-
a04-07	건축물(시설물)	NameOfTheBuilding _NameOfTheFacility			-
a04-09	저작(창작)물/정보	Information			-
b01-02	교육활동	EducationActivity			-
d01-01	이론/사상	Theory_Ideology	ObjectType	관련자	isConceptuallyRelatedTo(↔hasTheory)
			ObjectType	해당국가	hasLocation(↔isCoreAreaOf)
			ObjectType	주창자	advocatedBy(↔Advocate)
			ObjectType	시대	hasEra(↔isBeginningPeriodOf)
			ObjectType	대상주제	hasSubject(↔isSubjectIn)
			ObjectType	반대이론	isOppositionOf(↔isOppositionOf)
d01-04	개념(정의)	Concept_Definition	ObjectType	주창자	advocatedBy(↔Advocate)
			ObjectType	관련자	isConceptuallyRelatedTo(↔hasTheory)
			ObjectType	해당국가	hasLocation(↔isCoreAreaOf)
			ObjectType	시대	hasEra(↔isBeginningPeriodOf)
			ObjectType	대상주제	hasSubject(↔isSubjectIn)
			ObjectType	반대이론	isOppositionOf(↔isOppositionOf)
e01-01	문학장르	Genre			-
x01-02	국가명	CountryName	DataType	지역	area
x01-03	도시/구/동명	NameOfCity_Town_ State_Street_Avenue	ObjectType	위치	hasLocation(↔isPlaceof)
			ObjectType	원지명	PT(↔LT)
			DataType	지역	area
x02-01	시대	Era			-
y01-01	실존인물	RealPerson	ObjectType	시대	hasEra(↔isActivityPeriodOf)
			ObjectType	직업	hasJob(↔isJobOf)
			ObjectType	출생지	hasBirthPlace(↔isBirthPlaceOf)
			ObjectType	국적	hasNationality(↔isNationalityOf)
			ObjectType	이론/사상	hasTheory(↔isConceptuallyRelatedWith)
			ObjectType	소속	isMemberOf(↔hasMember)
			ObjectType	저작	hasWork(↔hasCreator)
y02-01	문헌명	TitleOfLiterature	ObjectType	저작자	hasCreator(↔hasWork)
			ObjectType	시대	hasEra(↔isCreationEraOf)
			ObjectType	장르/유형	hasGenre(↔isGenreOf)
			ObjectType	이론/사상	isConceptuallyRelatedTo(↔ isConceptuallyRelatedTo)
			ObjectType	원서명	hasOriginalWorks(↔isOriginalWorksOf)
			ObjectType	판본	hasForm(↔isFormOf)
			ObjectType	수상명	winAward(↔isAwardedTo)

클래스 코드	클래스 명	영문명	속성 타입	속성 명	속성 영문명 ( $\leftrightarrow$ inverseOf)
y02-02	예술작품명	TitleOfCreativeWorks	ObjectType	저작자	hasCreator( $\leftrightarrow$ hasWork)
			ObjectType	시대	hasEra( $\leftrightarrow$ isCreationEraOf)
			ObjectType	장르/유형	hasGenre( $\leftrightarrow$ isGenreOf)
			ObjectType	이론/사상	isConceptuallyRelatedTo( $\leftrightarrow$ isConceptuallyRelatedTo)
			ObjectType	원서명	hasOriginalWorks( $\leftrightarrow$ isOriginalWorksOf)
			ObjectType	수상명	winAward( $\leftrightarrow$ isAwardedTo)
y03-01	사건명/회담명	EventName_TitleOfAgreement	ObjectType	주관자	isManagedBy( $\leftrightarrow$ Manage)
			ObjectType	관련자	hasRelatedPerson( $\leftrightarrow$ hasRelatedEvent)
			ObjectType	장소	hasLocation( $\leftrightarrow$ isCoreAreaOf)
			ObjectType	시대	hasEra( $\leftrightarrow$ isPeriodOf)
			ObjectType	대상자	hasPatient( $\leftrightarrow$ hasAgent)
			ObjectType	원인	isCausedBy( $\leftrightarrow$ hasResult)
			ObjectType	결과	hasResult( $\leftrightarrow$ isCausedBy)
y03-04	수상명	NameOfTheAward	-		
Y04	기념물명(문화재)	MonumentName_CulturalAssetName	ObjectType	제작시대	hasEra( $\leftrightarrow$ isCreationEraOf)
			ObjectType	소재지	hasLocation( $\leftrightarrow$ isPlaceOf)
			ObjectType	유형/장르	hasGenre( $\leftrightarrow$ isGenreOf)
			ObjectType	소유자(처)	hasOwner( $\leftrightarrow$ isOwnerOf)
			ObjectType	구분	hasType( $\leftrightarrow$ isTypeOf)
			ObjectType	제작자	isProducedBy( $\leftrightarrow$ Produce)
			ObjectType	관리자(처)	isManagedBy( $\leftrightarrow$ Manage)
			DataType	지정호수	designationNumber
			DataType	지정일자	designationDate
			DataType	형태	shape
			DataType	제작(설립)년도	creationYear
			DataType	주소	address
y05-01	법률/법령	Law_Legislation	ObjectType	이전법률명	PT( $\leftrightarrow$ LT)
			ObjectType	이후법률명	LT( $\leftrightarrow$ PT)
y06-01	조직/단체명	OrganizationName_GroupName	ObjectType	소재지	hasLocation( $\leftrightarrow$ isPlaceOf)
			ObjectType	시대	hasEra( $\leftrightarrow$ isActivityPeriodOf)
			ObjectType	설립자	hasFounder( $\leftrightarrow$ isFounderOf)
			ObjectType	설립목적	hasPurpose( $\leftrightarrow$ isPurposeOf)
			ObjectType	근거법/제도	baseOn( $\leftrightarrow$ isBaseFor)
			ObjectType	이전기관명	PT( $\leftrightarrow$ LT)
			ObjectType	이후기관명	LT( $\leftrightarrow$ PT)
			DataType	주소	address

※ 모든 속성 혹은 ObjectType 속성이 설정되지 않고 DataType 속성만 설정된 클래스는 슬어부(Range)로 연결되는 클래스임

## [부록 2] DL(Description Language)의 표현력 표기 규칙

표기 형식	표기 형식별 의미
$\mathcal{AL}$	Attributive language. This is the base language which allows: <ul style="list-style-type: none"> <li>• Atomic negation (negation of concept names that do not appear on the left hand side of axioms)</li> <li>• Concept intersection</li> <li>• Universal restrictions</li> <li>• Limited existential quantification</li> </ul>
$\mathcal{FL}$	Frame based description language, allows: <ul style="list-style-type: none"> <li>• Concept intersection</li> <li>• Universal restrictions</li> <li>• Limited existential quantification</li> <li>• Role restriction</li> </ul>
$\mathcal{EL}$	Allows: <ul style="list-style-type: none"> <li>• Concept intersection</li> <li>• Existential restrictions (of full existential quantification)</li> </ul>
$\mathcal{F}$	Functional properties, a special case of uniqueness quantification
$\mathcal{E}$	Full existential qualification (Existential restrictions that have fillers other than $\top$ )
$\mathcal{U}$	Concept union
$\mathcal{C}$	Complex concept negation
$\mathcal{H}$	Role hierarchy (subproperties - rdfs:subPropertyOf)
$\mathcal{R}$	Limited complex role inclusion axioms; reflexivity and irreflexivity; role disjointness
$\mathcal{O}$	Nominals (Enumerated classes of object value restrictions - owl:oneOf, owl:hasValue)
$\mathcal{I}$	Inverse properties
$\mathcal{N}$	Cardinality restrictions (owl:cardinality, owl:maxCardinality), a special case of counting quantification
$\mathcal{Q}$	Qualified cardinality restrictions (available in OWL 2, cardinality restrictions that have fillers other than $\top$ )
( $\mathcal{D}$ )	Use of datatype properties, data values or data types

※ 출처: 위키백과 ([https://en.wikipedia.org/wiki/Description\\_logic#TBox](https://en.wikipedia.org/wiki/Description_logic#TBox))

### [부록 3] 시나리오에 대한 SPARQL 쿼리

---

시나리오 1: [시대 ee]에 [도시/구/동명 또는 국가 bp]에서 출생하여 [직업 jj]을 가진 [실존인물 pp]

---

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX : <http://www.stnet.re.kr/ontology#>

```
SELECT ?Era ?BirthPlace ?Nationality ?Job ?RealPerson
WHERE {
    ?RealPerson rdf:type: y01-01__RealPerson.
    ?Era rdf:type :x02-01__Era.
    ?BirthPlace rdf:type :x01-03__NameOfCity__Town__State__Street__Avenue.
    ?Nationality rdf:type :x01-02__CountryName.
    ?Job rdf:type :a01-06-01__Human__Job.
    ?RealPerson :hasEra ?Era.
    ?RealPerson :hasBirthPlace ?BirthPlace.
    ?RealPerson :hasNationality ?Nationality.
    ?RealPerson :hasJob ?Job
}
```

---

시나리오 2: [시대2 e2]의 [이론/사상 ti]에 반대하여 [실존인물 pp]은 [시대3 e3]에 [이론/사상 ti]을 주창하였다.

---

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX : <http://www.stnet.re.kr/ontology#>

```
SELECT ?Era1 ?TheoyIdeology1 ?RealPerson ?Era2 ?TheoyIdeology2
WHERE {
    ?RealPerson rdf:type :y01-01__RealPerson.
    ?Era1 rdf:type :x02-01__Era.
    ?TheoyIdeology1 rdf:type :d01-01__Theory__Ideology.
    ?Era2 rdf:type :x02-01__Era.
    ?TheoyIdeology2 rdf:type :d01-01__Theory__Ideology.
    ?RealPerson :hasEra ?Era1.
    ?TheoyIdeology1 :advocatedBy ?RealPerson.
    ?TheoyIdeology1 :isOppositionOf ?TheoyIdeology2.
    ?TheoyIdeology2 :hasEra ?Era2
}
```

---

시나리오 3: [실존인물 pp]은 [실존인물2 p2]가 설립한 [조직/단체 og]에 소속되어 활동하였다.

---

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX : <http://www.stnet.re.kr/ontology#>

```
SELECT ?RealPerson1 ?RealPerson2 ?OrganizationName__GroupName
WHERE {
    ?RealPerson1 rdf:type :y01-01__RealPerson.
    ?RealPerson2 rdf:type :y01-01__RealPerson.
    ?OrganizationName__GroupName rdf:type :y06-01__OrganizationName__GroupName.
    ?RealPerson1 :isMemberOf ?OrganizationName__GroupName.
    ?OrganizationName__GroupName :hasFounder ?RealPerson2
}
```

---

---

시나리오 4: [실존인물 pp]은 [시대4 e4]에 [문헌1 l]과 [문헌2 l2] 등을 저술하였다.

---

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
 PREFIX : <http://www.stnet.re.kr/ontology#>

```
SELECT ?RealPerson ?Era ?TitleOfLiterature
WHERE {
    ?RealPerson rdf:type :y01-01__RealPerson.
    ?Era rdf:type :x02-01__Era.
    ?TitleOfLiterature rdf:type :y02-01__TitleOfLiterature.
    ?RealPerson :hasEra ?Era.
    ?RealPerson :hasWork ?TitleOfLiterature
}
```

---

시나리오 5: [실존인물 pp]은 [시대5 e5]에 [국가2 bp2]에서 발생한 [사건명/회담명 ea]을 주도하였다.

---

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
 PREFIX : <http://www.stnet.re.kr/ontology#>

```
SELECT ?EventName__TitleOfArgreement1 ?Era ?placeName ?EventName__TitleOfArgreement2
WHERE {
    ?EventName__TitleOfArgreement1 rdf:type :y01-01__RealPerson.
    ?Era rdf:type :x02-01__Era.
    ?placeName rdf:type :X01__PlaceName.
    ?EventName__TitleOfArgreement2 rdf:type :y03-01__EventName__TitleOfAgreement.
    ?EventName__TitleOfArgreement2 :isManagedBy ?EventName__TitleOfArgreement1.
    ?EventName__TitleOfArgreement2 :hasEra ?Era.
    ?EventName__TitleOfArgreement2 :hasLocation ?X01__PlaceName
}
```

---

시나리오 6: 사건명/회담명2 ea2] 때문에 발발한 [사건명/회담명 ea]의 관련자로 [실존인물2 p2]와 [실존인물3 p3]이 있다.

---

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
 PREFIX : <http://www.stnet.re.kr/ontology#>

```
SELECT ?EventName__TitleOfArgreement1 ?EventName__TitleOfArgreement2 ?RealPerson
WHERE {
    ?EventName__TitleOfArgreement1 rdf:type :y03-01__EventName__TitleOfAgreement.
    ?EventName__TitleOfArgreement2 rdf:type :y03-01__EventName__TitleOfAgreement.
    ?RealPerson rdf:type :y01-01__RealPerson.
    ?EventName__TitleOfArgreement2 :hasResult ?EventName__TitleOfArgreement1.
    ?EventName__TitleOfArgreement1 :hasRelatedPerson ?RealPerson
}
```

---

