

온톨로지 기반의 주제-객체관계를 이용한 국가 R&D 지식맵 구축*

Development of a National R&D Knowledge Map Using the Subject-Object Relation based on Ontology

양명석 (Myung-Seok Yang)** , 강남규 (Nam-Kyu Kang)***
김운정 (Yun-Jeong Kim)**** , 최광남 (Kwang-Nam Choi)*****
김영국 (Young-Kuk Kim)*****

초 록

최근 효과적인 정보검색을 제공하기 위해 시맨틱 웹을 비롯한 다양한 검색기법들을 사용하고 있다. 이 중에서 효과적인 방법은 온톨로지를 이용한 검색기술을 적용하는 것이라 할 수 있다. 본 논문에서는 국가과학기술지식정보서비스(NTIS)에서 구축한 국가R&D정보를 분석하여 온톨로지를 구축하고, 이용자가 관심있어 하는 주제분야(과제, 인물, 성과, 기관)를 중심으로 온톨로지의 객체관계를 표현하고 정보를 탐색하기 위한 국가R&D지식맵(knowledge map)을 구축하였다. 국가R&D지식맵은 사용자가 선택한 객체를 중심노드로 설정하여, 주제분야를 노드로 표현하고, 객체와 주제분야간의 관계를 분석하여 사용자가 관심 있어 하는 질의를 주제분야의 하위노드로 표현하였다. 사용자가 하위노드의 질의를 선택하면 시스템에서는 선택한 질의를 온톨로지로부터 추론할 수 있는 SPARQL 질의어를 생성하고 추론엔진으로부터 검색결과를 받아 사용자에게 제시하였다.

ABSTRACT

To develop an intelligent search engine to help users retrieve information effectively, various methods, such as Semantic Web, have been used. An effective retrieval method of such methods uses ontology technology. In this paper, we built National R&D ontology after analyzing National R&D Information in NTIS and then implemented National R&D Knowledge Map to represent and retrieve information of the relationship between object and subject (project, human information, organization, research result) in R&D Ontology. In the National R&D Knowledge Map, center-node is the object selected by users, node is subject, subject's sub-node is user's favorite query in National R&D ontology after analyzing the relationship between object and subject. When a user selects sub-node, the system displays the results from inference engine after making query by SPARQL in National R&D ontology.

키워드: 온톨로지, 시맨틱 웹, 국가R&D정보, 국가과학기술지식정보서비스, 지식맵
ontology, semantic web, national R&D information, NTIS, knowledge

* 본 연구는 2012년 국가과학기술지식정보서비스 구축사업에 의한 것임.

** 한국과학기술정보연구원, NTIS센터(msyang@kisti.re.kr)

*** 한국과학기술정보연구원, NTIS센터(ngkang@kisti.re.kr)

**** 한국과학기술정보연구원, NTIS센터(miso@kisti.re.kr)

***** 한국과학기술정보연구원, NTIS센터(knchoi@kisti.re.kr)

***** 충남대학교 컴퓨터공학과 교수(ykim@cnu.ac.kr) (교신저자)

■ 논문접수일자: 2012년 11월 19일 ■ 최초심사일자: 2012년 11월 23일 ■ 게재확정일자: 2012년 12월 14일

■ 정보관리학회지, 29(4), 123-142, 2012. [http://dx.doi.org/10.3743/KOSIM.2012.29.4.123]

1. 서론

1.1 연구동향

최근 시맨틱 웹(Semantic Web)을 기반으로 정보를 제공하는 서비스들이 많이 연구되고 있다. 기존 서비스들은 데이터를 데이터베이스에 축적하고, 검색엔진을 활용하여 검색결과를 제공하는 방식에서 데이터 간 혹은 정보 간의 의미관계를 파악하고 분석하여 이를 지식맵(Knowledge Map)이나 온톨로지(Ontology) 등으로 표현하여 사용자의 질의 의도를 파악하고 검색결과에 의미를 더하여 제공하는 의미기반의 검색서비스를 제공하는 방식으로 진화해 나가고 있다.

키워드 기반의 검색서비스는 기존에 구축된 데이터를 데이터베이스로부터 추출하여 색인 데이터를 만들고 색인된 결과 값을 별도의 장치에 저장한 후, 사용자로부터 키워드를 입력받아 검색엔진을 통해 색인된 결과값들 중에서 키워드와 매칭하여 검색결과를 제공한다. 이러한 방식의 검색서비스는 데이터가 가지고 있는 의미나 사용자가 원하는 검색결과들을 단순 키워드 매칭을 통해 제공하는 것으로 한계가 있다. 이러한 서비스들은 시맨틱 웹이 등장하면서 특정 도메인에 맞는 온톨로지(ontology)를 구축하여 데이터의 의미 혹은 데이터들 간의 관계를 파악하여 보다 정확하고 의미있는 검색결과를 제공하기 위해 많은 노력을 기울이고 있다.

1.2 연구목적

기존 키워드 중심의 검색은 이용자가 검색어

를 입력하면 해당 키워드의 의미를 고려하지 않고 단순히 매칭되는 검색결과만을 제공하기 때문에 이용자가 원하는 검색결과를 제공하는 데에는 한계가 있다. 또한 1차적으로 나온 검색결과 중에서 검색결과와 관련있는 다른 정보를 검색하기 위해서 또 다른 질의어를 입력하거나 다른 정보원천을 찾아서 검색하는 불편함이 있다. 예를 들어, 국가R&D정보에서 특정 “줄기세포”라는 검색어를 입력했을 때, 이용자는 줄기세포를 연구하는 과제정보나, 줄기세포를 연구하고 있는 인물, 줄기세포를 연구한 결과물인 연구성과 등을 검색결과로 보여주는데, 만일 과제정보를 검색하고 이 과제를 연구한 연구자가 연구한 다른 과제나 혹은 이 과제를 통해 나타난 연구성과 등은 무엇이 있을까 하는 의문이 생겼을 때, 단순 키워드 검색에서는 다시 키워드를 입력하거나, 링크된 정보를 클릭하면서 찾아가는 불편함을 겪게 된다.

온톨로지 기반의 검색은 앞서 언급된 불편함을 해소하기 위해 과제나 성과, 인물, 기관간의 의미관계를 파악할 수 있는 기반을 제공함으로써 이용자가 관심있어 하는 정보를 추론규칙을 통해 검색결과를 제공할 수 있다. 그러나 추론규칙을 생성하거나 의미기반의 검색서비스를 이용한다하더라도 사용자가 추론규칙을 이해하지 못한다면 직접 의미기반 서비스를 이용하는데 어려움을 겪게 될 것이다. 이러한 의미관계나 지식을 표현하는데 있어서 직관적으로 살펴볼 수 있는 인터페이스가 있다면 이용자는 별다른 추론규칙을 알지 못한다 하더라도 지식을 표현하는 인터페이스를 통해 의미있는 혹은 관심있는 검색결과를 찾아볼 수 있을 것이다.

본 연구의 목적은 국가과학기술지식정보서

비스(NTIS)에서 구축한 국가R&D정보를 분석하여 온톨로지를 구축하고 이용자가 관심있는 주제분야(과제, 인물, 성과, 기관)를 중심으로 온톨로지의 객체관계를 표현하는데 있으며, 정보 탐색은 지식맵 인터페이스를 제공하여 이용자가 원하는 질의를 정확히 인지하고, 검색결과를 보다 쉽게 제공하는 방법에 그 목적이 있다. 또한 온톨로지를 구축한 사례를 예시로 제공하여 온톨로지 구축을 쉽게 할 수 있도록 하며, 지식맵과 온톨로지의 제공방식을 융합하여 서비스를 제공함으로써 온톨로지에 대한 추론이나 질의어 생성 등을 잘 모르는 일반 사용자도 지식맵을 통해 서비스를 활용함으로써 사용자의 편의성을 제공한다는 점에서 의의가 있다.

1.3 연구방법

본 연구의 방법은 첫 번째 단계에서, 국가과학기술지식정보서비스(NTIS)에서 구축중인 국가R&D정보의 데이터 스키마, 객체 관계 등을 분석한 후, 연구정보 분야의 CERIF 모델을 기반으로 국가R&D 온톨로지를 구축하기 위한 대상을 정의한다. 두 번째 단계에서는 국가R&D 온톨로지를 구축하기 위해 Ontology Development 101과 METHONTOLOGY 개발 방법론의 절차들 중에서 필요한 부분을 선정하여 국가R&D 온톨로지 방법론을 정의하고, 방법론을 중심으로 온톨로지 구축절차를 수립하고, 온톨로지 생성규칙 등을 정의하여 국가R&D 온톨로지를 구축한다. 세 번째 단계에서는 국가R&D 온톨로지를 기반으로 기존의 키워드 검색방식에서 나타난 문제점을 보완하기

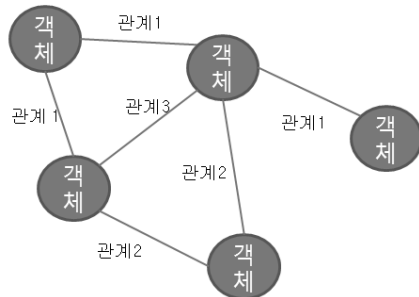
위한 방법을 분석하고, 온톨로지를 이용하여 이용자에게 국가R&D와 관련한 지식을 전달하기 위해 효율적이면서 의미전달을 명확히 하기 위한 인터페이스를 설계한다. 네 번째 단계에서는 지식맵 인터페이스를 활용하여 이용자에게 정확한 의미전달과 관심있는 질의 등을 주제분야(과제, 성과, 인물, 기관)와 객체(과제, 성과, 인물, 기관)간의 관계를 고려하여 추론규칙 등을 정의한다. 마지막으로 정의된 주제분야와 객체, 추론규칙 등을 지식맵 인터페이스를 구축하여 이용자에게 국가R&D정보에 대한 지식을 쉽고 편리하게 이용할 수 있는 서비스를 구현한다.

2. 관련연구

2.1 지식맵(Knowledge Map)

지식맵(Knowledge Map)은 “개념을 표현하는 이미지 혹은 모양과 개념들 사이의 관계를 표현하는 화살표의 연속성에 의해 데이터들 간의 관계를 표현하는 그래픽 인터페이스이다. 지식을 표현하는 형태이자, 개념맵으로 알려져 있다.”라고 위키피디아에서 정의하고 있다(Wikipedia).

일반적으로 지식맵은 노드와 링크로 구성되는데, 노드는 객체를 표현하고 링크는 객체들 간의 관계를 표현한다. 노드는 객체가 가지고 있는 개념에 대한 정의와 상세정보를 담고, 링크는 객체들간의 관계나 의미 등을 표현하는데 활용된다. 본 논문에서의 지식맵은 링크의 관계보다는 노드와 노드들 간의 관계를 의미하는 질의어를 표현하는 관점에서 활용하였다(〈그림 1〉 참조).



〈그림 1〉 지식맵-노드와 링크

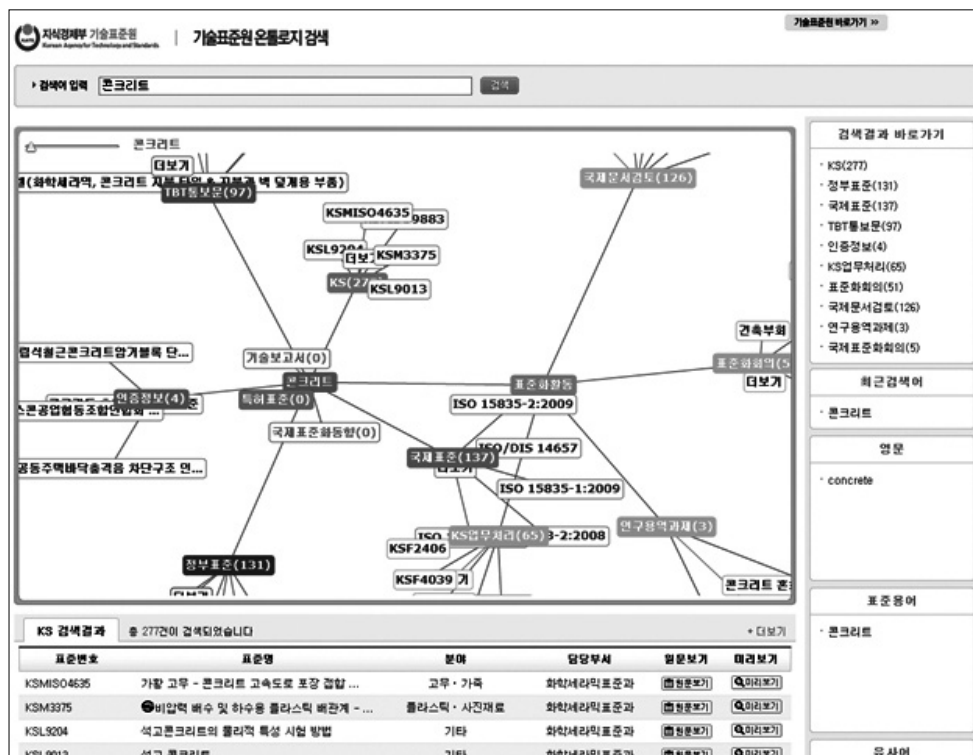
정부 규격, KS 표준, 국제 표준 등 표준 정보와 KS 표준 분류체계, ICS 표준 분류체계 등 다수의 분류체계 정보를 중심으로 관련 정보를 연계·검색할 수 있도록 하였다. 기술표준에 대한 통합 및 연계검색이 가능하고 특화된 도메인에 대하여 서비스 활용도가 높다(〈그림 2〉 참조).

2.2 지식맵(Knowledge Map) 서비스 사례

2.2.1 기술표준원 의미기반서비스 사례 기술표준원에서 제공하는 의미기반서비스는

2.2.2 화학공학연구정보센터 연구자네트워크 사례

화학공학연구정보센터가 보유한 논문정보를 분석하여 공저자 관계를 네트워크 맵 형태로 서비스한다. 중심인물 외 타 인물을 더블클릭하여 인물 간 내비게이션이 가능하고, 마우스 온(Mouse On)시 공저논문 수, 소속기관 정보화



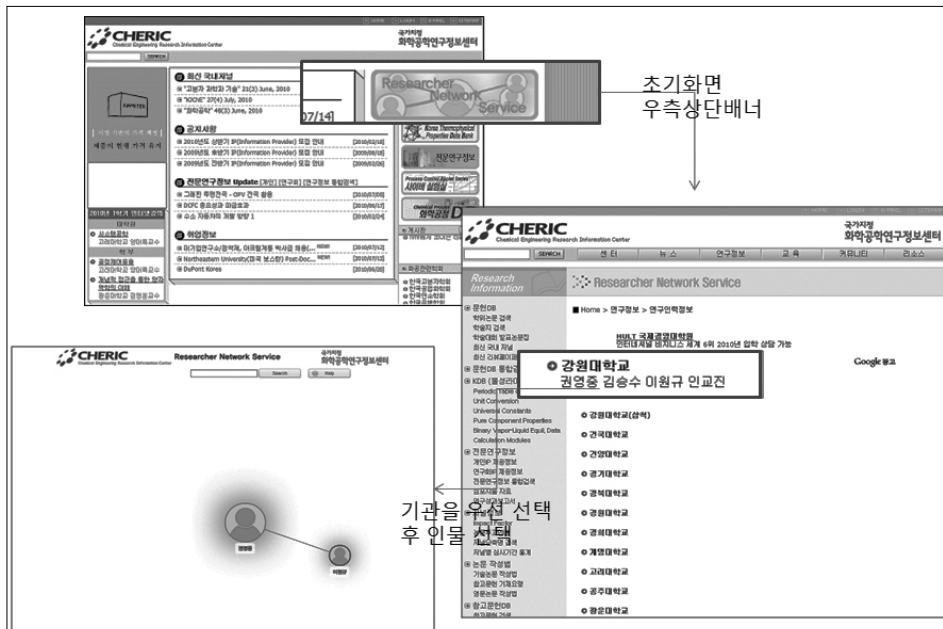
〈그림 2〉 기술표준원 의미기반서비스 사례

인이 가능하다. 다른 서비스로의 링크가 없어 인물 상세정보 및 논문 상세 정보 제공은 미흡한 편이다(〈그림 3〉 참조)(<http://www.cheric.org/research/people/rns.php>).

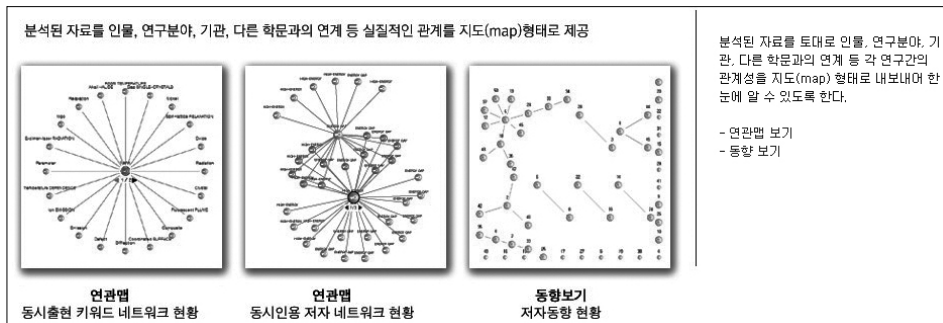
2.2.3 물리화학연구정보센터의 지식맵서비스

물리화학연구정보센터의 지식맵서비스는 WoS

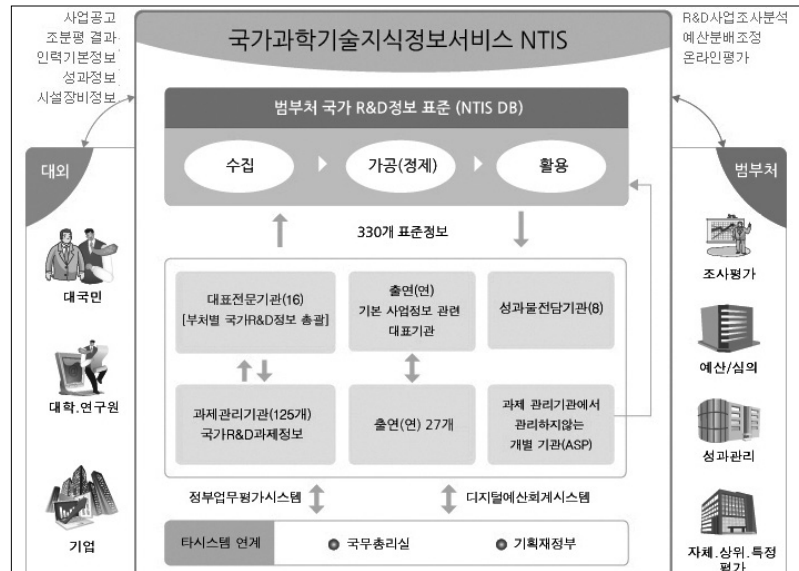
(Web of Science)의 물리분야 논문 데이터에 대한 키워드 관계, 논문에 대한 피인용 관계를 활용하여 논문의 연관관계와 저자간의 연관관계를 맵형태로 서비스하고 있다. 연구자와 키워드에 대하여 검색하고, 연구자 및 논문에 대한 상세정보 연계기능을 제공한다(〈그림 4〉 참조)(<http://km.icpr.or.kr/>).



〈그림 3〉 화학공학연구정보센터 연구자네트워크 서비스



〈그림 4〉 물리화학연구정보센터 지식맵서비스



〈그림 5〉 NTIS 개념도

2.3 국가과학기술지식정보서비스 (NTIS)

국가과학기술지식정보서비스(NTIS)에서는 2006년부터 국가R&D정보를 각 정부부처와 연계하여 데이터베이스를 구축하여 연구자, 정책결정자, 정부부처관계자들에게 다양한 서비스를 제공하고 있다. 2010년 이후부터는 고도화 사업을 통해 국가R&D정보를 보다 효율적으로 제공하기 위한 노력을 기울이고 있다.

NTIS에서는 국가R&D정보 객체간 의미를 파악하여 서비스 제공을 하기 위해 시맨틱 웹 기술을 도입하고 있다. 국가R&D정보를 서비스 할 때 기존에 정보검색엔진을 통하여 검색결과 리스트를 보여주고, 사용자의 선택에 의해 상세정보와 관련 있는 정보는 링크를 통해 제공하는 방식에서 벗어나 정보객체간의 관계를 분석하여 국가R&D 온톨로지를 구축하고, 온톨

로지 데이터를 활용하여 의미기반의 서비스를 제공하고 있다(〈그림 5〉 참조)(<http://www.ntis.go.kr>).

3. 국가R&D 온톨로지

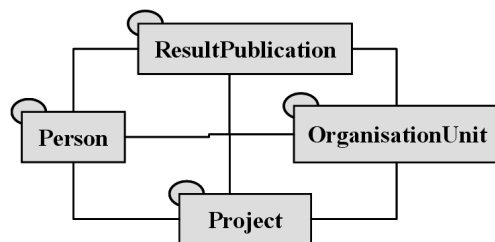
3.1 국가R&D 온톨로지 구축 범위

국가R&D 온톨로지 구축 대상은 NTIS에서 구축한 데이터들을 바탕으로 활용 가능한 데이터베이스의 내용을 포함한다. NTIS의 데이터베이스는 과제정보, 참여인력정보, 평가위원정보, 성과정보, 연구시설·장비기자재 정보 등 330개 세부항목으로 구성되어 있다. 330개 세부항목 중 연구정보 분야의 CERIF 모델을 기반으로 국가R&D 온톨로지 모델을 도출하였다(〈표 1〉 참조)(<http://www.cordis.lu/cerif>).

<표 1> NTIS데이터베이스 정보항목

구분	정보항목
과제정보	사업정보, 기본정보, 연구비정보, 위탁과제/공동연구, 역할매핑정보, 참여인력분포, 제재정보, 연구개발 공고, 연차정보
성과정보	연수지원, 연수지원 통계, 기술거래, 논문(게재성과), 사업화, 산업재산권, 인력양성(배출실적), 생물자원(성과물), 생명정보(성과물), 화합물(성과물), 소프트웨어(성과물), 연구보고서(성과물), 기술요약정보
참여인력정보	경력사항, 기본정보, 논문실적, 자격사항, 재직기관사항, 지식재산권, 최종학위사항, 학력사항
평가위원정보	소속학회현황, 지역서 실적, 훈포장 사항

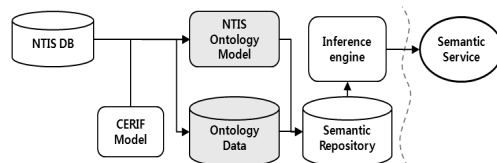
CERIF 모델은 유럽 국가들의 연구정보를 효율적으로 관리하고 서로 공유하기 위해 개발된 표준 포맷으로, 유럽연합 산하 각국의 연구기관들이 생성한 다양한 형식의 연구성과물들의 상호교환 및 관리를 위한 종합 모델이다. CERIF는 유럽연합의 연구정보시스템인 CRIS (Current Research Information Systems)에서 제공하고 있는 다양한 개체(Entities) 및 속성(Attributes)들 간의 상호호환성을 지원할 수 있는 종합모델(Full Model)이라는 점에서 의의를 가진다(<그림 6> 참조).



<그림 6> CERIF 코어 엔티티

가능한 부분을 검토한 후 적용하고, 이외 부분에 대하여 새롭게 설계하여 국가R&D 온톨로지 모델을 도출하였다.

국가R&D 온톨로지 모델링 시 정의된 NTIS DB로부터의 관계를 활용하여 온톨로지 데이터를 수집하기 위한 데이터 변환(매핑) 규칙을 정의하고, 이를 바탕으로 온톨로지 데이터(인스턴스)를 수집하고, 수집된 온톨로지 데이터(인스턴스)를 온톨로지 저장소에 적재하여 데이터가 서비스 가능한 형태로 구성하였다. 온톨로지 저장소에 적재된 데이터를 추론엔진을 통하여 추론 서비스를 구축하여 시맨틱 서비스가 이루어질 수 있도록 하였다(양명석 외, 2011)(<그림 7> 참조).



<그림 7> 국가R&D 온톨로지 구축 절차

3.2 국가R&D 온톨로지 구축 절차

국가R&D 온톨로지 구축 절차는 CERIF 모델에 정의되어 있는 개념 및 관계정보에 대하여 NTIS 환경에 적합성 여부를 검토하여, 반영

3.3 국가R&D 온톨로지 모델링

국가R&D 온톨로지 모델링 방법은 일반적인 온톨로지 모델링 절차에 따라 구축되는데 특히,

Ontology Development 101과 METHONTOLOGY 개발 방법론의 일부를 적용하여 수립하였다.

Ontology Development 101은 인공지능 분야 연구기관인 스탠포드 대학 지식 시스템 연구소의 핵심 연구과제 중 하나로서, 와인 및 음식분야의 지식을 바탕으로 온톨로지를 구축하였고, 처음 온톨로지를 구축하고자 하는 개발자에게 도움을 주는 방향을 제시하였다. 그러나 온톨로지 구축에 필요한 전체적인 구축 절차를 포함한 완성도를 갖춘 방법으로 판단할 수는 없다. 개념 구조화의 기초적인 절차를 설명하는 방법이다.

METHONTOLOGY 개발방법론은 소프트웨어공학과 지식공학 분야의 방법론에서 많은 영향을 받아, 온톨로지 개발에 있어서도 포괄적인 방법론을 제공하고 있다. 생명주기 모형을 지향함으로써, 개발 과정에서 반드시 수행해야 하는 활동들의 순서 집합을 가지고 있다. METHONTOLOGY는 온톨로지 개발에 있어 프로토타이핑 기법을 사용한다. 프로토타이핑이란 점진적 접근법이라 할 수 있는데, 제한된 범위에서 시행착오를 반복하여 온톨로지를 만들어 가는 과정이다. 이를 통해 온톨로지 개발자는 요구사항을 단순화하여 온톨로지를 구축할 수 있다(Natalya & McGuinness, 2001).

국가R&D 온톨로지 모델링 방법론은 온톨로지 구축단계를 크게 5단계(온톨로지 명세화, 온톨로지 개념화, 온톨로지 구현, 온톨로지 구축, 유지보수)로 이루어진다.

첫째, 온톨로지 명세화 단계에서는 온톨로지의 사용 목적, 구축 범위, 온톨로지의 정형성 단계, 잠재 사용자에게 대한 명세 등을 정의한다.

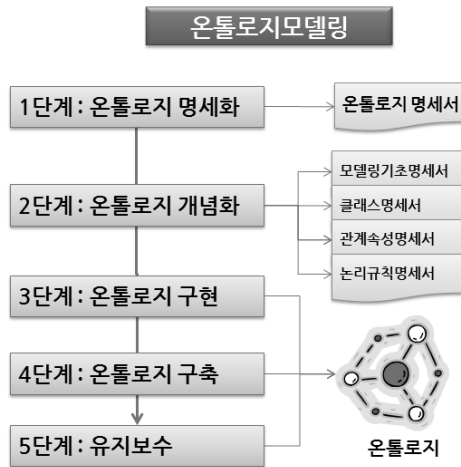
둘째, 온톨로지 개념화(설계) 단계에서는 모델링 기초명세화, 클래스 명세화, 관계속성 명세화, 논리규칙 명세화 단계를 거치는데, 온톨로지 모델링 대상이 되는 데이터를 기반으로 개념과 개념간의 관계 등을 도출하고, 온톨로지의 클래스, 클래스 관계, 클래스 속성 등을 정의한다. 모델링 기초명세화 단계에서는 우선 국가R&D 관련 정보에서 해당분야에서 사용되는 용어를 정리하여 용어사전을 정의하고, 정의된 용어 중 개념으로 분류될 수 있는 것들의 상하위 관계를 트리구조로 구성하는 개념분류 트리를 정의한다. 클래스 명세화 단계에서는 개념간의 관계에 대하여 개념관계를 정의하고, 앞서 정의한 개념들 중 클래스로 구분된 클래스들에 대하여 이름, 속성, 관계, 계층 등을 정의한다. 관계속성 명세화 단계에서는 기초 명세서, 클래스 명세서를 바탕으로 클래스간의 관계(Object property), 클래스의 속성(Datatype Property), 제약조건(Restriction) 등을 상세히 정의한다. 관계속성 명세화단계에서는 온톨로지를 정의할 때 중요한 개념들간의 의미나 관계 등을 고려하여 추론할 수 있는 규칙등을 정의하는데, 온톨로지에 포함되어야 하는 공리(RDF, RDFS, OWL-LITE 등의 추론 규칙), 규칙(기본 추론 규칙 이외의 해당 온톨로지 분야에서 통용되는 명시적 규칙) 등을 정의한다.

셋째, 온톨로지 구현 단계에서는 실제 개념화 과정을 통해 도출된 개념과 관계, 속성, 제약조건 등을 이용하여 실제 온톨로지 언어를 이용하여 온톨로지(모델)를 구현한다.

넷째, 온톨로지 구축 단계에서는 구현된 온톨로지 모델을 이용하여 실제 온톨로지 데이터를 추출하고, 추출된 온톨로지 데이터는 온톨

로지 저장소에 적재하고 추론기를 이용한 추론 서비스를 제공할 수 있도록 한다.

마지막으로, 유지보수단계에서는 환경과 데이터 변화 또는 논리적 오류 발생으로 인한 개념, 관계, 속성, 제약조건 등의 변화에 대응하여 온톨로지 모델을 수정 보완하면서 온톨로지를 구축한다(〈그림 8〉 참조).



〈그림 8〉 온톨로지 모델링 방법론

국가R&D 온톨로지를 구축하기 위한 용어사전은 국가R&D정보를 분석하여 용어들의 리스트를 정리하고 온톨로지에서 사용할 용어와 영문용어, 용어에 대한 설명, 그리고 용어의 유형을 정의한다(〈표 2〉 참조).

국가R&D 온톨로지의 클래스는 thing을 기

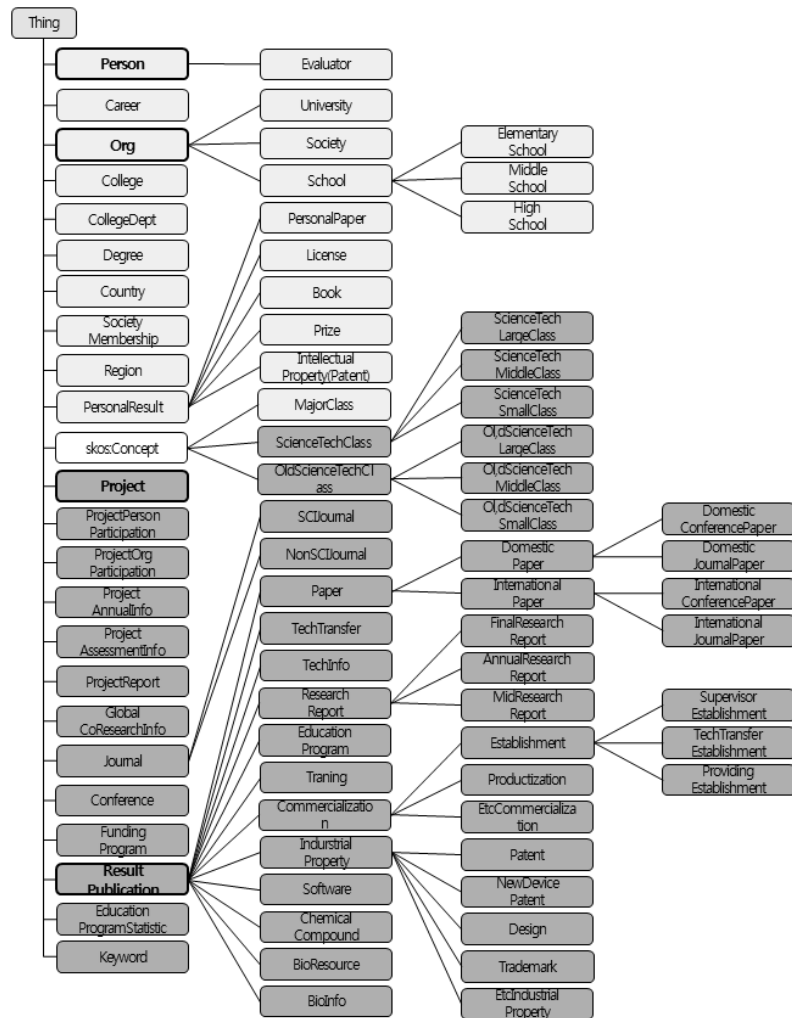
본으로 NTIS에서 제공하는 핵심정보인 인물(Person), 기관(Org), 과제(Project), 성과(Result Publication) 등을 기본개념 클래스로 정의하고, 인물이 가지는 특성 학력사항, 경력사항, 지역사항 인물이 가지는 실적정보 등을 하위개념으로 정의하였으며, 과제와 관련된 과제참여정보(ProjectPerson participation), 과제연차정보(Project Assessment Info), 사업정보(Funding Program) 등을 정의하였으며, 성과정보와 관련된 논문(paper), 기술정보(Techinfo), 또한 NTIS에서 사용하는 기본 분류 과학기술표준분류(ScienceTechClass), 지역코드(Region), 전공코드(MajorClass) 등을 개념클래스로 분류하여 추론시 이러한 코드정보를 활용할 수 있도록 구축하였다(〈그림 9〉 참조).

국가R&D 온톨로지의 클래스(Class)는 개념클래스를 중심으로 정의하고, 부모클래스와 클래스가 갖는 관계속성 등을 정의한다. 관계속성은 속성에 대한 이름, 용어, 대상(range), 그리고 속성이 갖는 설명 등을 정의한다(〈표 3〉 참조).

국가R&D 온톨로지의 클래스관계명세는 클래스간의 관계를 속성으로 정의하는데 학위 같은 경우, 학위는 '대학교, 단과대학, 학과의 코스를 마치다'라는 의미를 갖게 되어 학위(degree)라는 클래스와 대학교(University), 단과대학(college), 학과(CollegDept) 등의 클래스와의 관계를 정의하는 것을 말한다(〈표 4〉 참조).

〈표 2〉 용어사전 정의

도메인	용어	영문용어	설명	유형
공통	기관	Org	인력의 재직 기관	개념
- 생략 -				
공통	구성원을 갖다	hasOrgMember	기관은 사람을 구성원으로 갖는다.	관계



〈그림 9〉 국가R&D 온톨로지 개념트리

〈표 3〉 클래스 사전

클래스 명	Person	용어	인력	도메인	인력
설명	인력				
부모 클래스	Thing				
관계 속성					
속성명	용어	대상(Range)		설명	
hasCareer	경력을 갖다	Career		인력은 경력을 가진다.	
- 생략 -					
hasPersonalResult	개인실적을 갖다	PersonalResult		인력은 개인 실적(지역서, 논문, 특허, 자격증, 훈포장)을 갖는다.	

<표 4> 클래스 관계명세

속성명	용어
completedAt	학위를 완료하다.
Domain	Range
Degree	University, College, CollegeDept
설명	
대학, 단과대학, 학과로부터 학위가 완료되다(인력이 학위를 받기 위해서 대학, 단과대학, 학과의 Course를 완료하더라는 의미).	
Inverse Property	Equivalent Property
Super Property	Sub Property
characteristics	

국가R&D 온톨로지를 활용하기 위해서는 클래스들간의 속성정보 혹은 관계를 이용하여 의미를 파악할 수 있는데, 이것 이외에 많이 쓰이는 추론규칙 등을 논리규칙으로 정의하여 활용할 수 있다. “동일 과제의 공동 참여인력”이라는 논리규칙을 정의할 경우, 인물이 수행한 과

제를 찾고, 그 과제중에 인물이 아닌 다른 사람이라는 의미인데 질의어로 반복적으로 활용할 때 어려움을 겪게 된다. 이럴 경우 미리 논리규칙을 정의해 놓으면 반복적인 질의를 수행할 필요없이 간단하게 질의 처리를 할 수가 있게 된다(<표 5> 참조).

<표 5> 논리 규칙 관계명세

관계명	ntis:hasCoParticipant
설명	동일 과제의 공동 참여 인력
관계 (ORL)	(?Person1 ntis:participateIn ?Project) (?Project ntis:hasParticipant ?Person2) NotEqual(?Person1, ?Person2) -> (?Person1 ntis:hasCoParticipant ?Person2)

<표 6> 인스턴스 변환 규칙 명세(기관정보)

데이터 수집 명	Org		
수집 대상 Field	수집 SQL		
• ORGURI • CD_CODE • CD_NM	SELECT concat('Org_',REPLACE(REPLACE(REPLACE(cd_code,' ','_'),'(','_'),'_')) AS OrgURI, cd_code, cd_nm FROM tcd_code W HERE cd_gb='ORG'		
종류	Subject	Predicate	Object
객체	ORGURI	rdf:type	ntis:Org
속성	ORGURI	ntis:hasName	CD_NM (xsd:string)
속성	ORGURI	ntis:hasOrgID	CD_CODE (xsd:string)

국가R&D 온톨로지를 구축하기 위해서는 실제 NTIS DB에서 관련된 정보를 추출해내는 과정이 필요하다. 인스턴스를 추출하기 위해서는 미리 정의된 클래스 정보와 해당하는 DB내에서 정보를 매핑하여 온톨로지 인스턴스를 생성하게 된다. 즉, 실제 DB내에서 SQL 쿼리를 이용하여 데이터를 추출하는 과정이라고 생각하면 된다(〈표 6〉 참조).

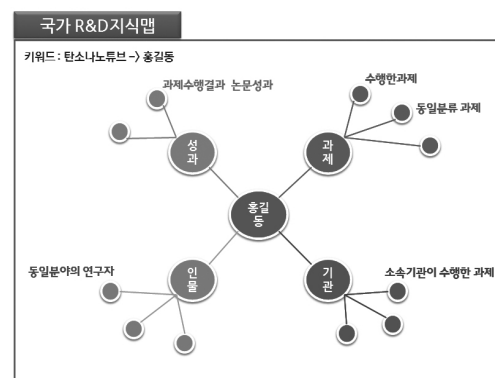
4. 온톨로지 기반의 지식맵 구축

4.1 국가R&D지식맵

국가R&D지식맵은 국가R&D정보 중에서 이용자가 관심있어 하는 객체를 중심으로 4가지 주제분야로 구성하였다. 주제분야는 국가R&D 정보의 가장 중심이 되는 과제와 과제를 수행하는 주체인 인물, 인물이 소속된 소속기관, 그리고 과제를 통해 수행되어 얻어진 연구성과 등 이렇게 4가지 주제분야를 선정하였다. 객체와 주제분야(과제, 인물, 성과, 기관)와 객체(Object), 그리고 객체들 간의 관계(Relation)를 고려하여 사용자가 원하는 정보를 미리 추론규칙에 정의하여 의미기반의 검색을 지원할 수 있도록 하였다.

국가R&D지식맵 구조는 일반적인 지식맵의 형태인 노드(Node)와 링크(Link) 등을 이용하여 인터페이스를 구성하는데, 노드에 상세정보를 제공하고 링크에는 노드들 간의 관계를 제공하는 것과 달리, 가운데에 이용자가 선택한 객체를 중심노드(예, 홍길동이라는 인물이 선택됨)로 하고 객체와 관련있는 주제분야(과제, 인

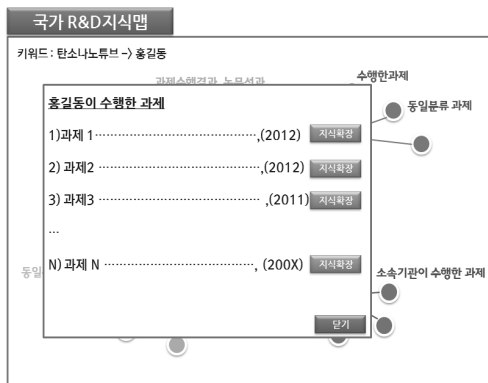
물, 성과, 기관)를 하위 노드로 하여 구성하였다. 객체와 주제분야간의 관계성을 고려하여 객체와 관련이 있거나 이용자가 검색하고자 하는 결과를 제공하기 위해 추론규칙(예: 수행한 과제)들을 하위노드로 표현하여 검색결과를 팝업형태로 제공하도록 설계하였다(〈그림 10〉 참조).



〈그림 10〉 지식맵 설계

예를 들면, 키워드 '탄소나노튜브'를 검색하였을 때 검색결과 중 홍길동을 선택하여 지식맵을 요청하면 지식맵을 볼 수 있다. 사용자가 홍길동과 관련된 과제 분야를 선택하면, 이용자에게 홍길동과 관련있는 과제들에 대한 질의를 할 수 있는 하위 노드들이 펼쳐진다. 하위 노드중에서 사용자가 "수행한 과제"를 선택하면 홍길동이 수행한 과제 검색 결과가 리스트로 나타난다. 이용자들이 특별히 SPARQL을 몰라도 "홍길동이 수행한 과제"를 시스템에서는 SPARQL 질의로 변환하고, 시맨틱 질의어 추론기를 이용하여 검색결과를 요청하면 추론엔진을 통해 검색결과 리스트를 XML로 형태로 얻게 된다. XML결과는 지식맵 인터페이스를 통해 사용자에게 검색결과를 알려준다. 또한 검색결과들 중,

과제제목 옆 “지식확장” 버튼을 선택하면 그 과제를 중심노드로 지식맵이 다시 그려지는 형태로 구성된다. 기존 검색에서는 탄소나노튜브를 검색하고 홍길동이라는 인물이 어떤 과제를 수행했는지 알기 위해 키워드로 만들어서 검색한 데 비하여 지식맵을 이용하면 사용자가 원하는 의도를 손쉽게 변환하고, 검색결과를 제공한다(<그림 11> 참조).



<그림 11> 지식맵 검색결과 화면 설계

4.2 추론규칙

추론규칙은 주제분야와 객체간의 관계를 고려하여 이용자가 관심있어 하거나 의미있는 질의를 선정하여 정의하였다. 지식맵의 중심노드에 어떤 주제분야가 오느냐에 따라 과제중심맵, 성과중심맵, 인물중심맵, 기관중심맵 등 4가지 맵으로 구성된다. 과제중심맵은 선택된 과제객체를 중심으로, 성과중심맵은 선택된 성과객체를 중심으로, 인물중심맵은 선택된 인물객체를 중심으로, 기관중심맵은 선택된 기관을 중심으로 주제분야의 정보들을 추론규칙으로 미리 정의하여 파라미터 값에 따라 추론결과를 제공할 수 있도록 구성하였다. 객체가 과제로 선택되었을 경우, 과제중심맵의 추론규칙들이 보이고, 이용자가 관련 주제분야를 선택하면 정의된 추론규칙들이 하위노드로 표시되어 이용자가 하위 노드를 선택하면 검색결과를 리스트로 볼 수 있다(<표 7> 참조).

<표 7> 과제중심맵의 추론규칙

No	추론규칙	관련주제	과제객체와 관련된 추론 정보
1	동일 분류의 과제	과제	과제와 동일한 분류정보(최대4개) 선택된 분류의 과제
2	동일한 사업 과제	과제	동일한사업과제
3	연구책임자가 참여한 과제	과제	연구책임자가 참여한 과제
4	연속과제	과제	연속과제
5	주관 기관이 수행한 과제	과제	주관기관이 수행한 과제
6	도출한 연구성과	성과	도출한연구성과
7	동일분류과제의 연구성과	성과	과제와동일한분류(과제의 분류명) 선택된 분류명의 과제를 찾고 그과제로부터 나온 연구 성과
8	참여자의 연구성과	성과	과제에 참여한 인물 선택된 인물의 연구성과
9	주관 연구 기관	기관	주관연구기관
10	과제 분류와 같은 분야의 연구자	인물	과제분류 선택된 분류와 같은분야의 연구자
11	참여 연구자	인물	참여연구자

“동일 분류의 과제”의 추론규칙은 선택한 과제와 동일한 과학기술표준분류의 과제를 검색하는데, 1차적으로 과제의 분류정보를 추론하고, 선택한 과제분류에 대해서 동일분류의 과제를 검색하는 것이다. 이를 SPARQL 질의어로 구성하면 <표 8>과 같은데, 과제가 속한 표준분류의 분류명을 받고 해당하는 분류코드의 과제들을 선택하여 결과를 가지고 오는 질의어다. ntis:hasSameFieldProject는 미리 정의된 같은 필드를 갖는 과제를 의미하는 논리규칙이

다. ntis:hasName은 객체의 이름을 얻는 논리규칙이다.

이외에도 성과중심맵은 “유발과제”, “제출한 연구자” 외 9개의 추론규칙이 있으며, 인물중심맵은 “참여과제”, “참여한 과제의 성과” 등 9개, 기관중심맵은 “주관 연구기관으로 수행한 과제” 등 7개의 추론규칙을 미리 정의하여 서비스를 제공하고 있다. 각각의 추론규칙은 사용자의 객체선택에 따라 동적으로 구성하도록 <표 9, 10, 11>과 같이 SPARQL을 적용하여 구현하였다.

<표 8> 동일분류의 과제의 SPARQL 예(과제중심맵)

SARAQL	<pre> PREFIX ntis: <http://www.ntis.go.kr/ontology/ntis#> SELECT DISTINCT ?u ?name WHERE { <\$uri\$> ntis:hasSameFieldProject ?u. ?u ntis:hasName ?name. }</pre>
--------	--

<표 9> 유발과제의 다른 연구성과(성과중심맵)

SARAQL	<pre> PREFIX ntis: <http://www.ntis.go.kr/ontology/ntis#> select distinct ?Result ?ResultName where{ <\$uri\$> ntis:resultFrom ?Project. ?Project ntis:hasResultPublication ?Result. ?Result ntis:hasName ?ResultName. FILTER(?Result!=<\$uri\$>) }</pre>
--------	---

<표 10> 참여과제의 연구자(인물중심맵)

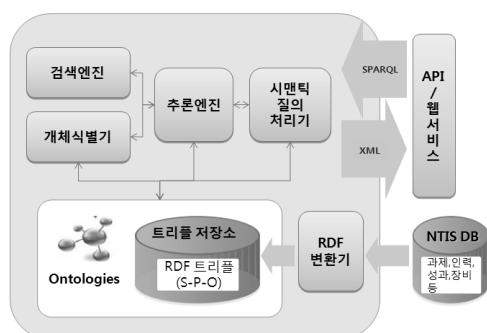
SARAQL	<pre> PREFIX ntis: <http://www.ntis.go.kr/ontology/ntis#> SELECT DISTINCT ?u ?name ?OrgName WHERE { <\$uri\$> ntis:hasCoParticipant ?u. ?u ntis:hasName ?name. ?u ntis:workFor ?Org. ?Org ntis:hasName ?OrgName. }</pre>
--------	--

〈표 11〉 유사분야 연구기관(기관중심맵)

SARAQL	<pre> PREFIX ntis: <http://www.ntis.go.kr/ontology/ntis#> SELECT DISTINCT ?u ?name WHERE { <\$uri\$> ntis:hasSameFieldProjectParticipatedOrg ?u. ?u ntis:hasName ?name. } </pre>
--------	--

4.3 온톨로지 기반의 지식맵 구현

온톨로지 기반의 지식맵은 앞 절에서 설계한 지식맵을 이용하여 사용자가 원하는 질의를 SPARQL 질의로 변환하여 시맨틱 질의처리기에 전달한다. 시맨틱 질의처리기는 SPAQL 질의를 분석하여 추론엔진에게 전달하고, 추론엔진은 이를 분석하여 트리플 저장소로부터 검색결과를 제공받고, 식별기를 통해 객체를 식별하고 검색결과를 XML 형태로 제공하면 지식맵 인터페이스에서는 XML을 분석하여 검색결과를 제공하도록 구성하였다(〈그림 12〉 참조).



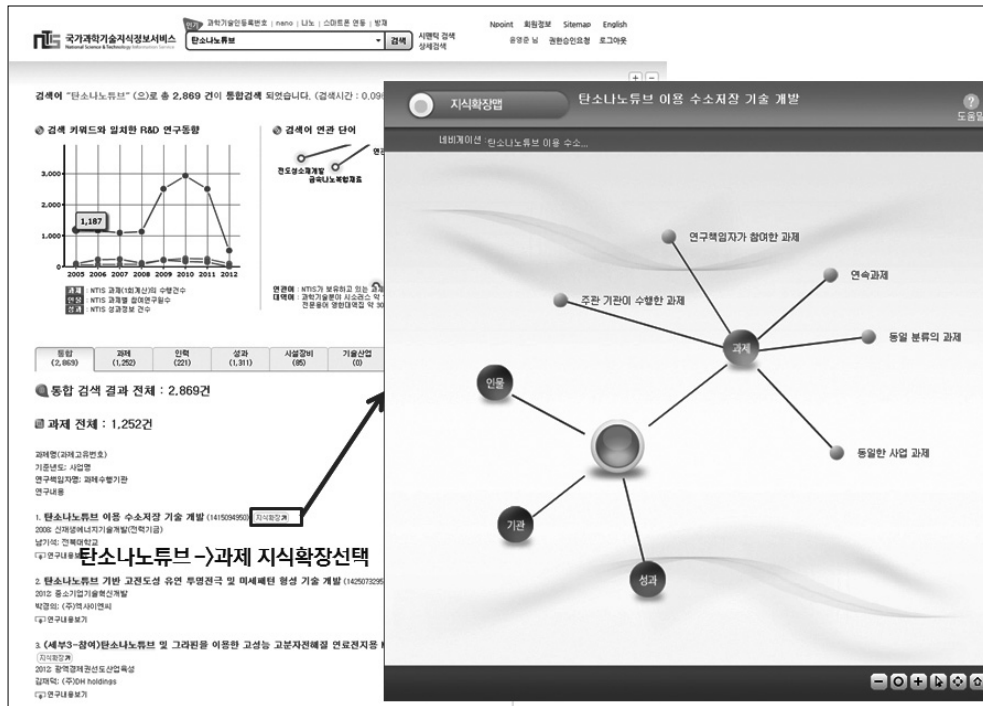
〈그림 12〉 온톨로지 기반 지식맵 서비스 구성도

국가R&D 지식맵은 NTIS의 통합검색에서

사용자가 입력한 키워드로 검색을 하고, 검색 결과들 중에서 사용자가 원하는 검색결과(과제, 인력, 성과 등) 중에서 “지식확장맵”을 버튼을 선택하면 팝업형태로 제공하도록 구현하였다(〈그림 13〉 참조).

‘탄소나노튜브’를 검색어로 입력한 결과들 중에서 ‘탄소나노튜브 이용 수소저장 기술개발’의 지식확장맵을 선택하면 과제중심의 지식맵이 제공된다. 중심노드를 이용자가 선택하면 해당하는 과제의 상세정보를 볼 수 있다(〈그림 14〉 참조).

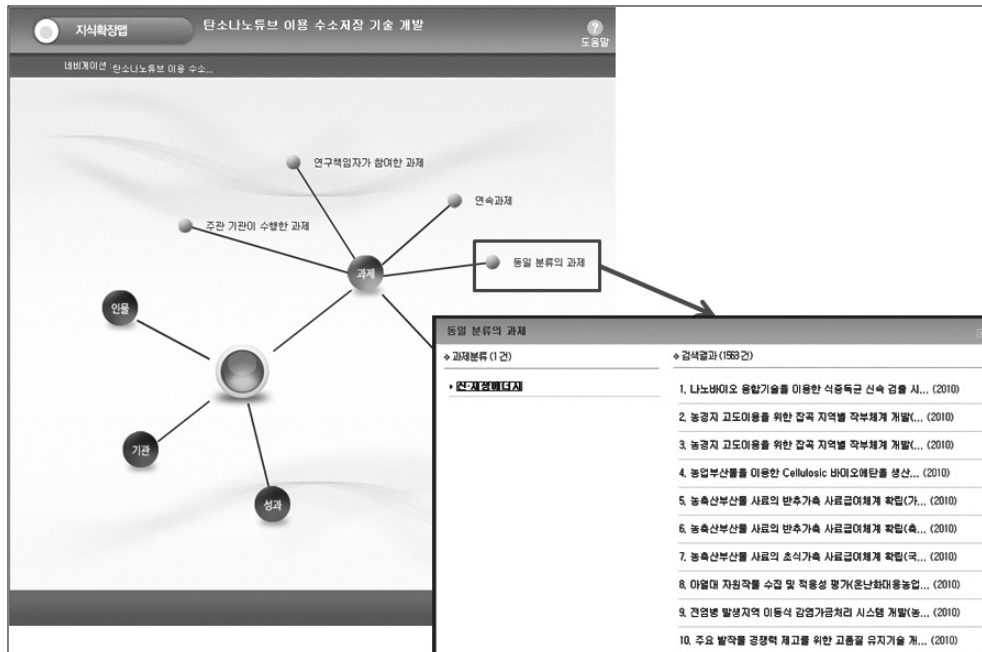
이용자가 선택한 과제 ‘탄소나노튜브 이용 수소저장 기술개발’에서 4가지 주제분야에서 과제 분야를 선택하면 앞서 정의한 과제 관련 추론규칙들을 제공하고, 이용자가 여러 가지 추론규칙들 중에서 특정 추론규칙을 선택하면 추론규칙들을 SPAQL언어로 변환하여 검색결과를 사용자에게 제공하도록 구현하였다. 만일 사용자가 추론규칙들 중에서 “동일분류의 과제”를 선택하면, 즉 “탄소나노튜브 이용 수소저장 기술개발”과 같은 분류의 다른 과제들을 검색하는 의미로 해석할 수 있다. 이용자는 선택한 과제가 어떤 분류를 갖고 있고, 이 분류에 해당하는 다른 과제들을 찾는다는 의미를 지식맵 인터페이스를 제공함으로써 시스템이 이해하여 검색결과를 제공한다. 기존 검색엔진을 활용한다면 이용자가 과제가 갖는 분류를 찾은 다음, 다시



〈그림 13〉 지식맵 서비스구현 화면(과제중심맵)



〈그림 14〉 지식맵 상세정보 보기



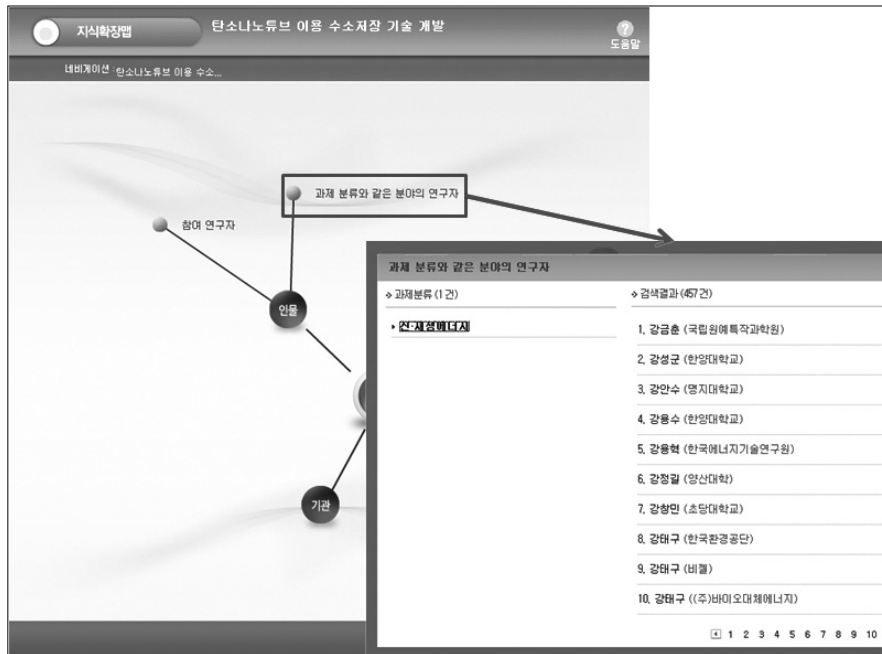
〈그림 15〉 “동일분류의 과제” 검색결과

분류코드로 키워드를 입력하여 검색결과를 찾거나 혹은 분류코드명으로 검색을 하여 결과를 얻어야 할 것이다. 이는 이용자가 원하는 의미를 이해하지 못하고 또한 정확한 검색결과도 제공하지 못하는 결과를 가져올 것이다(〈그림 15〉 참조).

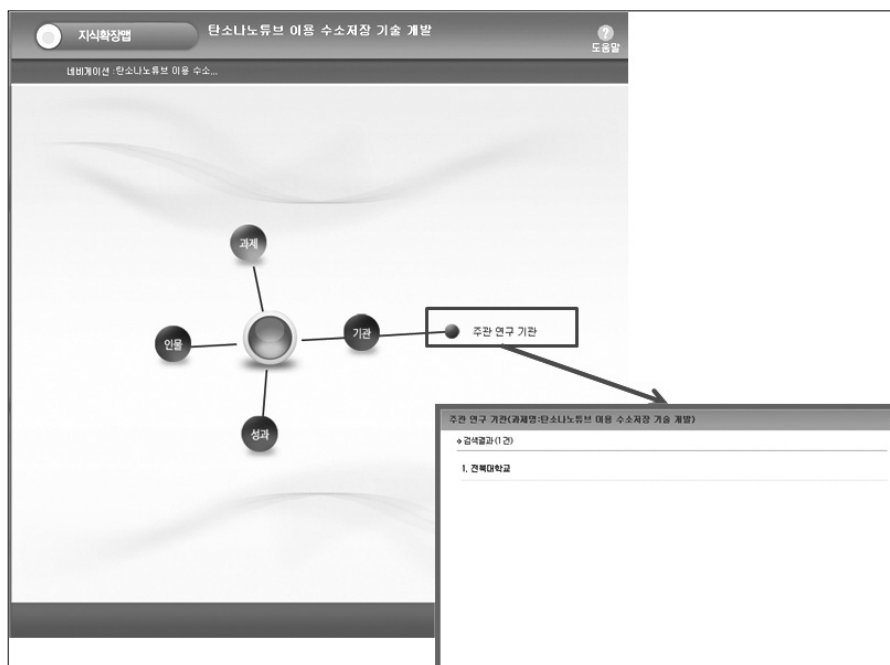
또한 이용자가 4가지 주제분야에서 인물분야를 선택하면 앞서 정의한 인물 관련 추론규칙들을 제공하고, 이용자가 여러 가지 추론규칙들 중에서 특정 추론규칙을 선택하면 과제분야와 마찬가지로 형태로 서비스를 이용할 수 있다. 추론규칙들 중 “과제분류와 같은 분야의 연구자”를 선택하면 과제의 분류를 보여주고 사용자가 분류를 선택하면 분류에 해당하는 인물을 검색할 수 있다. 기존의 키워드 중심의 검색서비스에서는 과제의 분류인 “신재

생에너지” 분야의 인물을 검색하기 위하여 통합검색으로 가서 “신재생에너지”라는 검색어를 입력하고, 인물에서 검색분야를 직접 찾아야 되는 번거로움이 있을 것이다. 이렇게 지식맵에서는 이용자에게 손쉬운 접근경로를 제공한다. 이는 기존 검색과 차별성을 보인다(〈그림 16〉 참조).

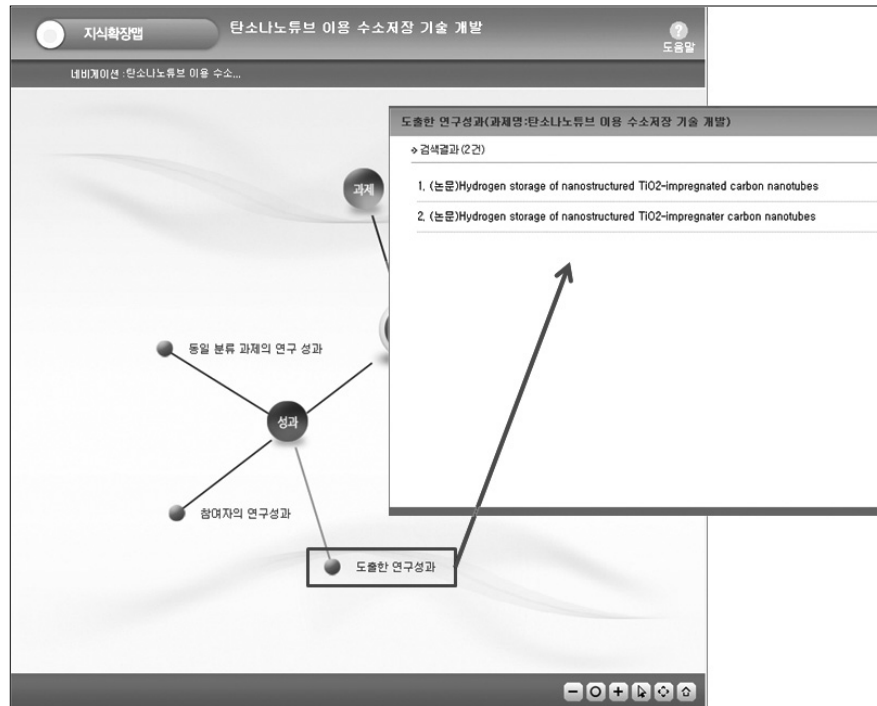
이밖에도 이용자가 기관분야를 선택하면 기관과 관련된 추론규칙을 볼 수 있으며, “주관연구기관”을 선택하면 과제를 전북대에서 수행하고 있음을 알 수 있다(〈그림 17〉 참조). 또한 이용자가 성과분야를 선택하면, 성과와 관련된 추론규칙들을 제공하고 추론규칙들 중에서 “도출한 연구성과”를 선택하면 과제에서 도출한 연구성과를 검색결과로 제공한다(〈그림 18〉 참조).



〈그림 16〉 “과제분류와 같은 분야의 연구자” 검색결과



〈그림 17〉 “주관연구기관” 검색결과



〈그림 18〉 “도출한 연구성과”의 검색결과

이용자가 검색결과들 중에서 인물객체를 선택하면, 앞서 살펴본 과제중심맵과 같이 주제분야를 선택하여 같은 방식으로 인물객체와 주제분야들 간의 추론규칙을 보고 원하는 검색결과를 살펴볼 수 있다. 또한 성과나 기관 객체도 역시 마찬가지이다.

5. 결 론

본 논문에서 제안하고 있는 온톨로지 기반의 국가R&D 지식맵은 사용자가 국가R&D 정보와 관련하여 관심있어 하는 객체와 주제를 중심으로 관련 정보를 손쉽게 추론하거나 다른 정보를 탐색할 수 있다는 점에서 의의가 있다

고 할 수 있다. 또한 온톨로지에 대한 전문가가 아니라도 데이터에 대한 의미관계를 쉽게 파악하고 사용자가 의도하는 질의나 관심정보를 쉽게 표현한다는 점에서 국가R&D 지식맵은 장점이 있다고 할 수 있다. 온톨로지를 표현하는 틀이나 방식은 여러 가지가 있으나 본 논문에서 제안하는 지식맵은 지식에 대한 정보를 표현하거나 관계를 표현하는 점에서 많은 장점이 있다.

향후, 국가R&D 주제분야와 객체들간의 관계를 고려하여 추론규칙을 더 확대하고, 데이터의 성능이나 사용자 로그를 분석하여 의미 없는 추론규칙들은 배제를 하고, 새로운 추론규칙을 추가하여 의미있고 다양한 검색결과를 얻을 수 있도록 할 계획이다.

참 고 문 헌

- 국가과학기술지식정보서비스 홈페이지. Retrieved from <http://www.ntis.go.kr/>
- 물리학연구정보센터(n.d.). 물리학지식지도. Retrieved from <http://km.icpr.or.kr/>
- 양명석, 최희석, 신수미, 한승우, 노현숙, 김재수 (2011.7.8). 온톨로지 기반의 국가R&D지식 네비게이션 서비스 구축에 관한 연구. 2011년 한국융합학회 하계학술대회 발표자료, 제주국제컨벤션센터, 제주.
- 화학공학연구정보센터(n.d.). 연구자네트워크 서비스.
Retrieved from <http://www.cheric.org/research/people/rns.php>
- Common European Research Information Format (2004).
Retrieved from <http://www.cordis.lu/cerif>
- McGuinness, D., & van Harmelen, F. (2004). OWL web ontology language overview. W3C Recommendation.
Retrieved from <http://www.w3.org/TR/2004/REC-owl-features-20040210/>
- Natalya F. Noy, & McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. Stanford, CA: Stanford University.
- WIKIPEDIA(n.d.). Knowledge map.
Retrieved from http://en.wiktionary.org/wiki/knowledge_map

• 국문 참고문헌에 대한 영문 표기 (English translation of references written in Korean)
--

- Chemical Engineering Research Information Center(n.d.). Research Network Service.
Retrieved from <http://www.cheric.org/research/people/rns.php>
- Information Center for Physics Research(n.d.). Knowledge Map of Physics.
Retrieved from <http://km.icpr.or.kr/>
- National Science & Technology Information Service Homepage(n.d.).
Retrieved from <http://www.ntis.go.kr/>
- Yang, Myung-Seok, Choi, Hee-Seok, Shin, Su-Mi, Han, Seung-Wu, Noh, Hyunsuk, & Kim, Jae-Soo (2011.7.8). A study on national R&D knowledge navigation service using ontology. Paper presented at the 2011 Summer Conference of Korean Society of Convergence, ICC Jeju, Jeju Special Self-Governing Province, Korea.