

Topic Map을 활용한 연구개발정보의 연계 모델 개발*

Development of Cross Reference R&D Information Model using Topic Map

김재성** · 윤종민***
Jae-Sung Kim · Chong-Min Yoon

차 례

- | | |
|------------------------|-----------------|
| 1. 서론 | 5. 연계 모델의 적용예제 |
| 2. 연구개발정보의 유형 및 연계 필요성 | 6. 결론 및 향후 연구과제 |
| 3. 선행연구 | · 참고문헌 |
| 4. 연구개발정보의 연계 모델 개발 | |

초 록

최근 국가적 필요성의 인식 하에 과학기술정보, 특히 연구개발정보의 종합적 수집·관리 및 유통을 위한 국가 과학기술 종합정보 시스템(NTIS)의 개발이 논의되고 있다. 국가 과학기술 종합정보 시스템은 연구개발 프로세스, 연구개발활동, 연구개발정보, 연구개발 데이터 레벨의 구성요소들 간의 유기적 연계를 그 근간으로 한다. 본 논문에서는 국가 과학기술 종합정보 시스템의 개발에 기반이 되는 연구개발정보들 간의 상호연계 모델을 제안한다. 연계를 위해 고려되는 연구개발정보는 연구과제정보, 연구인력정보, 연구성과정보로 연구개발활동에 있어 매우 중요한 정보라 할 수 있다. 제안된 연계 모델을 통해 연구개발정보들 간의 상호 참조 및 탐색이 가능하며, 다양한 관점에서 다양한 형태의 질의를 수행할 수 있다. 연계 모델은 ISO13250 표준인 XML Topic Map을 이용하여 개발되었으며, 예제를 통해 연계 모델의 우수성 및 향후 활용가능성을 살펴보았다.

키 워 드

국가 과학기술 종합정보 시스템, 연구개발정보, XML Topic Map, 연구개발 프로세스

*본 논문은 KOSTI Workshop 2004 발표논문을 수정 및 보완한 것임.

** 한국과학기술정보연구원 정책연구실 선임연구원
(Senior Researcher, Policy Study Dept., KISTI, jaesungkim@kisti.re.kr)

*** 한국과학기술정보연구원 정책연구실 책임연구원
(Principal Researcher, Policy Study Dept., KISTI, yoon1215@kisti.re.kr)

· 논문접수일자 : 2005년 9월 26일

· 게재확정일자 : 2005년 12월 7일

ABSTRACT

Recently, under the pressure of national necessity, the national S&T information system(NTIS) for collecting, managing and distributing various R&D information is under discussion. The close inter-connections and cooperations among elements in R&D process, activity, information and data levels are the basis of national S&T information system. In this paper, we propose a cross reference model of R&D information. The proposed model includes research project, person and output information which are essential in the R&D activities. By using the proposed model, a cross navigating and referencing among R&D information is possible. And various queries on the models from various viewpoints are also possible. XML Topic Map, ISO13250 Standard, is used for development of the proposed model. The efficiencies and further practical usages of proposed model are discussed by demonstrating the proposed model with an example.

KEYWORDS

National R&D Information System, R&D Information, XML Topic Map, R&D Process

1. 서 론

국가 연구개발정보(national R&D information)는 한 국가의 과학기술 경쟁력 평가를 위한 중요한 척도와 수단으로 인식되고 있다. 연구개발 인력, 연구개발 성과, 분야별 특허정보의 양과 그 질적 수준 등은 그 국가의 기술 경쟁력을 대변하고 있다고 여겨지고 있기 때문이다. 따라서 국가 과학기술 혁신과 발전의 핵심기반이 되는 국가 연구개발정보의 효율적 유통체제의 구축은 21세기 지식기반경제 시대의 국가경쟁력을 좌우하는 매우 중요한 사업이라 할 수 있다.

국가 연구개발정보의 효율적 유통체제 구축

을 위해서는 국내·외에서 발생하는 관련 연구개발정보를 신속하게 수집하여 이를 효율적으로 관리하고, 정보를 필요로 하는 국내 연구자들에게 적절히 제공하는 것이 중요하다. 이를 위해, 미국과 유럽 연합(EU)은 오래 전부터 국가 연구개발정보의 수집·관리 및 유통을 위한 종합 시스템을 개발하여 운영 중에 있는데, 미국의 RaDiUS(R&D in US)(RAND 2004)와 유럽 연합의 CRIS(Current Research Information System)(euroCRIS 2004)가 대표적인 예라 할 수 있다.

하지만, 우리나라는 국가 연구개발사업의 연구개발정보의 체계적인 수집·축적 및 관리·유통과 관련된 범 국가적 차원의 시스템

구축과, 이를 뒷받침하기 위한 제도적 근거가 매우 미흡한 실정이다. 각 부처가 과제관리 차원에서 국가 연구개발사업의 성과정보를 중심으로 관련 정보를 수집·관리하고 있으나, 범부처차원 더 나아가 국가적 차원의 종합적인 보존과 관리 및 유통이 제대로 이루어지지 않고 있을 뿐만 아니라, 수요자의 입장에서 원활한 접근 및 활용이 어려운 실정이다(윤종민 외 2002).

최근 들어, 국가혁신체제(National Innovation System)의 일환으로 국내에서 발생하는 각종 연구개발정보를 효율적으로 수집·관리·유통하기 위한 국가 과학기술 종합정보 시스템(NTIS, National Technical Information System)을 구상하여 한국과학기술정보연구원(KISTI), 한국과학기술기획평가원(KISTEP) 등 관련 연구기관들을 중심으로 시스템 개발에 노력을 기울이기 시작하고 있다.

이러한 국가 과학기술 종합정보 시스템의 개발에 있어서 가장 핵심적인 기능 중 하나가 바로 시스템의 기반이 되는 연구개발정보의 상호연계 모델(Cross Reference Model)의 개발이라 할 수 있겠다. 연구개발정보의 상호연계를 통해 다양한 관점에서의 연구개발정보와 지식의 탐색·추출 및 응용이 가능해지기 때문이다.

본 논문에서는, 연구개발정보들 중 특히 핵심이 되는 과제정보, 인력정보 및 연구 성과정보의 상호연계 모델을 제안한다. 상호연계 모델의 개발을 위해 ISO 13250 표준인 XML Topic Map을 활용하였으며, 간단한 예제를 통

해 개발된 연계 모델의 특징과 적용 가능성을 살펴보았다.

2. 연구개발정보의 유형 및 연계 필요성

국가 연구개발사업은 국가 전체의 과학기술 발전을 위하여 국가 예산 또는 공공 자금을 투자하여 일정한 목적에 따라 계획을 수립하고 추진하는 연구개발사업을 말한다(윤종민 외 2002). 일반적으로 연구개발 프로세스는 연구개발의 기획·조사 및 조정·선정이 이루어지는 연구기획 단계와 연구관리를 중심으로 직접적인 연구를 수행하게 되는 연구수행 단계, 그리고 연구개발 결과로 생성된 성과의 분석·평가 및 이전·확산이 이루어지는 연구성과관리 단계로 구성된다.

국가 연구개발정보는 이러한 국가 연구개발사업의 기획, 수행 및 관리의 연구개발 전 주기에서 생산·관리 및 활용되는 모든 정보를 뜻한다.

〈표 1〉은 연구개발 프로세스별 주요 연구개발정보의 유형을 나타내고 있다. 표에서의 각 정보들은 프로세스에서 활용·관리되는 대표적 주요정보를 나타낼 뿐이며, 특정정보가 특정 프로세스에 국한되어 활용됨을 뜻하지는 않는다. 즉, 특정 프로세스의 정보가 타 프로세스의 정보로도 활용되어 지기도 한다. 예를 들어, 학술연구정보나 특허정보 등은 연구수행 단계 뿐만 아니라 연구기획단계에서도 활용 될 수

〈표 1〉 연구개발 프로세스별 주요 연구개발정보의 유형

R&D 프로세스	연구기획	연구수행	성과관리
주요 정보유형	<ul style="list-style-type: none"> 기술·특허정보 기술·동향분석정보 산업·시장분석정보 기술수준정보 기술타당성평가정보 	<ul style="list-style-type: none"> 사업·과제정보 학술연구정보 연구성과정보 	<ul style="list-style-type: none"> R&D사업성정보 기술가치정보 기술활용성평가정보 기술이전정보 시장성평가정보 기술·기업평가정보
기반정보	<ul style="list-style-type: none"> 연구인력정보 연구장비정보 		

있는 것이다. 특히, 인력 및 연구장비정보는 연구개발 전 주기에 걸쳐 관리·활용되는 기반정보라 정의한다. 이점에 비추볼 때, 연구개발정보는 연구개발 프로세스별, 정보 및 데이터별로 상호연계되어 관리·운영되어야 한다.

그 동안 우리나라는 과학기술 연구개발에 대한 투자의 확대와 성과의 확산·확대를 위한 지속적 노력을 기울인 결과, 양적 측면에서는 어느 정도 그 규모를 갖추게 되었다. 그러나 1990년 이후 국가 연구개발사업의 추진체제와 내용이 다양화되고 확대·분산되는 과정에서 이를 관리하기 위한 시스템을 각 부처별 혹은 전담 기관별로 상호 호환성 없이 중복개발 및 운영됨에 따라 이와 관련된 정보·시스템의 국가적 차원의 종합적 관리가 불가능한 상황이다. 특히, 연구개발활동의 효율화를 직·간접적으로 지원하는 연구개발정보의 종합적 관리와 유통에 있어서는 여전히 문제점들을 안고 있는 실정이다.

서론에 언급한 바와 같이, 이러한 문제점들의 인식 하에 최근 우리나라에서도 국가적 차

원의 연구개발정보의 종합적 수집·관리 및 유통을 위한 종합정보 시스템의 개발이 시작되고 있는 단체에 있다. 즉, 그 동안부처별·기관별로 독립적·분산적으로 수집·관리·서비스되어 오던 연구개발정보의 유통체제를 일원화·종합화하고 하부 정보 단위에서는 각 정보들 간의 상호연계를 통하여 정보의 가치성과 활용성을 제고한다. 이를 통하여 국가 연구개발활동, 국가 연구개발 인프라 구축 등의 현황을 종합적으로 분석·평가함으로써 국가 연구개발의 전반적인 효율성 제고를 도모하고자 하고 있다. 따라서 종합정보 시스템 구축의 기반이 되는 연구개발정보들 간의 상호연계 모델의 개발이 필요하다.

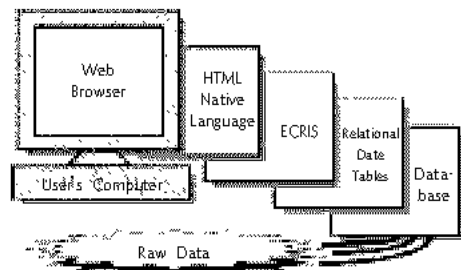
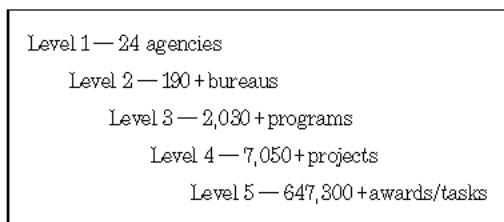
3. 선행연구

미국과 유럽을 중심으로 하는 선진국의 경우에는 1990년대 이후 국가적 연구개발사업의 종합적이고 체계적인 관리와 연구개발 수행의 효율성 제고를 위한 각종 종합정보 시스템을

구축·운영하고 있다. 미국의 경우, 50여 주정부 및 연방정부 산하의 개별부처에서 수행되는 연구개발사업에 대한 연구개발종합정보 시스템인 RaDiUS(Research and Development in United States)(RADN Corporation 2004) 시스템을 구축·운영해 왔다. RaDiUS 시스템은 연구개발과 관련된 기술의 벤치마킹(Benchmarking technologies), 이전가능 기술 선별(Identifying transferable technologies), 연구개발활동 현황파악(Profiling R&D activity) 등을 통해 연구개발사업의 체계적 관리와 연계·조정역할을 수행하고 있다. RaDiUS 시스템의 DB는 연방부처 수준의 연구개발활동으로부터 부처(agency) 내 실·국(bureau) 및 프로그램을 거쳐, 세부 프로젝트 및 수행과제까지 추적해 내려가는 종합적 정보 서비스를 제공하고 있다(과학기술부 2002). 이는 5개의 레벨로 구성된 데이터 모델(〈그림 1(a)〉)과 이를 상호연계하는 관계형 DB 및 Web 기반 정보의 검색을 위한 ECRIS 시스템(〈그림 1(b)〉)을 통해 연구개발정보의 종합적

관리와 정보 서비스를 제공하게 된다.

CRIS(Current Research Information System)는 유럽 내 각 국가·기관별 출연 연구개발정보를 담고 있는 연구개발정보 시스템의 종합적 연계를 통해 유럽 연합(EU)에서 출연되는 연구개발사업들을 종합적으로 관리코자 개발된 시스템이다. CRIS의 주된 목적은 진행 중인 프로젝트들에 대한 과학자들간의 커뮤니케이션을 향상시키는 것과, 국가 연구개발 프로그램의 관리자들에게 효과적인 정보기반을 제공하는 것으로(류범중, 최기석 2004), 연구개발사업, 기관, 연구자, 연구개발성과 등의 다양한 정보를 매우 포괄적으로 제공하고 있다. CRIS의 연구개발정보는 CERIF(Common European Research Information Format)(Asserson, Jeffery, and Lopatenko 2002)를 통해 정의되며, 이를 기반으로 각 국가 및 기관에 적합한 CRIS 시스템을 개발하여 종합적으로 연계한다. CERIF는 유럽전역의 연구개발정보를 포괄적으로 표현하는 공통의 언어로서, 메타데이터, 데이터베이스 스키마 및 시스템



(a) Data levels and number of records

(b) The Architecture of RaDiUS

〈그림 1〉 RaDiUS System (RAND 2004)

템개발과 필요한 각종 가이드라인을 담고 있으며, 이질적 정보자원을 포괄적으로 정의하는 Full CRIS Data Model과 데이터/시스템 간의 교환을 위한 Ex-change Data Model, 그리고 공통의 데이터를 정의하는 Metadata Data Model의 세 가지 데이터 모델로 구성되어 있다(Lopatenco 2001). 각각의 데이터 모델은 기본적으로 프로젝트, 인력, 조직의 3개 데이터를 중심의 다대다(many to many) 관계를 기반으로 관련 정보들이 상호연계적으로 구성되어 정의되어 시스템의 개발과 운영에 성공적으로 적용되고 있다.

4. 연구개발정보의 연계 모델 개발

본 논문에서는 연구과제정보(project information), 연구인력 정보(person informaton) 그리고 연구성과정보(output information)를 중심으로, 이들 세 가지 정보와 관련된 연구개발정보에 대한 연계 모델을 제안한다.

제안된 연구개발정보의 연계 모델을 언급하기 전에, 먼저 연구개발정보 모델의 기본적인 요구조건을 살펴해보도록 한다.

4.1 연구개발정보 모델의 요구조건

국가 과학기술 종합정보 시스템은 국가적 차원의 다양한 정보자원의 수집·구축·관리 및 유통은 물론, 구축된 정보자원의 실태, 활용

현황, 파급효과 등에 대한 상호연관적 종합분석과 평가를 수행하게 된다. 따라서 시스템의 핵심이 되는 연구개발정보 모델은 다음의 조건을 갖추어야 한다.

첫째, 확장성이 뛰어나야 한다. <표 1>에서 국가 과학기술 종합정보 시스템에서 고려하는 주요 연구개발정보유형을 제시하였으나, 국가적 필요성에 따라 대상 정보의 유형이 추가되거나 삭제되는 등의 변경사항이 발생 할 수 있다. 따라서 대상 연구개발정보 모델은 정보유형의 추가와 삭제와 같은 변경사항에 견고하고 효율적으로 대처할 수 있도록 설계되어야 한다.

둘째, 호환성을 갖추어야 한다. 국가 과학기술 종합정보 시스템은 18개 부처, 70여개 기관이 보유한 매우 다양한 정보자원/데이터베이스 및 시스템의 상호연계를 통한 종합정보 시스템의 구축에 그 기본방향을 두고 있다. 따라서 최소한 메타 정보/시스템 레벨에서의 용어(vocabulary), 데이터베이스 스키마(schema), 온톨로지(ontology) 등의 정의 및 이를 근간으로 하는 시스템 개발에 있어서의 표준화가 요구된다. 따라서, 연구개발정보 모델은 제시된 표준적 지침 및 가이드라인을 충실히 반영할 수 있어야 한다.

셋째, 다양한 질의의 수행이 가능해야 한다. 앞서 언급한 바와 같이, 국가 과학기술 종합정보 시스템의 가장 중요한 역할 중 하나가 바로 국가적 차원의 연구개발활동의 종합적 상황판기능의 수행이다. 이러한 기능의 수행을 위해서는 주어진 정보자원으로부터 다양하고 의미

있는 결과를 이끌어 내기 위한 유연한 질의 (flexible query)를 수행할 수 있어야한다. 이러한 기능은 국가적 차원의 연구개발정보자원, 활동 등의 종합적 분석·평가의 기본이 되기 때문이다. 따라서 연구개발정보 모델은 다양한 질의·응답이 가능하도록 설계 되어야 한다.

넷째, 다양한 관점(viewpoint)에서의 정보의 표현이 가능해야 한다. 국가 과학기술 종합 정보 시스템은 국내 연구개발자, 정부부처 정책입안자는 물론, 일반국민을 그 이용대상자로 삼고 있다. 이러한 다양한 이용자들은 서로 다른 관점에서 국가 과학기술 종합정보 시스템을 활용하게 된다. 예를 들어, 연구개발자는 논문, 특허 등의 특정분야의 연구결과정보를 중심으로 관련 과제정보, 과제수행자 등의 정보를 탐

색하는 반면, 정책입안자는 국가적 연구사업·과제를 중심으로 과제로부터 파생된 연구개발 결과와 과제수행자 등의 정보를 탐색할 수 있다. 따라서 연구개발정보 모델은 이러한 다양한 이용자들의 다양한 관점에서의 탐색이 가능하도록 정보의 정의·기술이 가능해야 한다.

4.2 연구개발정보 연계 모델

본 논문에서 제안되는 연구개발정보 연계 모델은 정보의 성격에 따라 <표 2>와 같이 두 가지 레벨로 구분되어 정의된다. 제안되는 연구개발정보 연계 모델의 핵심 정보인 연구과제 정보, 연구인력정보, 연구성과정보의 1st 레벨 정보와 1st 레벨 정보와 개념적 의미관계를 가

<표 2> 레벨에 따른 연구개발정보의 구분

Information Types	
1 st Level	2 nd Level
연구과제정보	행정부처, 발주·수행기관, 사업유형 정보
연구인력정보	사업책임자, 참여연구원 정보
연구성과정보	결과보고서, 논문, 특허, 장비, 수치·사실정보



<그림 2> The cross reference model of R&D information: 1st Level Model

지는 2nd 레벨 정보로 구분된다. 예를 들어, 연구성과정보는 1st 레벨 정보이며, 연구성과정보의 하위개념인 연구결과보고서, 논문, 특허 등은 2nd 레벨 정보로 표현된다.

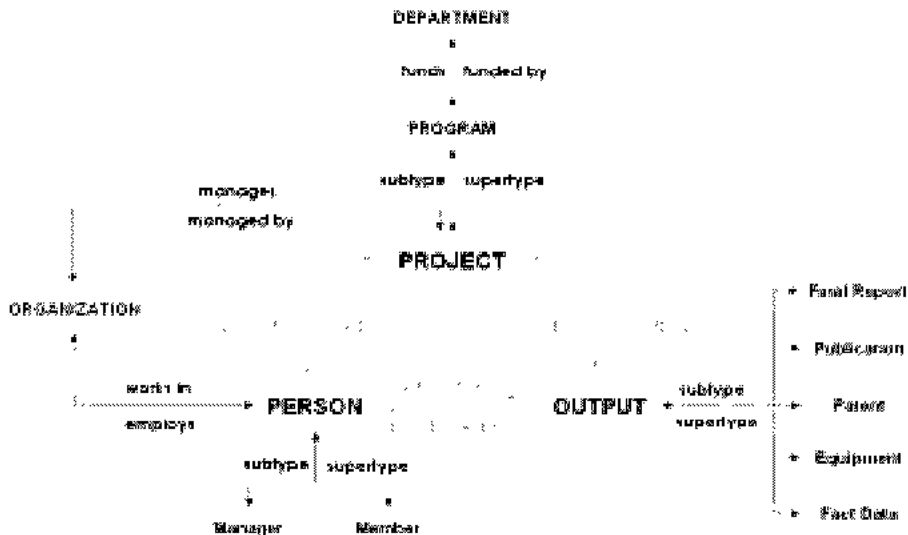
제안된 연계 모델은 각 레벨 내의 정보자원들 간의 상호 연관관계 및 1st 정보자원과 2nd 정보자원들 간의 상호연관 관계의 기술을 통해 정의된다.

〈그림 2〉는 이들 1st 레벨 정보들 간의 상호 연계 모델을 보이고 있다. 그림에서 PROJECT는 연구과제 정보를, PERSON은 연구인력정보를, OUTPUT은 연구성과정보를 나타낸다.

각 정보는 Subject로서 사각형으로 묘사되었으며, 각 정보들 간의 연관관계는 화살표와 화살표 위의 술어(predicate)로 묘사되었다. 여기서 1st 레벨 정보들 간에는 상호참조 관계

(cross reference relationship)를 가진다. 즉, 연구과제(PROJECT)는 연구인력(PERSON)에 의해 수행되며(conducted by), 연구인력(PERSON)은 연구과제(PROJECT)를 수행(conducts)하는 관계를 가진다. 마찬가지로, 연구인력(PERSON)은 연구성과(OUTPUT)를 생산(produces)하며, 연구성과(OUTPUT)는 연구인력(PERSON)에 의해 생산되어지는 관계(produced by)를 가진다. 또한, 연구과제(PROJECT)는 연구성과(OUTPUT)를 가지며(has), 연구성과(OUTPUT)는 연구과제(PROJECT)의 결과(result of)라는 관계를 가진다.

〈그림 3〉은 2nd 레벨에서의 연구개발정보의 연계 모델을 보이고 있다. 2nd 레벨 연계 모델에서는 PROJECT, PERSON, OUTPUT의 상



〈그림 3〉 The cross reference model of R&D information: 2nd Level Model

위·하위 정보 및 주요 관련 정보들 간의 상호연계관계를 표현한다. 그림에서 알 수 있듯이, 연구과제(PROJECT)는 상위개념(supertype)인 연구사업(PROGRAM)의 하위개념(subtype)이며, 연구사업(PROGRAM)은 주관행정부처(DEPARTMENT)로부터 연구자금을 지원받는다(funded by) 관계를 가진다. 연구과제와 연구사업, 주관행정부처 간의 그 역의 관계도 성립한다. 연구인력(PERSON)은 책임연구자(Manager)와 참여연구원(Member)으로 나누어지며 이들 간은 상위·하위 관계를 가진다. 마찬가지로, 연구성과(OUTPUT)는 연구결과보고서(Final Report), 연구논문(Publication), 특허(Patent), 연구장비(Equipment), 실험결과 데이터(Fact Data)의 5가지 타입의 성과정보로 나누어지며, 연구성과(output)와는 상위·하위 관계를 가진다. 연구수행기관(ORGANIZATION)은 연구과제(PROJECT)를 관리·감독(supervises)하며, 연구를 수행하는 연구인력(PERSON)을 고용(employs)한다. 역으로, 연구과제(PROJECT)는 연구수행기관(ORGANIZATION)에 의해 관리·감독되며(supervised by), 연구인력(PERSON)은 연구기관(ORGANIZATION)에서 연구를 수행(works in)하는 관계를 가진다. 위에서 설명된 연구개발모델은 연구개발 프로세스 상에서 가장 기본이 되고, 중요한 정보에 대한 기초적인 관계들만을 정의한 것이다. 실질적인 연구개발 정보의 상호연계 모델에는 매우 다양한 수많은 종류의 정보들 간의 상호연계 관계가 기술되어

질 것이다.

4.3 XML Topic Map을 활용한 연계 모델 정의

제안된 연구개발정보의 상호연계 모델은 ISO13250 표준인 XML Topic Map(XTM) (Pepper 2002a; Pepper 2002b)을 활용하여 개발되었다. XML Topic Map은 다양한 정보 자원들과 그 정보들 간의 관계를 기술함으로써 정보의 연계·공유를 통한 지식의 표현 및 관리를 가능케 하는 국제표준정보 기술언어이다. Topic Map은 토픽(Topic)과 토픽 간의 관계(Association), 그리고 토픽과 관련된 리소스(Occurrence)의 3가지 기본 개념으로 구성된다.

〈그림 4〉는 Topic Map으로 기술된 IST 레벨 연계 모델의 일부를 보이고 있다. 그림에서는 PROJECT, PERSON, OUTPUT 토픽과 이들 간의 관계를 기술하기 위한 conducts, produces, has 관계 토픽을 정의하고 있다. 정의된 토픽들을 바탕으로 토픽의 객체(instance)를 생성하여 객체들 간의 관계를 〈Association〉 Tag를 이용하여 정의해 주게 된다.

제안된 연계 모델은 〈그림 5〉와 같은 구조로 개발되었다. 그림에 나타난 바와 같이, 각 정보에 대한 Topic Map 모델들을 작성하고, 작성된 모델을 종합(merge)하는 RNDINFO.XTM 모델을 만든다. 이때 각 Topic Map 모델에서

```

<?xml version="1.0" encoding="EUC-KR"?>
<topicMap xmlns="http://www.topicmaps.org/xtm/1.0"
  xmlns:xlink="http://www.w3.org/1999/xlink" id="
  mndinfo-tm">

<!-- R&D information cross referece model -->
<!-- Topic: PROJECT, PERSON, OUTPUT -->
<!-- Association: conducts, produces, has -->

<topic id="project">
  <subjectIdentity>
    <subjectIndicatorRef
      xlink:href="http://psi.ontopia.net/xtm/topic/project"/>
  </subjectIdentity>
  <baseName>
    <baseNameString>Project</baseNameString>
  </baseName>
</topic>

<topic id="person">
  <subjectIdentity>
    <subjectIndicatorRef
      xlink:href="http://psi.ontopia.net/xtm/topic/person"/>
  </subjectIdentity>
  <baseName>
    <baseNameString>Person</baseNameString>
  </baseName>
</topic>

<topic id="output">
  <subjectIdentity>
    <subjectIndicatorRef
      xlink:href="http://psi.ontopia.net/xtm/topic/output"/>
  </subjectIdentity>
  <baseName>
    <baseNameString>OUTPUT</baseNameString>
  </baseName>
</topic>

  <topic id="conducts">
    <subjectIdentity>
      <subjectIndicatorRef xlink:href="#conducts"/>
    </subjectIdentity>
    <baseName>
      <baseNameString>Conducts</baseNameString>
    </baseName>
    <baseName>
      <scope><topicRef xlink:href="#project"/></scope>
      <baseNameString>Conducted by</baseNameString>
    </baseName>
  </topic>

  <topic id="produces">
    <subjectIdentity>
      <subjectIndicatorRef xlink:href="#produces"/>
    </subjectIdentity>
    <baseName>
      <baseNameString>Produces</baseNameString>
    </baseName>
    <baseName>
      <scope><topicRef xlink:href="#output"/></scope>
      <baseNameString>Produced by</baseNameString>
    </baseName>
  </topic>

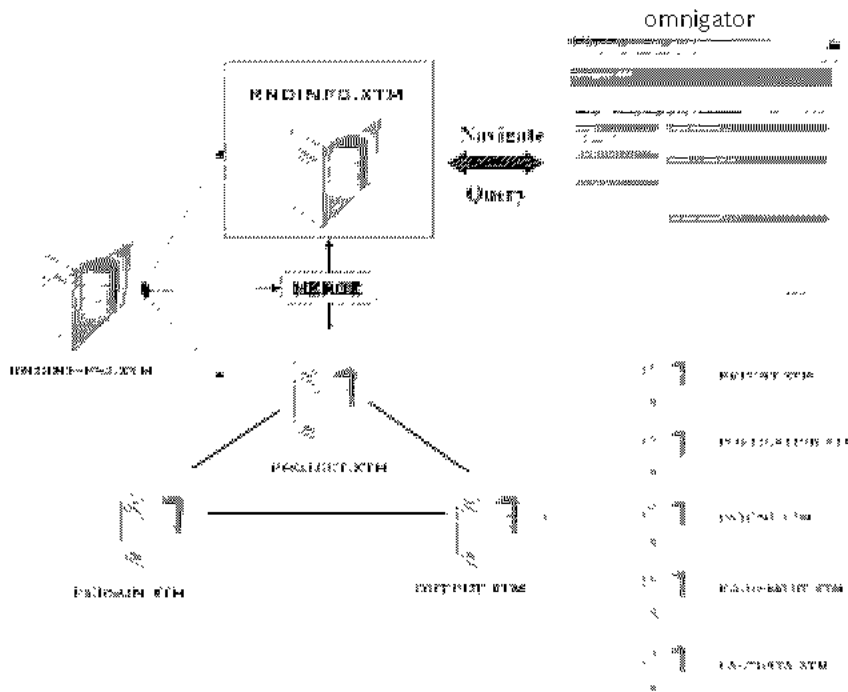
  <topic id="has">
    <subjectIdentity>
      <subjectIndicatorRef xlink:href="#has"/>
    </subjectIdentity>
    <baseName>
      <baseNameString>Has</baseNameString>
    </baseName>
    <baseName>
      <scope><topicRef xlink:href="#output"/></scope>
      <baseNameString>Result of</baseNameString>
    </baseName>
  </topic>

```

〈그림 4〉 The description of 1st level cross-reference model using XML Topic Map

활용되는 각종 Topic, Association 및 Occurrence의 표준적 타입을 정의한

RNDINFO PSI.XTM을 참조하게 된다. Topic Map에서 제공되는 Subject



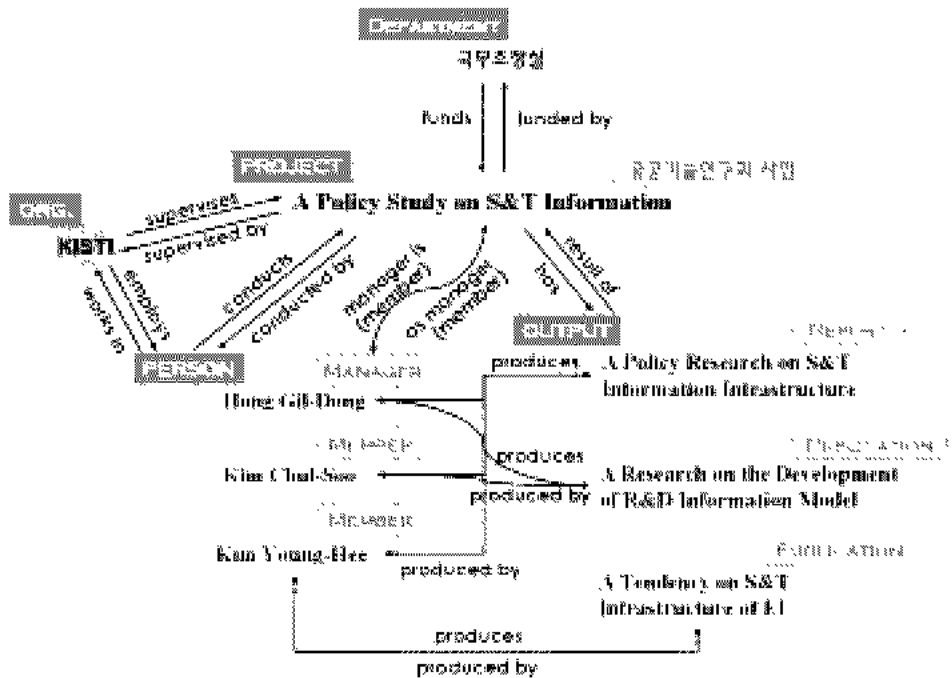
〈그림 5〉 An development architecture of R&D information Topic Map

Identifier(URI)를 활용하여 표준적 타입들의 집합체인 RNDINFO PSI를 정의함으로써, 여러 Topic Map 모델의 결합(merge) 시 유사개념의 Topic, Association, Occurrence 간의 충돌을 방지할 수 있다. 얻어진 RNDINFO 모델을 DB화하여 인덱싱(Indexing)한 후, 온토피아(Ontopia)(Ontopia 2004a)에서 제공되는 Free S/W인 Omnigator를 활용하여 정의된 모델을 탐색(navigate)하고, 다양한 형태의 질의(query)를 수행하게 된다.

5. 연계 모델의 적용 예제

프로젝트 “A policy study on S&T infor-

mation”이 있고, 이 프로젝트는 홍길동, 김철수, 김영희라는 사람들에 의해 수행되었다. 이 프로젝트의 연구책임자는 홍길동이며, 김철수와 김영희는 참여연구원으로서 프로젝트에 참여하고 있다. 이 프로젝트는 공공기술연구회 사업의 일환이며, 국무조정실로부터 연구자금을 지원받았으며, 프로젝트의 주관수행 연구기관은 한국과학기술정보연구원(KISTI)이다. 프로젝트의 수행결과, 1개의 최종연구보고서와 2개의 연구논문이 생산되었다. 이를 제안된 연구개발 연계 모델로 적용하여 도식화하면 〈그림 6〉과 같다.



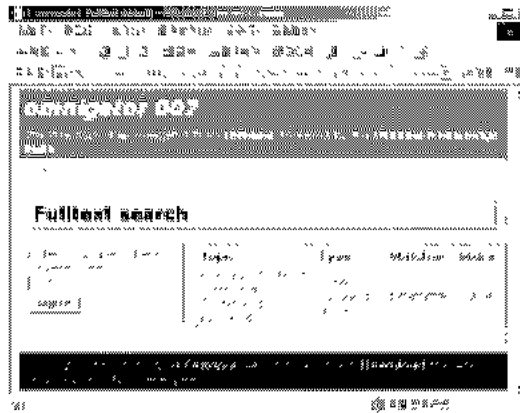
〈그림 6〉 An application of example to the cross-reference R&D information model

5.1 연계 모델 Navigation

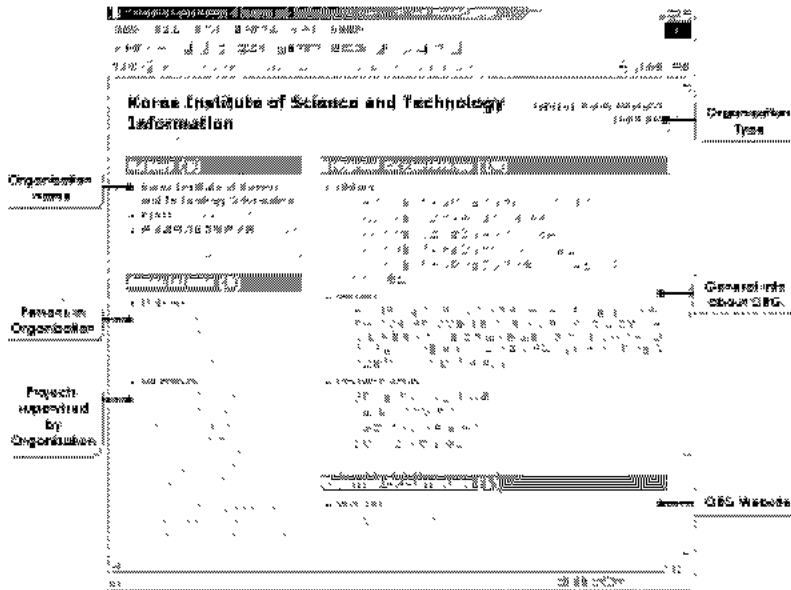
Topic Map Navigator인 Omigator를 활용함으로써 정의된 연구개발정보들이 어떻게 상호연계되어 있는 지 살펴 볼 수 있다. 〈그림 7〉은 위 예를 〈그림 5〉의 구성에 따라 Topic Map 모델을 작성한 후, 이를 Omigator를 활용하여 탐색하는 화면을 보인다. 〈그림 7(a)〉는 정의된 Topic Map에 대한 Omigator의 첫 화면에서 키워드 “KISTI”에 대한 검색(full text search)을 수행한 결과이다. 〈그림 7(b)〉는 검색결과 얻어진 Topic(Korean Institute of Science and Technology Information)에

대한 정보를 보이고 있다.

〈그림 7(b)〉를 살펴보면 Korean Institute of Science and Technology Information은 공공연구기관(public research institute)으로 한글(in korean) 명은 “한국과학기술정보연구원”이며, 약명(short name)은 “KISTI”임을 알 수 있다. 또한, 한국과학기술정보연구원은 Hong Gildong, Kim Chulsoo, Kim Duksoo, Kim YoungHee라는 사람들을 고용(employs)하고 있으며 “A Policy Study on S&T Information”, “Bioinformatics System Development”, “Construction of Computing Grid by Supercomputer Resources” 등의 프



(a) Query result : keyword "KISTI"



(b) Information on KISTI in the model

<그림 7> Navigating the RNDINFO model with Omnigator

로젝트 들을 관리(supervises) 혹은 관리해 왔다는 것을 알 수 있다. 이밖에도 한국과학기술정보연구원의 연혁(history), 미션(mission), 연구영역(research area) 및 웹 사이트 정보를

알 수 있다.

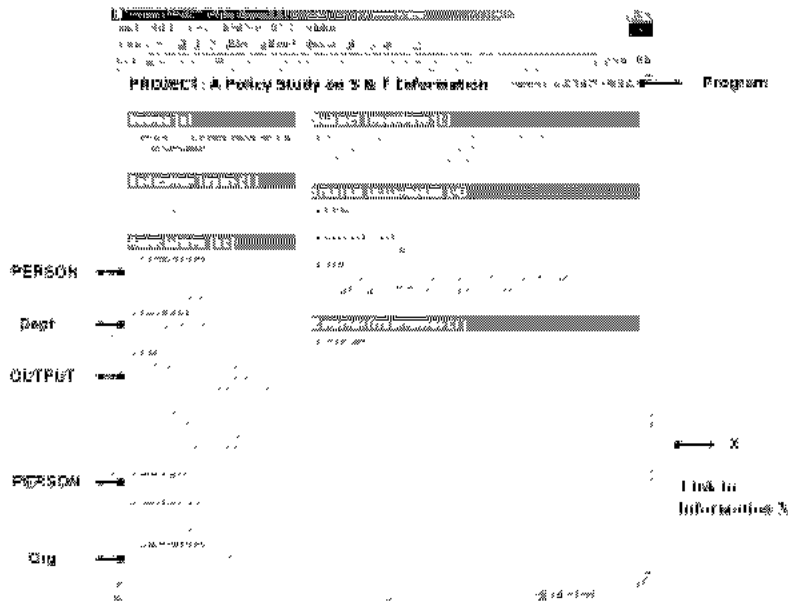
<그림 8>은 <그림 7(b)>의 KISTI에서 관리(supervises)하는 프로젝트 "A Policy on S&T Information"의 정보를 나타내는 화면이

다. 그림에서 알 수 있듯이, 본 프로젝트는 공공기술연구회사업의 일환으로 수행되었으며, Hong Gildong, Kim Chulsoo, Kim YoungHee에 의해 수행되었는데(conducted by), Hong Gildong은 프로젝트 책임자(manager)이며, Kim Chulsoo와 Kim YoungHee는 프로젝트의 참여연구원(member)으로서의 역할을 수행하였음을 알 수 있다. 이 프로젝트는 국무조정실(The Office for Government Policy Coordination)로부터 자금지원을 받았으며(funded by), 하나의 연구결과보고서("A Policy Research on National S&T Information Infrastructure")와 2개의 연구논문("A Research on the Development of R&D

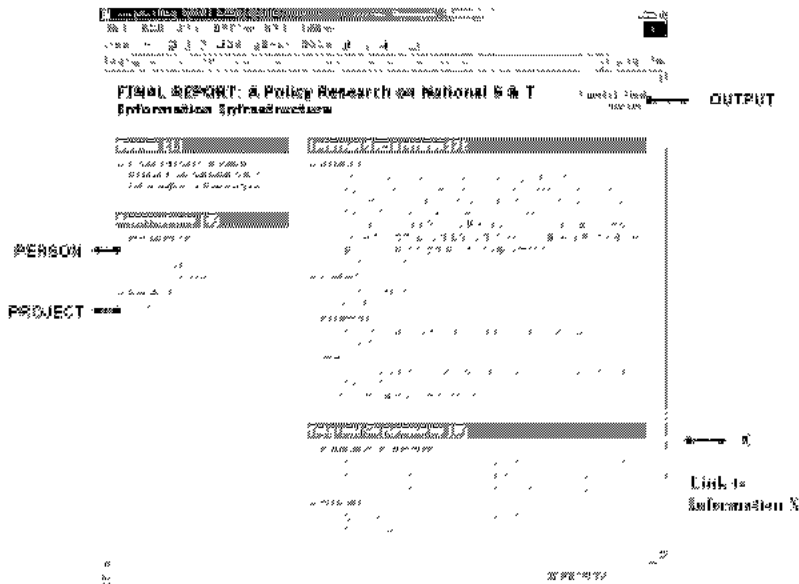
Information Model", "A Tendency on S&T Information Infrastructure of EU")를 가지고(has) 있다는 것을 알 수 있다. 이 프로젝트는 KISTI에 의해 관리·감독되었다(supervised by). 이 밖에도 연구결과보고서 제목, 연구기간, 연구사업비 및 프로젝트 관련 정보를 담고 있는 웹 사이트 정보를 살펴 볼 수 있다.

그림에 나타난 바와 같이, 프로젝트 정보로부터 연구인력정보(PERSON), 행정부처정보(Dept.), 연구성과정보(OUTPUT), 기관정보(Org.), 연구사업정보(Program) 정보로의 연계/탐색이 가능하다.

〈그림 9〉는 〈그림 8〉의 프로젝트가 가지고 있는(has) "FINAL REPORT: A Policy



〈그림 8〉 Information on the PROJECT : "A Policy Study on S&T Information"



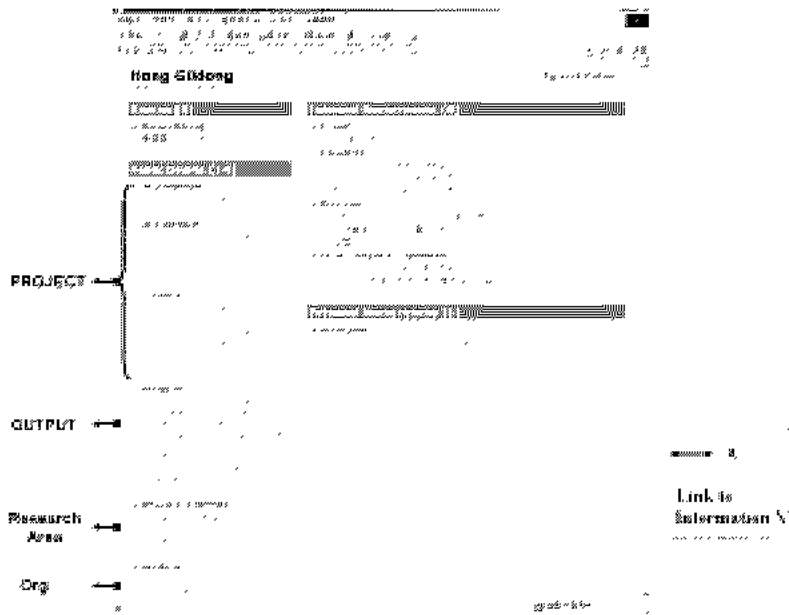
〈그림 9〉 Information on the OUTPUT : “A Policy Research on National S&T Information Infrastructure”

Research on National S&T Information Infrastructure”의 정보를 나타내고 있다. 그림에서 알 수 있듯이, 이 연구결과보고서는 Hong Gildong, Kim Chulsoo, Kim YoungHee라는 사람들에 의해 생산되었으며 (produced by), 프로젝트 “A Policy Study on S&T Information”의 결과이다(“PROJECT: A Policy Study on S&T Information”을 클릭하면 〈그림 8〉의 화면으로 돌아가게 된다). 이밖에도 결과보고서의 제목, 초록, 키워드, 연락처 및 원문정보(publication on the web), 관련정보를 담고 있는 웹 사이트 정보를 살펴볼 수 있다. 마찬가지로, 연구결과보고서 정보로부터 연구인력정보(Person), 프로젝트 정보(Project)로의 연계/탐색이 가능하다.

지금까지 기관정보 과제정보 연구결과보고

서 정보 간의 연계를 예를 통해 살펴보았다. 마지막으로 인력정보와의 연계를 살펴볼 것이다.

〈그림 10〉은 〈그림 9〉의 Hong Gildong(korean: 홍길동)이라는 연구자의 정보를 보이고 있다. 그림에 따르면 홍길동은 총 3개의 프로젝트를 수행(conducts)하였는데, 한 번은 “A Policy Study on S&T Information” 프로젝트에서 책임자(as a manager)로 참여하였고, “Bioinformatics System Development”, “Implementation of Domestic S&T Information Service System” 프로젝트들에는 참여연구원(as a member)으로 참여하였다. 또한, 홍길동은 1개의 연구결과보고서와 2개의 연구논문을 생산(produces)하였다. 홍길동은 Data Modeling, Information Management 등의 연구분야에 관심이 있으며, 한국



〈그림 10〉 Information on the PERSON : "Hong Gildong"

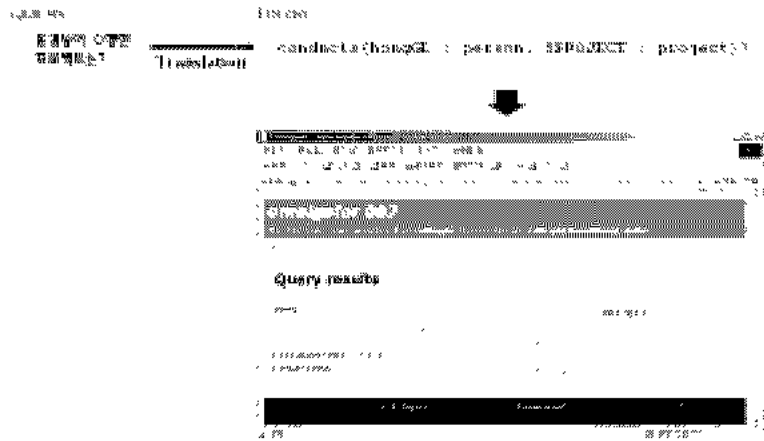
과학기술정보연구원에서 근무한다는 것을 알 수 있다. 이 밖에도 홍길동에 대한 신상사항, 업무이력, 연락처 및 홈페이지 정보를 살펴 볼 수 있다. 마찬가지로, 연구인력정보로부터 프로젝트 정보(Project), 연구성과정보(Output), 기관정보(Org.) 및 연구영역정보(Research Area)로의 연계/탐색이 가능하다.

5.2 연계 모델에 대한 Query

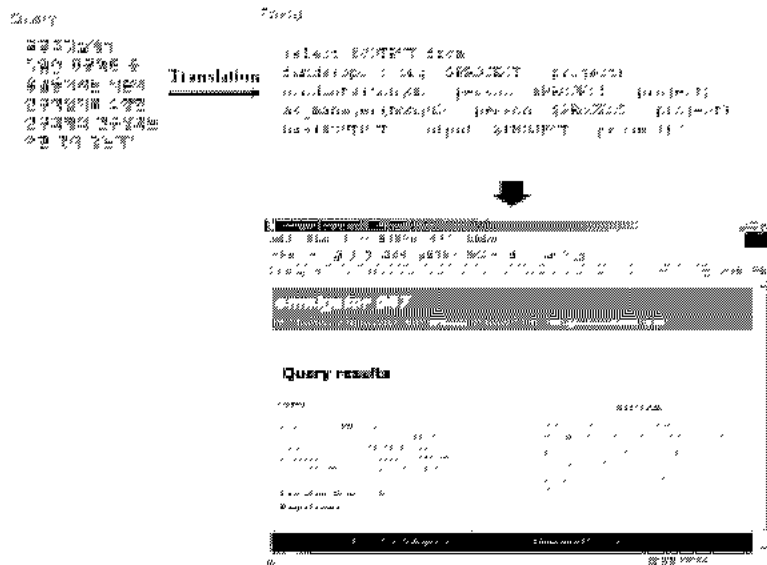
국가 연구개발종합정보 시스템에서는 다양한 관점에서 다양한 형식의 질의의 수행이 필요하다. 예를 들어, “국무조정실에서 지원한 연구과제는 어떤 것이 있는가?”에서부터 “국무조정실이 지원한 연구과제들 중, 한국과학기술정

보연구원에서 수행·관리한 과제연구과제에서 홍길동이라는 사람이 연구책임자의 역할로서 수행한 연구과제의 연구성과는 어떠한 것이 있는가?”라는 다소 복잡한 형태의 질의까지도 지원이 가능해야 할 것이다.

본 논문에서 제안된 RNDINFO 모델을 이용할 경우 위에서 언급된 복잡한 형태의 질의의 수행이 가능하다. Topic Map에 대한 질의는 Ontopia에서 제공하는 TMQL(Topic Map Query Language)인 Tolog를 활용하였다. Tolog는 Logic 기반 질의언어(Logic based query language)로서, 인공지능언어인 Prolog와 비슷한 문법을 가지며 질의에 대해 참(true)인 Fact의 리스트를 질의결과로 돌려준다. Tolog에 대한 구체적인 내용은 참고문헌



(a) A query example



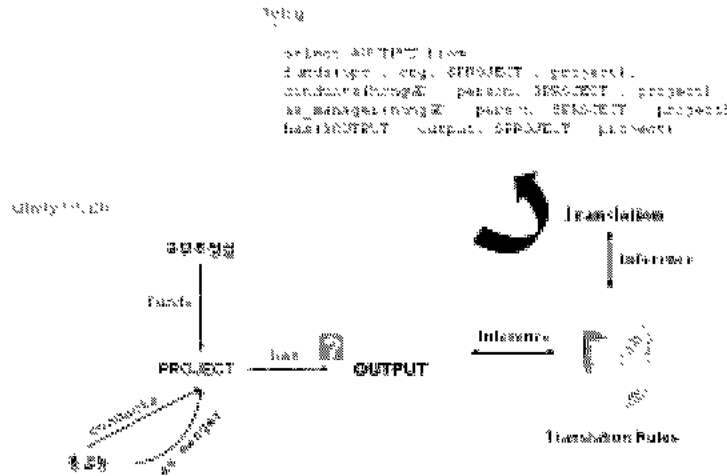
(b) A complex query example

〈그림 11〉 Examples of queries on the RNDINFO model

(Ontopia 2004b)을 참고하기 바란다. 〈그림 11〉은 Tolog를 활용하여 RNDINFO 모델에 대한 질의의 수행례를 보이고 있다.

사용자의 질의를 Tolog의 Syntax에 따라

Tolog 질의문으로 변환하고 이를 RNDINFO 모델에 전달하여 그 값이 참인 Fact (PROJECT)들의 리스트가 질의 결과로 도출되었다. 본 논문에서 제안된 RNDINFO 모델을



〈그림 12〉 The concept on the translation of graph query to Tolog query

활용할 경우, 예제에서 제시된 질의 외에도 매우 다양한 형태의 질의의 수행이 가능하다. Tolog는 RNDINFO 모델에서 사용되는 Topic, Association, Occurrence 등에 조합으로 질의문을 만듦으로 이들의 조합으로 수백, 수천 가지의 다양한 형태의 질의어를 만들어 낼 수 있는 것이다.

그러나 〈그림 11〉에 나타난 바와 같이 사용자의 임의의 문자 질의어를 Tolog 질의어로 변환하는 것은 매우 어려운 문제이다. 현재로서는 이러한 문제의 근본적인 해결은 불가능하다고 판단되며, 문자열 질의와 비슷한 효과를 달성하기 위해 사용자의 편의성을 고려한 질의 인터페이스(query interface)의 개발이 필요하다. 이 경우 그래프를 활용한 질의 인터페이스를 예를 들 수 있다. 즉, 〈그림 12〉와 같이 사용자가 그래프 형태의 질의문을 작성하고 시스템

이 이를 정해진 룰에 따라 해석하여 Tolog 질의문으로 자동 변환하여 질의를 수행하는 것이다. 이러한 방법을 통해, 사용자는 쉬운 방법으로 다양한 형태의 질의를 수행할 수 있게 될 것이다.

6. 결론 및 향후 연구과제

본 연구에서는 연구과제정보, 연구인력정보, 연구성과정보로 구성되는 연구개발정보의 상호연계 모델을 제안하였다. 연계 모델의 개발에 XML Topic Map을 활용하였으며, 예제를 통해 제안된 모델의 적용 가능성과 우수성을 보였다. 제안된 연계 모델은 다음과 같은 장점을 가지고 있다.

첫째, 연계 모델의 확장성이 뛰어나다. 본 논문에서는 연구과제정보/인력정보/성과정보

의 세 가지 주요 정보만을 대상으로 연계 모델을 정의 하였으나, 정보유형(topic) 및 정보들 간의 연관관계(association)의 추가 및 삭제를 통해 기술정보/기술이전정보/동향정보 등과 같은 연구개발정보를 얼마든지 추가·삭제 할 수 있다. 또한, 제안된 모델을 Topic Map을 활용하여 정보와 정보간의 연관관계를 매우 자연스러운 방식으로 정의함으로써 정보 모델에 대한 가독력과 구현의 수월성이 높으며, 향후 시맨틱 웹(semantic web)을 지향하는 지식기반 연구개발정보 시스템으로의 확장이 가능하다.

둘째, 5.2절에서 보인 바와 같이 다양한 형태의 질의를 수행할 수 있다. 이를 바탕으로 주어진 정보에 대한 다양한 분석과 평가 등의 활동을 지원함으로써, 국가 과학기술 종합정보 시스템의 주요 기능인 종합상황판 지원을 위한 하부 정보 모델로 활용될 수 있을 것으로 판단된다.

셋째, 다양한 관점에서 정보의 탐색이 가능하다. 5.1절의 탐색 예에서 보인 바와 같이, 이용자는 특정 정보를 중심으로 그와 관련된 다양한 정보를 살펴볼 수 있다. 물론, 다양한 관점에서의 정보의 탐색을 위해서는 관련 정보들의 관점별 표현을 위한 용어, 연관관계 등에 대한 추가적 정의와 구현이 필요하다. 이러한 기능의 구현은 연구개발정보 및 시스템의 표준화 작업과 병행하여 추진되어야 하며, 이는 국가적 차원에서 정부부처, 관련 연구기관, 연구자 등의 이해 당사자들 간의 구체적인 지속적인 논의를 통해 해결될 수 있으리라 판단된다.

논문에서 제안된 연구개발정보 연계 모델은 국가 과학기술 종합정보 시스템의 기반이 되는 연계 모델의 기본적 프레임워크에 해당된다. 실제로 본 시스템의 구현에 있어서는 더 많은 연구개발정보자원들(Topics)의 정의와 이들 정보 자원들 간의 다양한 연관관계들(Associations)을 규정하여 정의해야 할 것이며 이에 따른 연계 모델의 구조 및 계층 등의 확장에 대한 지속적 연구가 필요하다. 이를 통해 연구사업의 기획·선정에서부터 최종성과의 피드백에 이르는 연구개발 전주기에 걸쳐, 각종 정보 및 연구개발 인프라를 목표지향적으로 제공할 수 있는 종합적 정보유통체제의 구축을 기대할 수 있을 것이다.

참고문헌

- 과학기술부, 2002, 연구개발지식의 종합관리 시스템 구축방안 연구. 『과학기술부, 정책연구 2002-28』.
- 류범중, 최기석, 2004, 국가 R&D 지식정보관리 시스템 구축에 관한 연구 연구기획 및 관리를 중심으로. 『한국문헌정보학회지』, 38(1): 281-301.
- 윤종민, 정용일, 이성호, 이형진, 2002, 국가연구개발보고서 정보의 유통촉진을 위한 제도 연구. 『The 7th Workshop on Korean Infrastructure for Science and Technology Information』, 21-43.

- 한국과학기술기획평가원. 2001. 『국가 연구개발사업 통합관리 시스템 개발』. 서울: 한국과학기술평가원.
- 한국과학기술기획평가원. 2004. 『2004년도 정부연구개발예산 현황』. 서울: 한국과학기술평가원.
- Asserson, A., K. G. Jeffergy, and A. Lopatenko. 2002. "CERIF: Past, Present and Future: An Overview." *6th International Conference on Current Research Information Systems*, August 29 - 31, Germany
- euroCRIS. 2004. *euroCRIS in a nutshell*, [cited 2005, 8, 10].
<<http://www.eurocris.org/en/eurocris/>>.
- Fossum, D. 2004. *Identifying Federally Funded Research and Development on Information Technology*, RAND Corporation.
- ISO. 2002. *ISO/IEC 13250 Topic Maps: Second Edition*, ISO.
- Lopatenko, A. 2001. "Information Retrieval in Current Research Information Systems," *Workshop on Knowledge Markup and Semantic Annotation at K CAP 2001*, Oct. 21 - 23, Canada.
- Ontopia. 2004a. *The Ontopia Omnigator*, [cited 2005, 8, 10].
<<http://www.ontopia.net/omnigator/docs/navigator/userguide.html>>.
- Ontopia. 2004b. *tolog 2.0 : language tutorial*, [cited 2005, 9, 5].
<<http://www.ontopia.net/omnigator/docs/query/tutorial.html>>.
- Pepper, S. 2002a. *XML Topic Map*, [cited 2004, 10, 14].
<<http://www.topicmaps.org/xtm/1.0/>>.
- Pepper, S. 2002b. *The TAO of Topic Map: Finding the Way in the Age of Infogult*, [cited 2004, 10, 14].
<<http://www.ontopia.net/topicmaps/materials/tao.html>>.
- RAND. 2004. *What is RaDiUS*, [cited 2005, 8, 3].
<http://radius.rand.org/radius/radius_info.html>.