

메타데이터 기반의 연구성과정보 검색시스템의 개념적 설계 *

Conceptual Design of Metadata based Research Results Information Retrieval System

박 동 진** · 이 상 태*** · 최 기 석****

Dong-Jin Park · Sang-Tae Lee · Ki-Suk Choi

차 례

1. 서 론	4. 메타데이터 기반 시스템의 개념적
2. 연구의 배경	구조
3. 상호운용성 제고를 위한 시맨틱 기술	5. 결 론
의 적용기법	· 참고문헌

초 록

최근 연구성과정보의 원활한 교환 및 유통은 연구생산성 향상에 가장 핵심적인 요인으로 인식되고 있다. 그래서 국가적으로 그리고 각 연구기관에서는 연구성과정보 서비스를 위한 시스템을 계획하거나 구축 중에 있다. 그러나 연구정보들은 서로 다른 형태로 분산되어 있으며, 연구정보를 기술하는 방식에 있어서도 기관별 그리고 연구자 개인별로 많은 차이가 있어서 연구성과정보의 통합에 있어서 많은 제약이 있다. 따라서 본 연구는 시맨틱 웹 기술을 도입하여 연구성과정보의 메타데이터에 대한 상호운용성을 제고하는 방안을 제시함으로써 궁극적으로 연구성과정보의 통합을 지원하고자 한다. 특히 과학기술정보표준화위원회에서 제시한 연구성과정보 메타데이터 표준(안)을 기초

*본 연구는 "KOSTI 2005 한국과학기술정보인프라 워크숍" 학술발표대회에서 우수논문으로 선정된 논문을 수정·보완한 것임

**공주대학교, 산업시스템공학과 부교수
(Associate Professor, Dept. of Industrial and Systems Engineering, Kongju National University., mispdj@kongju.ac.kr)

*** 한국표준과학연구원 지식정보그룹 그룹장
(Team Leader, Technical Information Team, Korea Research Institute of Standards and Science., stlee@kriss.re.kr)

**** 한국과학기술정보연구원 R&D 시스템연구팀 팀장
(Team Leader, National R&D System Development Team, Korea Institute of Science and Technology Information, choi@kisti.re.kr)

· 논문접수일자 : 2006년 2월 19일

· 게재확정일자 : 2006년 5월 11일

로 개발되는 온톨로지는 용어들에 대한 구조적인 그리고 의미적인 통합을 가능하게 하며, RDF/RDFS를 이용한 메타데이터 지식표현은 정보 서비스의 검색과 유통에 있어서 지능적인 서비스를 가능하게 한다. 또한 본 연구를 통하여 제시한 시스템의 개념적 구조는 시맨틱 웹 기반의 검색시스템 개발에 있어서 주요 기능, 구성요소, 적용기술 및 상세 설계의 방향을 제시한다. 본 연구에서 제안한 메타데이터를 통한 시스템의 통합방식은 기존의 연구성과정보 시스템의 서비스 영역을 확대할 뿐 아니라, 더욱 정교해진 지능적인 검색 서비스도 가능하게 한다.

키 워 드

메타데이터, 시맨틱 웹, RDF/RDFS, RQL, 연구성과정보

ABSTRACT

It has lately been recognized that the sharing and exchanging of the research results information is the critical factor to improve the research productivity. So many institutions are planning or developing the information systems which provide the research information services for researcher. But it has very difficulty in integrating the research resources information due to the dispersion and heterogeneity in data sources, and semantic and structural difference in describing data. We propose the semantic web based methodology and conceptual framework for raising the interoperability of metadata about research results information, which will support the integration of the distributed research data for information services in the end. We first introduce the ontology which is developed based on Standard Metadata of Research Results Information published by STISC. Then to show the applicability in real-world environment, we express the metadata of research information in RDF/RDFS according to ontology. Finally we proposed the conceptual architecture of research information service system which shows the main components, the functional requirements, and the principal and design direction at implementing the system.

KEYWORDS

Metadata, Semantic Web, RDF/RDFS, RQL, Research Results Information

1. 서 론

최근 국가적으로 그리고 각 연구기관에서 R&D 투자의 확대와 아울러 연구생산성의 향

상을 위하여 많은 노력을 기울이고 있다. 특히 연구성과정보의 유통을 활성화하기 위하여 국가 R&D 종합관리체계구축의 일환으로서 국가 연구성과정보관리 시스템을 개발 중에 있으며,

각 연구기관에서도 연구개발 성과정보들을 수집, 교환 및 평가하는 정보 시스템을 계획하거나 운영 중에 있다. 그러나 현재 국가는 물론, 각 기관에서도 연구성과정보를 각자의 방법으로 정의하고, 또한 상이한 체제로 관리하고 있음으로 해서 구조적인(structural), 그리고 의미적인(semantical) 면에서 전 연구기관들의 정보들을 통합하여 서비스하기가 매우 어려운 실정이다. 즉, 연구성과정보의 호환성에 문제가 있는 것이다.

2005년 현재 과학기술정보표준화위원회(STISC)에서 검토 중이며 최종 발표될 예정인 “연구성과정보 메타데이터 표준”은 연구성과정보의 상호운용성 확보를 위한 단초의 역할을 할 수 있다. 즉, 과학기술분야의 연구성과물 메타데이터 표준은 국가적으로, 그리고 각 기관 별로 추진 중인 정보 시스템 개발에 있어서 연구성과 데이터를 일관성 있고 체계적으로 관리할 수 있는 기반구조로서 역할을 할 것이다. 또한 현재 국책사업으로 추진 중인 각 정부출연 연구기관의 연구성과 데이터베이스 통합사업에 있어서도 메타데이터베이스를 설계하고, 또한 분산된 데이터베이스로부터 정보를 수집하고 분석하여 저장하는 에이전트의 개발 시에도 활용되어 질 수 있다.

현재 연구성과정보의 서비스는 첫째, 국가적으로 각 연구기관의 데이터베이스를 통합한 후 포털을 통하여 정보를 서비스하는 방법과, 둘째, 각 연구정보 서비스 기관에서 수집한 연구정보를 구매하는 방법과, 셋째, 각 연구기관

의 웹 서버에 접속하여 해당 연구기관에서 제공하는 연구정보를 서비스받는 방법과, 마지막으로 사용자가 웹에서 범용검색 엔진을 통하여 연구정보를 수집하는 방법 등이 있다. 본 연구는 국가수준에서 연구성과정보를 서비스 해주는 첫 번째 방법의 보완차원에서 고려될 수 있다. 분산 데이터베이스 기술을 이용한 각 연구기관의 데이터베이스 통합 시 제도적으로 그리고 기술적으로 다음과 같은 상당한 제약이 있다. 먼저 데이터베이스의 보안차원에서 각 기관은 소극적으로 대처할 수 있으며, 다음으로 공개하는 정보제공의 적시성 및 다양성에도 많은 제약이 있다. 이는 각 기관 자체적으로 연구와 관련된 다양한 정보들을 체계적 관리하고 있지 않기 때문에 정보제공의 어려움이 있는 것이다. 즉, 많은 연구성과정보가 비구조적인 문서의 형태로 존재하기 때문에 데이터베이스화하기가 어렵다. 따라서 현재 추진 중인 연구성과정보 시스템이 완료되더라도 정보 서비스의 양과 질적인 면에서 연구자의 정보 서비스 욕구를 충분히 만족시킬 수 없을 것이다.

본 연구는 시맨틱 웹 기술을 도입하여 연구성과정보의 메타데이터에 대한 상호운용성을 제고하는 방안을 다음과 같이 제시함으로써 궁극적으로 연구성과정보의 통합을 지원하고자 한다. 첫째, 과학기술정보표준화위원회에서 제시한 연구성과정보 메타데이터 표준(안)을 기초로 용어들에 대한 구조적인 그리고 의미적인 통합을 가능케 하는 온톨로지를 개발하는 방안을 제시한다. 둘째, RDF/RDFS를 이용하여 메

타데이터의 지식표현을 하고자 한다. 셋째, 시맨틱 웹 기반의 검색시스템 개발에 있어서 필요한 주요기능, 구성요소, 적용기술 및 상세 설계의 방향을 제시하고자 한다. 제2장에서는 연구의 배경을, 제3장에서 시맨틱 웹 기술의 적용 방법을, 제4장에서는 시스템의 개념적 구조를 제시하고 제5장에서 결론을 맺는다.

2. 연구의 배경

2.1 연구성과정보 상호운용성 수준

정보의 상호운용성(interoperability)이란 사용자가 두 개 또는 그 이상의 기술적인(technical) 시스템에서 만족할 수 있는 방법으로 정보를 직접 교환할 수 있는 상태를 말한다(Mooney 2001). 즉, 상호운용성은 이기종 간 분산환경의 정보 시스템을 서로 연동하여 단일의 통합환경으로 서비스를 제공하는 개념이며 동시에 기능적 요구사항이다(이수상 2004). 통합의 기능적인 면은 3가지 면이 있다. 첫째, 연합(federation) 수준으로 상호간에 엄격한 프로토콜을 준수하여 통합을 하는 것이다. 둘째, 다소 느슨한 연합형태인 수확(harvesting) 수준으로서 각 시스템별로 메타데이터를 교환함으로써 상호연동이 가능한 수준의 통합이다. 마지막으로 수집(gathering) 수준으로 검색 엔진으로 수집하여 서비스하는 수준의 통합이다(Maly et al. 2001).

현재 국내의 각 연구기관 연구정보 시스템

의 데이터베이스는 이질적(heterogeneous)이고, 분산된(distributed) 환경이다. 또한 한 조직 내에서도 연구정보들이 데이터베이스화 되어 있지 않고 다양한 형태의 문서로 흩어져서 저장 관리되고 있다. 정보의 통합에 있어서 첫 번째 방법인 연합에 의한 통합은 연구정보의 접근에 있어서 매우 효율적인 방법이나, 국내의 연구성과정보가 이질적으로 분산된 데이터베이스에, 그리고 일부는 웹상에 흩어져 있기 때문에 이를 논리적, 그리고 물리적으로 통합하기가 매우 어렵다. 또한 세 번째 통합 방법인 일반 검색 엔진을 이용한 수집 수준으로는 소위 정보과다(overload)의 문제와 정보결핍(deficiency)의 문제가 동시에 발생되므로 전문 연구자들에게 있어서 정보검색 만족도는 매우 낮을 수밖에 없다. 따라서 현실적으로 연구성과정보 서비스는 상호운용성의 기능적인 면에서 보면 수확 수준에서 고려되어야 한다. 즉, 각 기관의 연구성과정보관리의 다양성을 인정하고 단지 메타데이터의 상호운용성을 제고하면서 연구성과정보의 통합을 촉진하는 방안이 필요한 것이다.

2.2 연구성과정보 서비스의 요구사항

연구성과정보는 R&D 전주기에 걸쳐서 다양한 문서 및 매체로 발생되며, 그 범위나 대상에 있어서도 통일된 의견은 없다. 연구성과정보 서비스 사용자인 전문 연구자의 관점에서 보면 정밀한 검색을 통한 구체적이고 정확한 정보를

요구한다. 즉, 검색목적 및 내용이 다양하고 검색의 대상도 폭도 넓어질 필요가 있다. 특히 연구성과물 측면에서 보면 공식적인 채널을 통하여 발표된 연구보고서나 논문 뿐 아니라 공식적인 출판경로를 통하여 입수하기 힘든 회색문헌(grey literature)에 대한 정보까지 요구한다(Jeffery 1999). 즉, R&D 관점에서의 회색문헌인 출판 전 배포문, 연구계획 제안서, 콘퍼런스 발표논문, 기술보고서, 학위논문, 회의/세미나/워크숍 자료, 기술계약에 관한 정보, 그리고 특허정보, 연구원 및 연구기관 등에 관한 정보까지 포함된다(Almeida 1999; Auger 1996). 국내 연구소 연구원들에 대한 설문조사의 결과에 따르면 회색문헌은 최신의 연구결과이므로 과학기술의 발전을 위해서 회색문헌의 학술적 가치 및 활용상의 가치가 매우 크며, 각 기관은 물론, 국가적 차원에서 이를 체계적으로 관리해야 할 필요가 있음을 주장하였다(남영준 2002).

그러나 현재 국가적으로 제공되는 연구정보 서비스인 과학기술통합검색(www.yeskisti.net)과 국가연구개발종합 시스템(www.kordi.go.kr/index.jsp)을 분석하면 다음과 같은 문제점이 있다. 첫째, 정보검색 측면에서 연구성과물 중 연구보고서와 같이 공식적으로 발표된 것들에 대한 정보만이 검색 가능하고, 각종 회의 및 세미나 자료, 논문, 단행본 등을 포함하는 기타의 연구성과물에 대한 정보의 검색이 불가능하다. 둘째, 연구성과측정 측면에서 특정 과제와 관련하여 산출된 모든 성과물에 대

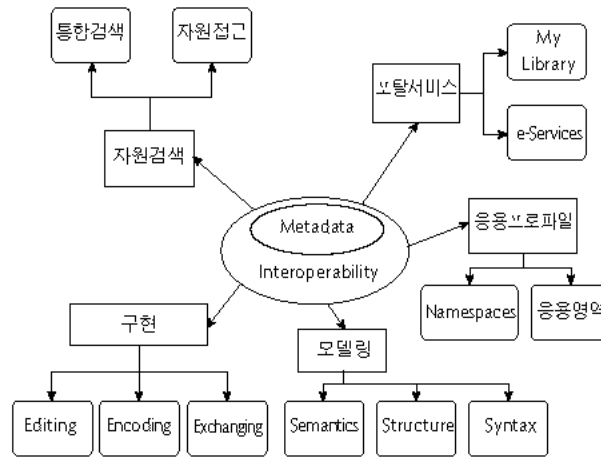
한 추적 및 관리가 시스템적으로 불가능하다. 셋째, 연구기관별 연구성과물 유통과 정보교환이 불가능하다.

또한 웹의 활용으로 인하여 많은 연구정보 자원이 웹상에 존재한다. 이는 연구자가 연구정보 시스템에 연구정보를 입력하기보다는 연구자 자신 혹은 연구과제의 웹 페이지에 연구성과물과 및 연구에 관한 정보를 저장하기 때문이다. 따라서 웹상에 넓게 분포되어 있는 연구정보의 접근도 중요하게 다루어져야 한다.

2.3 연구성과정보 메타데이터의 표현

메타데이터는 데이터에 관한 구조화된 데이터로서 자원과는 독립적으로 존재하면서 자원에 대한 다양한 접근점과 네트워크 주소를 포함하는 데이터이다(Miller 1998). 메타데이터는 다음과 같이 활용된다. 첫째, 자원의 의미를 요약해 준다. 둘째, 자원의 검색을 가능하게 한다. 셋째, 자원의 필요성에 대한 판단을 가능하게 해 준다. 넷째, 다른 자원들과의 연관성을 말해준다(Steinacker et al. 2001).

대상이 되는 자원을 메타데이터로 기술하기 위해서는 먼저 적절한 메타데이터 요소의 집합(element set)을 선택하고, 각 데이터 요소에 대하여 어휘를 정의하고, 그리고 표현방법인 스킴(scheme)을 결정한다. 많은 경우 메타데이터 스킴은 기존의 표준이나 온톨로지(ontology)들로부터 도입한다. 대표적인 웹 자원을 기술하는 표준 메타데이터가 Dublin



〈그림 1〉 메타데이터 상호운용성을 위한 기반기술

Core이다(Dublin Core Metadata Initiative 2005). 특히 과학기술분야의 연구개발 성과물을 위해서는 최근에 CRIS(Current Research Information System)에서 CERIF(Common European Research Information Format) 2004 메타데이터를 제시하였다.

2.4 메타데이터 상호운용성을 위한 기반기술

메타데이터의 상호운용성 보장을 위한 요소 기술은 〈그림 1〉과 같다(이수상 2004). 응용 프로파일(application profiles)은 응용영역의 목적에 부합하도록 메타데이터 스키마를 결정하는 것을 말한다. 메타데이터 스키마는 기존의 표준 메타데이터들로부터 추출한 데이터 요소들과 그들의 스킴을 도입하고 추가적으로 응용영역에서 적합한 데이터 요소들을 포함시키는 것이다.

다음으로 모델링(modeling)은 응용 프로파

일에서 정의된 메타데이터 요소를 실제 시스템의 데이터 형식에 맞춰 상세화하는 과정이다. 즉 다양한 메타데이터 요소 간의 연결을 위하여 메타데이터에 대한 의미론(semantics), 물리적 구조(structure), 그리고 구문론(syntax)에 대하여 결정하고, 실제 개체의 메타데이터가 시스템에서 어떻게 표현되어야 하는가를 제시하는 것이다. 메타데이터의 표현방식으로는 RDF 모델, RDF Schema 모델, XML DTD 모델, XML Schema 모델 등 다양한 방법이 있다.

구현(implementation)은 프로그램 코딩 영역으로 메타데이터 편집기를 제작하는 것이다. 마지막으로 검색시스템(resource discovery)의 구현 및 포털 서비스 개발도 메타데이터 상호운용성을 위해서 필요한 기술들이다.

2.5 시맨틱 웹을 활용한 연구성과정보 서비스

시맨틱 웹은 기존의 웹 서비스의 한계를 극

복하기 위하여 1998년 Berners Lee가 주창한 차세대 웹 기술이다. 시맨틱 웹은 웹이 제공하고 있는 정보를 잘 정의된 온톨로지를 기반으로 하여 표현함으로써 물리적·논리적으로 분산되어 있는 애플리케이션 간의 상호운용성을 제공하기 위한 것이다(이승희 외 2003). 즉, 시맨틱 웹에서는 다른 데이터 구조를 갖고 있는 애플리케이션도 온톨로지를 통해 상대방의 정보를 이해하고 처리할 수 있다. 이때 사람이 인식하고 이해하는 것이 아니라 기계, 즉 컴퓨터가 이해하고 정보를 처리하는 것을 말한다. 이와 같이 지능형 처리 즉, 컴퓨터가 이해할 수 있는 정보를 만든다는 것은 메타데이터에 대한 접근법에서 출발한 것이다. 시맨틱 웹은 메타데이터를 기반으로 한 온톨로지 및 지식표현(knowledge representation)이 핵심이다(이재호 2002).

시맨틱 웹은 계층화된 구조(layered structure)이다. 제일 하단에는 구문(syntax)을 전달할 수 있도록 XML이 기초를 이루고 있다. RDF는 정보표현 프레임워크를 제공하고 있으며, RDFS는 데이터 모델링의 구조로서 클래스와 속성을 정의하고, 그들 사이의 관계를 정의한다. 그 위의 온톨로지는 RDF/RDFS를 포함하며, 지능적 처리가 가능하도록 정보의 논리관계도 표현하는 것이다(Berners_Lee et al. 2001).

온톨로지란 특정 도메인에 있어서 어휘들의 의미를 기술하고 어휘들 간의 상호관계들에 대한 표현으로 어떤 영역을 기술하기 위해 사용

하는 개념과 그리고 유사한 어휘들을 모아 놓은 어휘집에 관한 규격이다(W3C 2005). Lopatenko와 그의 동료들에 따르면 연구정보 검색 시 다음과 같은 문제가 있을 때 시맨틱 웹 기술이 활용되어 질 수 있다고 주장한다(Lopatenko et al. 2002). 첫째, 데이터 소스가 분산되어 있으며, 데이터가 구조적으로 그리고 의미적으로 서로 차이가 나며, 정보검색 시 이를 고려해야 하는 경우, 둘째, 정교한(sophisticated) 정보검색이 요구될 때, 셋째, 기존의 데이터 구조와 새로운 데이터의 구조가 서로 호환될 필요가 있을 때, 넷째, 기존의 시맨틱 웹 데이터를 활용하여야 할 때, 다섯째, 데이터베이스에 직접적인 접근이 불가능할 때, 마지막으로 다른 시맨틱 웹 기반의 구조와 호환될 필요가 있을 때이다.

Lopatenko와 동료들에 의해서 유럽의 연구기관 간의 연구정보교환을 위해 연구성과정보 시스템 프로토타입인 AURIS MM(Austrian Research Information System Multimedia Enhanced)이 개발되었다(http://derpi.tuwien.ac.at/~andrei/AURIS_MM_plan.html). 이것은 기존의 오스트리아 연구정보 시스템인 AURIS에 시맨틱 웹 기술을 채택하여 기능적으로 확장한 시스템이다. 이 시스템은 유럽에서 표준 연구성과정보 메타데이터로 인정되는 CERIF 2000을 기반으로 온톨로지를 개발하고 이를 RDF/RDFS로 표현하였다. 본 연구는 AURIS MM과 같은 취지에서 우리나라의 특수성을 반영한 시맨틱 웹

기반의 연구성과정보 서비스 체계를 제안하고자 한다.

3. 상호운용성 제고를 위한 시맨틱 기술의 적용기법

3.1 온톨로지의 개발

1) 메타데이터 생성

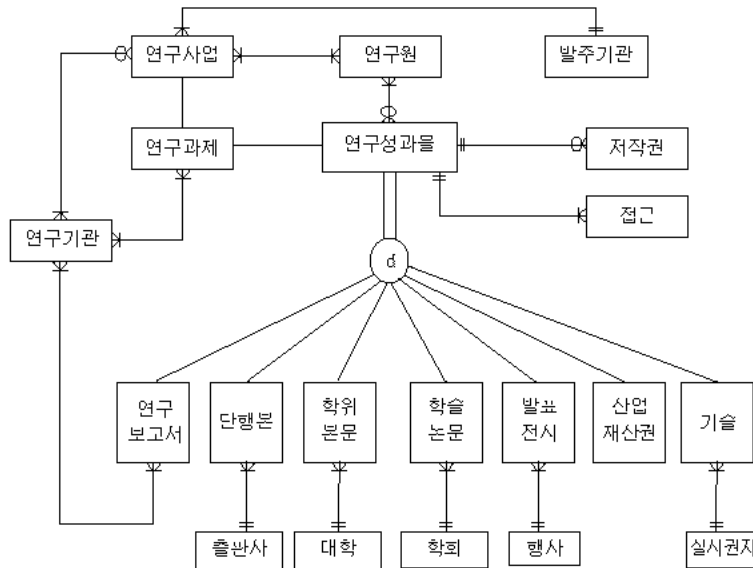
제2장에서 언급되었듯이 분산 데이터베이스 기반의 연구정보검색은 효율적인 검색을 가능하게 하나, 폭넓은 정보 및 최신의 정보를 확보하기가 매우 어렵다. 따라서 이를 보완하기 위해서는 웹상에 제공되어 있는, 그리고 계속적으로 증가될 연구정보자원을 확보하는 것이 매우 중요하다. 그러나 2.1 소절에서처럼 범용 검색 엔진의 방법처럼 웹 정보를 확보하여 서비스하는 것에도 문제가 있음을 확인했다.

본 연구에서는 웹 페이지 주석(annotation) 삽입을 통하여 연구성과정보에 대한 시맨틱 메타데이터를 확보하는 방법을 먼저 제안한다. 즉, 연구자가 본인의 연구정보를 입력할 때 해당 연구성과정보의 의미적 메타데이터를 자동으로 생성하게 한 후 이를 웹 페이지에 포함시키거나 따로 특정장소에 저장하는 방법이다. 다음으로 소프트웨어 에이전트가 웹 페이지에 포함된 메타데이터나 저장된 연구성과정보 메타데이터를 인터넷을 통하여 수집하고 분석한 후, 정보검색을 위하여 하나의 통합 메타데이터베이스에 연구성과정보 메타데이터를 저장하는 방식이다. 웹 페

이지 뿐 아니라 연구자가 각 조직의 기간 시스템인 연구정보 시스템에 정보를 입력할 때도 메타데이터를 자동으로 생성하고 이를 특정 장소에 저장할 수 있다. 따라서 이 방법은 웹 자원으로 만들어진 연구성과정보 뿐 아니라 연구정보 시스템에 직접 입력한 연구성과정보에 대해서도 메타데이터를 작성할 수 있다. 메타데이터의 작성 시 혹은 소프트웨어 에이전트가 메타데이터의 구성 및 어휘들의 의미를 파악하고 분석하기 위해서는 온톨로지가 필요하다.

2) 표준 메타데이터를 이용한 온톨로지의 개발

연구성과 정보검색시스템은 다음과 같은 목적으로 사용한다. 즉, 연구자의 경우에는 기존의 연구성과물에 대한 정보를 검색하여 본인의 연구에 활용하고자 하는 것이다. 또한 특정연구가 어떤 목적으로, 누구에 의해서, 어떤 규모로, 얼마 동안, 무슨 장비를 사용하여 진행되었는가에 대한 정보를 검색할 수도 있다. 뿐만 아니라 연구를 통하여 개발된 기술이 특허를 취득하였는지, 기술이전이 가능한 것인지에 대한 정보가 필요할 수도 있으며, 그리고 공동 연구를 위해서 전문성이 있는 파트너를 찾고자 할 수 있다. 반면 정책입안자나 연구성과 평가자의 경우에는 해당 연구과제에 대한 연구성과물의 평가 및 측정을 위해서도 사용되어 질 수 있다. 이와 같이 연구성과정보를 검색할 때 같은 목적이나 사용하는 어휘가 서로 차이가 있을 수 있으며, 같은 어휘라 할지라도 서로 다른 관점에서 사용할 수 있다. 또한 시스템 관점에서 보면, 정보를



〈그림 2〉 연구성과물 데이터 모델

서비스하는 사람 및 기관에서도 서로 다르게 정의된 어휘들을 사용할 수도 있는 것이다. 따라서 정보의 제공 및 검색이 지능적인 기능을 갖기 위해서는 이러한 구조적, 그리고 의미적 차이를 극복하는 온톨로지의 사용이 필수적이다.

3) 연구성과물 데이터 모델

데이터 모델은 메타데이터의 정적인(static) 면을 보는 것으로 메타데이터에 포함될 개체를 파악하고 데이터 모델링을 통하여 개체 유형의 구분, 연관관계, 세부 데이터 요소들을 파악하는 것이다. 연구성과물에 대한 메타데이터는 크게 ① 연구성과물 자체, ② 연구성과물 생산 주체, 그리고 ③ 연구성과물 유통과 관련된 요인들로 그룹화할 수 있다. 이상의 내용을 ERD로 표현하면 〈그림 2〉와 같다.

각각의 연구성과물을 세분하면 아래와 같다.

- ① 연구보고서(중간보고서/최종보고서/n차 보고서)
- ② 학술논문(국내 대학 및 기관논문지, 학술진흥재단등록, 학술진흥재단등록 후보, 전국 규모 학회지, 지역학회지, 해외 학회)
- ③ 각종 발표회 및 전시회(국내학술대회발표/해외학술대회발표/심포지엄/포럼/세미나/강연회/회의/전시회)
- ④ 단행본 및 간행물
- ⑤ 학위논문
- ⑥ 산업재산권(국내 특허/국제 특허/국내 실용신안/국제 실용신안/디자인/소프트웨어 등록)
- ⑦ 기술개발(기술등록/기술이전)

4) 연구성과정보 온톨로지 표준용어

연구성과정보를 위한 온톨로지의 개발을 위해서는 먼저 표준화된 용어(terms)들에 대한

정의가 선행되어야 한다. 현재 과학기술정보표준화위원회(STTSC)에서 검토 중이며 최종 발표될 “연

〈표 1〉 메타데이터 구성

		온톨로지 용어
공통 메타데이터 요소	식별기호 범주	연구성과물 식별기호(연구성과물 공개 식별기호, 연구성과물 내부관리 식별기호)
	연구성과물 일반 범주	연구성과물 제목(연구성과물 국문 제목, 연구성과물 영문 제목), 언어, 연구성과물 분류, 날짜(연구성과물 작성일, 연구성과물 발표일), 연구성과물 작성자(연구성과물 작성참여 인원수, 연구성과물 책임작성자 이름, 연구성과물 책임작성자 소속, 연구성과물 참여작성자 이름, 연구성과물 참여작성자 소속), 연구성과물 발행자(연구성과물 발행자 이름, 연구성과물 발행기관, 연구성과물 발행국), 과학기술 정보분류(계열, 대주제, 중분류), 연구성과물 주제어, 연구성과물 내용기술(연구성과물 목차, 연구성과물 요약기술), 연구성과물 저장매체 형태, 연구성과물 연결(연구성과물 접근 및 파일 이름, 연구성과물 관련정보 URL), 연구성과물 소장기관, 연구성과물 저작권, 연구성과물 공개여부
개별 연구 성과물 메 타데이터 요소	연구보고서 범주	연구보고서 유형코드, 연구보고서 면수
	학술논문 범주	학술지 서명, 학술지 권 번호, 학술지 호 번호, 학술지 발행계절, 시작 페이지, 끝 페이지, 학술지 분류, 학술지 간행빈도
	각종 발표회 및 전시회 범주	이벤트 명칭, 이벤트 주제, 이벤트 주최기관, 이벤트 공식언어, 세션주제, 발표자(발표방식, 발표자 이름, 발표자 소속), 행사일(이벤트 시작일, 이벤트 종료일)
	간행물 범주	ISBN, 출판사이름, 판매가격, 시작 페이지, 끝 페이지
	학위논문 범주	학위논문 식별자, 수여대학, 전공분야, 학위유형, 지도교수
	산업재산권등록 범주	IPC code, 출원(출원자, 출원번호, 출원일자), 최종처분, 등록(등록번호, 등록일자), 지정국
	기술개발 범주	기술활용분야, 계약당사자(기술실시 허락자, 기술실시권자), 기술이전형태, 기술실시 계약기간, 계약금액 및 지불방식(기본료, 경상기술료)
연구주체 메타데이터 요소	연구과제 범주	연구과제 관리번호, 연구과제 명칭(연구과제 한글명칭, 연구과제 영문명칭), 연구과제 유형, 연구과제 예산, 연구과제 수행단계, 연구과제 수행기간, 연구과제 실시일(연구과제 시작일, 연구과제 종료일), 연구과제 참여연구원 수, 연구과제 수행상태, 연구과제 책임자(이름, 소속), 상위 연구사업(사업명, 사업형태)
	연구기관 범주	연구기관 식별자, 연구기관 명칭, 연구기관 타입, 연구기관 웹 사이트, 연구기관 역할
	발주기관 범주	발주기관 식별자, 발주기관 명칭, 발주기관 웹 사이트, 발주기관 과제 관리번호, 주관부처
	연구원 범주	연구원 개인식별자, 연구원 이름, 연구원 성별, 연구원 개인 웹 사이트 URI, 연구원 전화번호, 연구원 이메일, 연구원 최종학위, 연구원 직급, 연구원 국적, 프로젝트 참여역할, 프로젝트 참여시기(프로젝트 참여 시작일자, 프로젝트 참여 종료일자)

구성과정보 메타데이터 표준”은 온톨로지 개발을 위한 표준용어를 제공한다. 위에서 설명한 데이터 모델을 기반으로 개발된 STISC의 연구성과정보 메타데이터 표준(안)이 온톨로지 개발의 기반으로 사용되어 질 수 있다.

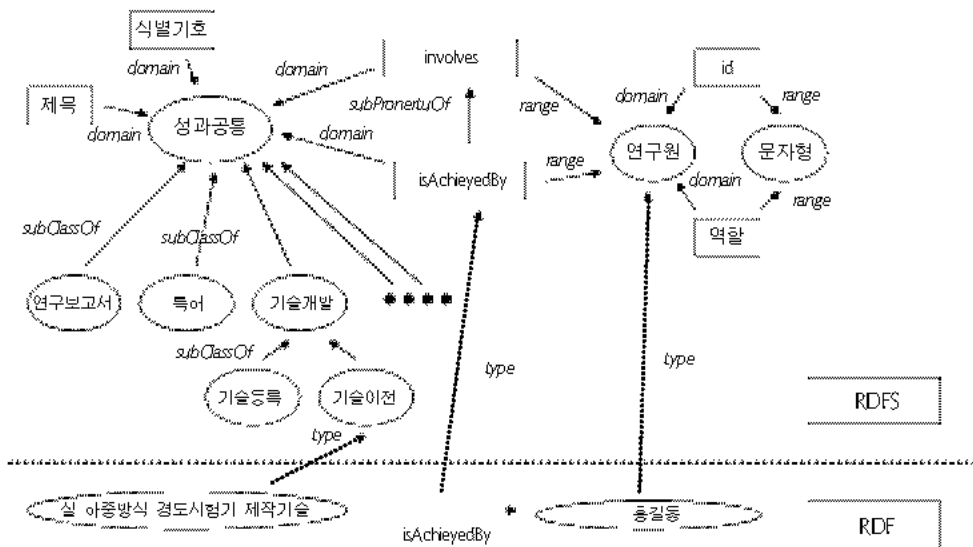
연구성과정보 메타데이터 표준(안)의 메타데이터 요소는 Dublin Core의 기본요소 및 하위요소를 근간으로 하고, 국제적으로 연구성과물 관리의 대표적인 메타모델인 CERIF에서 사용되는 요소를 수용하며, 국내 연구기관에서 연구성과물 관리에 적합한 기술요소를 선택적으로 적용하여 구성되었다. 연구성과물 메타데이터는 13개의 범주로 구분되고, 총 82개의 주요소, 47개의 하위요소로 구성되어 있다. 그러나 이상과 같은 메타데이터를 정보검색용 시스템

에서 필요로 하는 온톨로지 개발에 활용하기 위해서는 각 범주에 소속된 용어(메타데이터에서는 데이터 요소에 해당됨)들에 대한 추가적인 정의가 필요할 것이다. 예를 들면, 온톨로지 개발과정을 거친 연구성과정보 온톨로지는 <표 1>과 같은 용어들을 포함할 것이다.

3.2 연구성과정보 지식의 표현

1) 연구성과정보 도메인 모델링

RDF는 W3C에서 추천하는 시맨틱 웹 환경의 메타데이터 표현언어이며, 문서에 의미를 더하는 방법을 제공한다. 즉, 구조화된 메타데이터를 인코딩하고, 교환하고, 재사용하는 것을 가능하게 하는 기반구조인 것이다. 따라서 RDF는 이



<그림 3> 연구성과물 메타데이터 도메인 모델링

질적이고 분산된 데이터를 축적하는 연구성과정보 메타시스템의 구현에 적합한 언어이다.

RDF는 RDFS와 짝을 이루어 필요에 따라 확장할 수 있는 모델링 근원어(primitives)를 제공하며, 문법적으로 RDF 모델은 XML로 저장하고 교환할 수 있다. 본 연구에서는 장기적 측면

에서 연구성과물의 상호운용성 제고를 위하여 검색 엔진, 지능형 에이전트 등의 개발과 시맨틱 웹 환경을 고려하여 RDF/RDFS를 이용하기로 한다. 특히 RDFS는 용어들 간의 관계를 표현하는 기본 어휘들과 제약방식을 활용하여 도메인의 기본적인 온톨로지를 구성할 수 있게 한다.

```

<rdf:RDF
  xmlns:rdf: "http://www.w3c.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs: "http://www.w3.org/2000/01/rdf-schema#"
  <rdfs:Class rdf:ID="researchResultCommon">
    <rdfs:comment>
      연구성과물 공통요인 class
    </rdfs:comment>
  </rdfs:Class>
  <rdfs:Class rdf:ID="technologyDevelopment">
    <rdfs:comment>
      기술개발 class
    </rdfs:comment>
  <rdfs:subClassOf rdf:resource="#ResearchResultCommon" />
  </rdfs:Class>
  ... 중략 ...
  <rdfs:Property rdf:ID="researcher">
    <rdfs:comment>
      연구원 property
    </rdfs:comment>
  </rdfs:Property>
  <rdfs:Property rdf:ID="isAchievedBy">
    <rdfs:subPropertyOf rdf:resource="#involves" />
    <rdfs:domain rdf:resource="#researchResultCommon" />
    <rdfs:range rdf:resource="#researcher" />
  </rdfs:Property>
</rdf:RDF>

```

〈그림 4〉 RDFS로 기술한 도메인 모델

〈그림 3〉은 RDF/RDFS를 이용하여 연구성과물 영역의 도메인을 모델화 한 것의 일부이다. RDFS로는 연구성과정보 클래스의 계층적 구조, 클래스와 클래스, 클래스와 요소(properties), 요소와 요소 간의 관계를 표현한 것의 일부이다.

〈그림 3〉을 설명하면 다음과 같다. 기술개발 클래스는 성과공통 클래스의 하위 클래스(subClassOf)이고 기술등록은 기술개발의 하위 클래스이다. 그리고 연구성과물은 연구원에 의해서 완성된다는 관계를 보여주는 isAchievedBy는 involves 의 하위요소(subPropertyOf)이다. 또한 각 관계형 요소는 관계상의 주체(subject)가 되는 클래스와 값(value)을 갖는 클래스에 대하여 domain과 range를 이용하여 대상에 있어서 제약을 가할 수 있다. 즉 연구성과물은 연구원에 의해서만 작성된 것 만을 인정한다라는 제약이 있다면

IsAchievedBy 의 주체인 성과공통이 domain 이 되고, 연구원이 range가 된다. 또한 type을 이용하여 RDFS에서 정의된 객체들과 RDF의 실제 데이터 간의 관계를 맺어준다.

2) RDF/RDFS를 이용한 도메인 지식의 표현

〈그림 4〉는 〈그림 3〉에서 방향성 그래프로 모델링한 도메인의 일부를 RDFS 타입으로 변환한 것이다. 즉, 연구성과물 공통요인(researchResultCommon) 클래스와 하위 클래스인 기술개발(technologyDevelopment) 클래스, 그리고 연구원(researcher) 요소와 연구성과물 공통요인 클래스 간의 관계를 설명하는 isAchievedBy 요소를 정의한다. 특히 isAchievedBy에는 domain과 range를 이용하여 적용되는 대상 클래스를 구체적으로 표시한다.

도메인 전체에 대한 명세(specification)는 본 연구의 범위를 벗어나지만, 이상과 같은 방

```

<xml version="1.0" encoding="UTF-16"?>
<rdf:RDF
  xmlns:rdf="http://www.w3c.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:result="http://www.mydomain.or.kr/result-ns#"
  <rdf:Description rdf:about="http://www.mydomain.org/result/#200410-4112">
    <dc:title>실하중식 경도시험기 제작기술 </dc:title>
    <result: isAchievedBy> 홍길동 </result:isAchievedBy>
    .....
  </rdf:Description>

```

〈그림 5〉 RDF 기반의 메타데이터

법으로 메타데이터의 모든 어휘들과 그들 간의 관계 및 구조를 정의한다.

다음으로 RDF 편집기를 이용하여 <그림 5>와 같은 RDF 메타데이터를 생성한다. 먼저 선언부를 기술하고 모델링 스키마를 준수하며 메타데이터를 생성한다. 선언부는 RDF 문서가 XML 기반이고 이름공간(namespace)을 이용하므로 반드시 선언부에 명시하여야 한다. Dublin Core 표준인 dc와 <그림 5>에 포함된 result가 이름공간이다. 선언부를 표시한 후 <rdf:Description>과 about 속성을 이용하여 자원을 표시한다. <그림 5>에는 기술이전 성과물로 파악되는 식별기호가 200410 4112인 “실 하중방식 경도시험기 제작기술”에 대한 메타데이터 기술이다.

비록 본 논문에서는 연구성과물 메타데이터 스키마 전체와 RDFS로 표현한 도메인 구조 및 RDF로 기술한 연구성과물 메타데이터를 전체는 언급하지 못하였지만, 연구성과물의 상호용성의 제고를 위한 기반구조 구축에 방향을 개략적으로 언급하였다. 특히 RDFS/RDF에 의한 메타데이터 표현은 해당분야의 온톨로지의 개발을 가능케 하고, 궁극적으로는 “지능적인(intelligent)” 정보 서비스를 가능케 할 것이다.

3.3 연구성과정보 메타데이터 처리

1) RDF 메타데이터 문서의 생성 및 저장

연구성과정보를 검색하기 위해서는 먼저 RDF로 표현된 각 연구성과정보의 메타데이터

가 수집되고 분석된 후 저장되어야 한다. 즉, 연구자에 의해서 메타데이터가 작성되는데 이때 웹 페이지 주석 제작 및 삽입을 위한 틀이 사용된다. 이 틀은 웹 문서의 경우 문서의 시작 태그(opening tag)와 마감 태그(closing tag) 사이에 있는 텍스트에 대한 의미를 파악하여 RDF 형식의 태그를 부여하고 문서를 작성한 후 이를 저장하는 기능이 있다. 즉, 이 틀은 웹 환경에서 연구자가 연구성과정보를 입력하고, 아울러 온톨로지 브라우저를 열어서 온톨로지 용어, 관계, 속성 등을 이용하여 연구성과정보에 대한 메타데이터를 기술하게 한다. 또한 자동으로 RDF 시맨틱 메타데이터를 생성하고 이것을 웹 페이지에 포함시키거나, 정해진 장소에 RDF 메타데이터를 보관한다. 다음으로 저장된 RDF 메타데이터는 소프트웨어 에이전트(일반적으로 crawler라고 함)에 의해서 수확(harvesting)된 후 분석되고 통일된 포맷으로 RDF 데이터 저장소(repository)에 이를 저장한다. 연구자가 웹 페이지가 아니라 연구정보 시스템에서 직접 연구정보를 입력하는 경우에는 관계형 데이터베이스를 이용하여 메타데이터를 생성하여 저장하고 이를 한꺼번에 RDF 형식의 메타데이터로 변환한다.

2) RQL을 이용한 질의

RDF 문서는 구문적으로 XML의 표현기법을 따르지만, XML 기반의 질의언어로는 RDF로 표현된 메타데이터의 의미를 파악할 수 없다. 즉, 질의어 역시 RDF 스키마의 의미적 구성을 알고 있어야 되는 것이다. RDF용 질의언

(type 연구원 class)	(domain isAchievedBy 성과공통)
(type 성과공통 class)	(range isAchievedBy 연구원)
(type 기술개발 class)	(type "200410-4112" 기술이전)
(type 기술이전 class)	(type "730245" 연구원)
(subClassOf 기술개발 성과공통)	(isAchievedBy 실허중경도시험기제작기술 홍길동)
(subClassOf 기술이전 개술개발)	
(type involves property)	단, "200410-4112"는 "실허중경도시험기제작기술" ID 이며, "730245" 는 연구원 홍길동의 ID 이다.
(type isAchievedBy property)	
(subPropertyOf isAchievedBy involves)	

〈그림 6〉 트리플로 표현한 RDF 문서

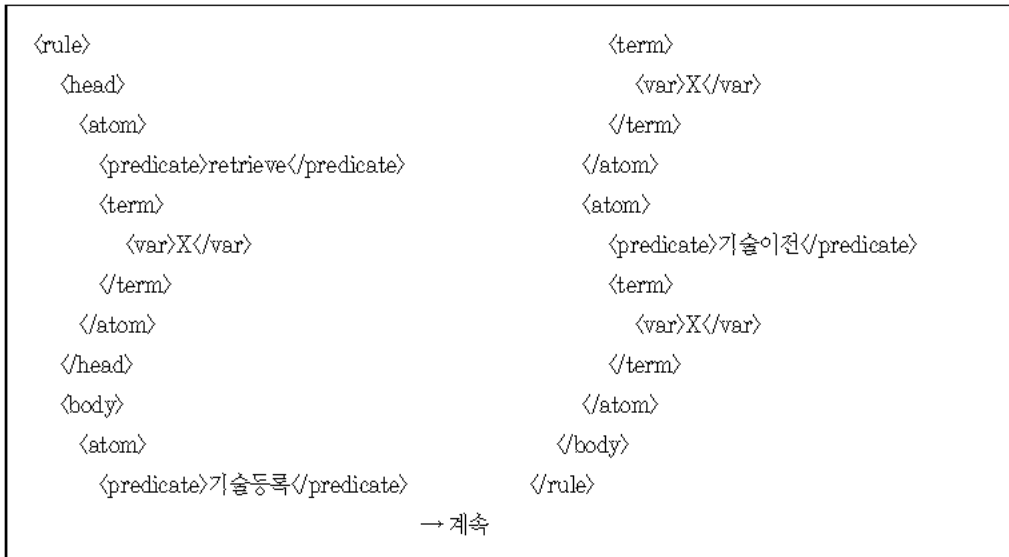
어로 RQL(RDF Query Language)이 범용적으로 사용된다. 관계형 데이터베이스의 표준질 의언어인 SQL과 마찬가지로 사용자가 직접 RQL을 사용하여 RDF 문서에 대한 질의는 어렵다. 따라서 RQL을 지원하는 질의용 툴을 선택하여 검색시스템을 개발하는 것이 바람직하다. RDF 질의용 툴은 사용자가 온톨로지를 포함하는 질의용 화면을 통하여 RDF 문서의 구조적인 내용은 물론, 의미적인 내용까지 질의할 수 있는 것이다.

RQL을 통한 질의처리의 예를 보면 다음과 같다. 3.2절의 1)에 있는 〈그림 3〉은 연구성과 정보 도메인을 그래프로 표현한 것이며, 3.2절의 2)에 있는 〈그림 4〉와 〈그림 5〉는 이를 RDF/RDFS 문서로 변환한 것이다. 〈그림 6〉은 RQL에 의하여 어떻게 질의되는가를 예시하기 위하여 RDF 문서 일부를 개체(object) 속성(attribute) 값(value)으로 구성된 트리플(triple)의 집합으로 표현한 것이다.

예를 들면, "홍길동"이란 연구원에 의해서 개발된 모든 기술의 이름을 파악하고자 한다면 RQL을 이용하여 다음과 같이 질의할 수 있다. 〈그림 6〉에는 언급되어 있지 않지만 속성인 title은 성과공통 클래스가 도메인이고, 그리고 속성인 researcherName은 연구원 클래스가 도메인이다. 또한 변수 K와 Y는 연구원 객체의 식별자를, C와 X는 연구성과물 객체의 식별자를 값으로 갖으며, N은 성과물의 제목값을 갖는 변수이다.

```
select N
from 기술개발(X).isAchievedBy(Y), (C)title(N),
(K)researcherName(P)
where P="홍길동" and Y=K and X=C
```

또한 RQL은 각 객체들 간의 관계성을 보이는 스키마 정보도 데이터 정보와 함께 검색할 수 있다. 즉, 클래스 변수 앞에 \$를, 그리고 속



<그림 7> 규칙의 RDF 표현

성변수 앞에 @를 접두사로 붙임으로 해서 스카마 정보를 검색할 수 있다. 예를 들면, 앞의 질의문의 select 문을 “select \$N, N”라고 하면 결과는 다음과 같이 나오게 할 수 있다.

- (“기술등록”, “실하중식경도시험기제작기술”)
- (“기술이전”, “실하중식경도시험기제작기술”)
- (“기술등록”, “조직응고기용 초고주파 특수안테나개발”)

여기서 첫 번째 결과인 “실하중식경도시험기제작기술”은 기술이 등록되어 이전이 실시된 경우이며, 두 번째 기술인 “조직응고기용 초고주파 특수안테나개발”은 기술이 개발되어 등록만 된 경우이다.

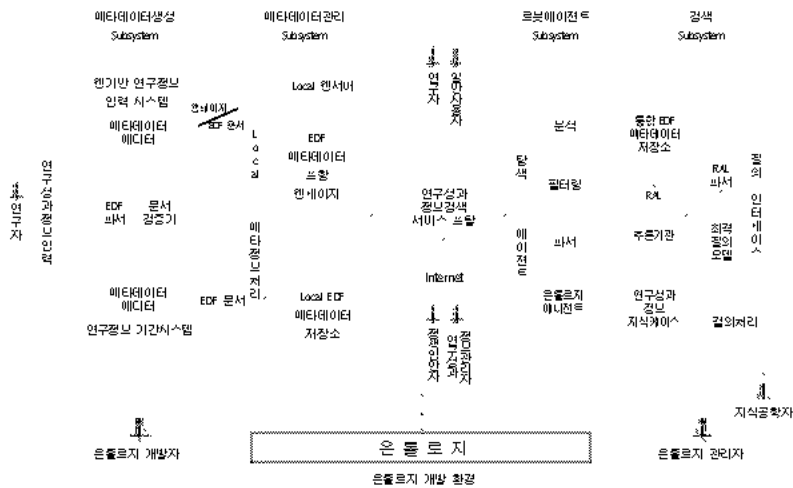
3) RDF 문서의 추론방법

RDF 로 표현되는 문장은 인공지능의 지식 표현 방법인 술어논리(predicate logic)의 특별한 형태이다. 따라서 추론을 위한 엔진을 사용

함으로써 RDF 문장의 집합으로부터 새로운 사실을 추론할 수 있으며, 이를 증명할 수도 있다. 또한 규칙(rule)을 RDF 문장으로 포함시킴으로써 더욱 정교한 검색을 가능하게 한다. 예를 들면, 기술개발실적을 검색하는 과정에서 “기술등록”이 되고 동시에 “기술이전”까지 성공한 기술에 검색의 우선순위를 주기 위해서는 다음과 같은 규칙을 포함할 수 있다.

- r1: 기술등록(X) ⇔ retrieve(X)
- r2: 기술이전(X) ⇔ retrieve(X)
- r3: 기술등록(X), 기술이전(X) ⇔ retrieve(X)
- r3 > r1
- r3 > r2

여기에서 r3 규칙을 RDF 문장으로 표현하



〈그림 8〉 메타데이터 기반 연구성과정보 시스템의 개념적 체제도

면 〈그림 7〉과 같다. 다음으로 r3가 r1, r2보다 우선순위가 있다는 규칙을 RDF 문장으로 표현 하면 된다.

4. 메타데이터 기반 시스템의 개념적 구조

본 장에서는 시맨틱 웹 기술의 적용관점에서 메타데이터 기반 연구성과정보 시스템의 개념적 구조를 제시한다. 〈그림 8〉에서처럼 시스템은 온톨로지를 기반 인프라로 하며, 기능적으로 크게 4개의 서브시스템으로 구성된다. 메타데이터 생성 서브시스템과 메타데이터 관리 서브시스템은 각 연구기관에서 관리되며, 로봇 에이전트 서버 시스템과 검색 서브시스템은 국가수준에서 연구성과정보를 서비스하는 기관에서 관리한다. 각 서브시스템의 구성요소 및

설계특성을 보면 다음과 같다.

4.1 메타데이터 생성 서브시스템

메타데이터 생성 서브시스템은 각 연구기관의 웹 환경 및 기간 시스템 환경에서 구동되는데 핵심은 메타데이터 편집기와 RDF 파서 및 문서검정기이다. 메타데이터 편집기는 RDF 메타데이터 작성을 위한 환경을 제공한다. 즉, 연구성과정보 온톨로지를 로드시켜 비주얼화 함으로써 사용자가 메타데이터용 엔티티를 선택하고, 세부사항인 속성들을 입력하게 한다. 또한 생성된 엔티티와 기존의 엔티티간의 관계도 설정한다. 또한 메타데이터 편집기는 웹 주석(annotation) 기능이 있어 웹 페이지에 작성된 메타데이터를 삽입하는 기능도 있다. 마지막으로 생성된 메타데이터는 메타데이터 저장소로

보내고, 메타데이터를 포함하고 있는 웹 페이지는 웹 서버로 보낸다. RDF 파서는 연구성과정보 메타데이터가 올바르게 작성되었는지를 확인하기 위해서 RDF 문서에 대하여 구문검사를 실시하는 기능이 있으며, 텍스트, 트리, 그래프에서 저작중인 RDF 및 XML 문서에 대한 파싱을 실시한다. RDF 문서 유효성 검증기(validator)는 RDF 문서를 생성하기 바로 전 단계에서 문서의 유효성을 검사한다. 유효성검사는 메타데이터에 포함된 의미(semantic)를 가진 여러 어휘들의 DTD와 스키마를 검증함으로써 표현의 정확성을 확보하는 것이다.

4.2 메타데이터 관리 서브시스템

메타데이터 관리 서브시스템은 생성된 메타데이터를 관리하는 기능들을 포함한다. 로컬 메타 정보처리 기능은 각 연구자의 홈페이지, 각 연구과제의 홈페이지, 그리고 연구정보 시스템 등으로부터 작성된 RDF 문서를 주기적으로 수집하여 웹 서버 혹은 메타데이터 저장소에 RDF 메타데이터 문서를 저장하고 기존의 저장된 메타데이터를 관리하는 기능을 한다. 그리고 RDF 문서가 포함된 웹 페이지는 웹 서버에 등록함으로써 연구성과정보를 연구기관에서 직접 외부에 공개할 수도 있다. 연구기관 내에서 작성된 모든 RDF 메타데이터를 저장하는 저장소(repository)는 다양한 유형의 데이터베이스가 사용될 수 있지만, 관계 객체혼합형 데이터베이스가 RDF 스키마의 구조를 모델

화하는 데 이점이 있다.

4.3 로봇 에이전트 서브시스템

로봇 에이전트는 RDF 메타데이터를 수확(harvesting)하는 기능을 하는 지능형 에이전트이다. 에이전트는 웹을 자동적으로 순화하는 프로그램으로서 온톨로지를 지식기반으로 각 연구기관에서 새로 등록한 RDF 메타데이터를 수집하고, 필터링하여 검색시스템의 통합 RDF 저장소로 가져오는 역할을 한다. 로봇 에이전트는 RDF 문서의 분석을 위해서 RDF 파서가 필요하다. 또한 로봇 에이전트는 기능적으로 보면 도메인 온톨로지와 RDF 구조를 조화하는 온톨로지 에이전트(Ontology Agent)와 RDF 문서를 탐색하여 수집하는 탐색 에이전트(Search Agent)로 구분된다.

4.4 검색 서브시스템

전체 시스템에서 가장 핵심적인 역할을 하는 것이 검색 서브시스템이며, 기술적으로 가장 많은 검토가 필요한 부분이다. 검색 서브시스템은 크게 질의처리 모듈, 지식 베이스 및 추론기관, 통합메타데이터 저장소 등으로 나누어진다. RQL을 기반으로 하는 질의 모듈은 사용자 검색 폼으로부터 입수된 정보검색 요구사항을 RQL 질의문장으로 변환시키고 RQL 파서를 이용하여 최적의 질의 모델을 생성시킨 후 검색을 요청한다. 지식 베이스와 추론기관은

지식 베이스에 포함된 규칙(rule)과 사실(fact)을 활용하여 새로운 결론을 추론(reasoning abilities)할 수 있는 능력과 온톨로지의 스키마 탐색(schema exploration) 능력이 검색 엔진의 정확도를 더욱 향상시킨다. 통합 RDF 메타데이터 저장소는 모든 연구기관으로부터 입수된 메타데이터를 저장하는 곳이다. 로컬 메타데이터 저장소와는 같은 구조지만, 대용량의 RDF 데이터를 처리해야 하므로 검색시스템의 성능에 매우 중요한 역할을 하나, 아직 기술적으로 완전히 해결되지 않고 있는 실정이다. RAL(Repository Abstraction Layer)은 관계형데이터베이스의 RDBMS의 역할을 하는 것으로 RDF 데이터와 시스템 간의 독립성을 유지하고, 각종 RQL 질의어를 처리하고, RDF 관리(administration) 기능을 포함한다.

5. 결 론

연구성과정보 시스템에 있어서 시맨틱 웹 기술의 적용은 기존의 문제점을 해결할 뿐 아니라 정보 서비스 및 상호운용성의 제고 차원에서 많은 가능성을 보인다. 본 연구에서는 기술적 관점에서 구체적인 시스템 구현 가능성을 검토하고 개념적 구조를 제시하였다. 특히 연구성과정보 메타데이터 표준(안)을 기초로 개발되는 온톨로지는 연구성과정보의 통합에 있어서 기반 인프라의 역할을 할 것이다. 또한 RDF/RDFS를 이용한 연구성과정보의 표현은 정보 서비스의 검색과 유통에 있어서 '지능적

인(intelligent)' 서비스를 가능하게 한다.

구체적으로 본 연구의 결과는 연구성과정보의 통합 면에서 다음과 같은 기대효과가 있다. 즉, 시맨틱 웹 기술의 채택은 이질적이고 분산된 데이터 소스로부터 정보를 수확(harvesting)하는 방법으로 시스템을 통합하게 한다. 또한 국가차원이나 각 연구기관 차원에서 야기되는 연구성과 데이터의 구조 및 의미의 비일치성(discrepancies)에 따른 호환성 및 상호운용성 문제를 해결할 수 있는 온톨로지 기반의 애플리케이션 개발을 가능하게 한다. 본 연구에서 제시한 시스템은 연구성과정보 시스템과 연동되어 기존의 문제점들을 해결할 수 있다.

다음으로 정보검색 면에서는 온톨로지 기반의 지능적 정보검색(intelligent information retrieval)을 가능하게 한다. 기존의 검색시스템에서는 어휘나 용어가 시스템에 따라서 다르게 표현되며, 정보 수요자나 제공자가 서로 다른 관점으로 정보를 이해한다. 그러나 온톨로지 기반의 검색은 정보 검색자로 하여금 도메인 지식을 적용하게 하여 더욱 정교하고 강력한 검색을 가능하게 한다.

그러나 본 연구에서는 개념적 방향과 핵심적인 기술사항만을 검토하였을 뿐이며, 구체적인 구현에 있어서 발생될 문제점들에 대한 검토는 없다. 아직 시맨틱 웹에 대한 기술체계가 완성되지 않았으며, 이를 지원하는 툴도 충분하지 않다. 따라서 본 연구에서 언급된 사항들을 구현하기 위한 기술적인 세밀한 검토가 추후에 필요할 것이다.

참고문헌

- 이수상. 2004. 메타데이터의 상호운용성 보장을 위한 요소기술, 『한국도서관·정보학회지』, 34(1): 91-109.
- 이승희, 신문수, 정무영. 2003. RDFS+OWL을 이용한 생물학적 데이터의 지식표현과 추출, 『한국경영과학회/대한산업공학회 2003 춘계공동학술대회』, 1136-1141.
- 이재호. 2002. 시맨틱 웹 기술을 적용한 전자상거래 표준 운용체계연구, 『전자상거래표준화통합포럼』.
- Almeida, Mario do G.G. 1999. "Control Access for Grey Literature in Brazil: A Proposal." *Proceeding on the fourth International Conference on Grey Literature*.
- Auger, C. P. 1996. *Information Sources in Grey Literature*, 4th ed, London.
- Berners-Lee T., J Hendle, and O. Lassila. 2001. *The Semantic Web*, Scientific American.
- Dublin Core Metadata Initiative, 2005.[cited 2006. 5.10].
<www.dublincore.org>.
- Jeffery, K. 1999. "An Architecture for Grey Literature in a R&D." *Proceeding on the fourth International Conference on Grey Literature*.
- Lopatenko, A., A. Asserson, G. Keith, and K. Jeffery. 2002. "CERIF Information Retrieval of Research Information in a Distributed Heterogeneous Environment." *Proceeding on the 6th International Conference on Current Research Information Systems*.
- Maly, K., M. Zubair, and X. Liu. 2001. "Kepler An OAI Data/Service Provider for the Individual." *D Lib Magazine*, Vol. 7, No. 1, 2005.[cited 2006. 5.10].
<www.dlib.org/dlib/april01/maly/04maly.htm>.
- Miller, E., "An Introduction to the Resource Description Framework," 2005. [cited 2006. 5.10].
<www.dlib.org/dlib/may98/miller/05miller.html>
- Mooney, S. 2001. "Interoperability." *D Lib Magazine*, Vol. 7, No. 1, 2005.[cited 2006.5.10].
<www.dlib.org/dlib/january01/mooney/01mooney.html>.
- Steinacker A., A. Ghavam and R. Steinmetzm. 2001. "Metadata Standards for Web Based Resource", *IEEE Multimedia*, Vol. 8, Issue 1, pp. 70-76.
- W3C, "OWL Web Ontology Language Overview." [cited 2006. 5.10].
<www.w3.org/TR/owl_features/>.