

연구보고 2007-1

온톨로지 기반 학위논문 서비스를 위한
메타데이터 표준화와 모델링에 관한 연구

2007년 4월

서울대학교중앙도서관



온톨로지 기반 학위논문 서비스를 위한
메타데이터 표준화와 모델링에 관한 연구

연구책임자: 김홍기 (서울대학교 치의학과 교수)

공동연구원: 서정욱 (서울대학교 의학과 교수)

최현정 (ISO/JTC1 SC34 한국전문위원)

연 구 원: 서울대학교중앙도서관 정보관리과

장석일, 김미향, 김기숙

보조연구원: 송승재, 이유진

2007년 4월

서울대학교중앙도서관

국립중앙도서관 출판시도서목록(CIP)

온톨로지 기반 학위논문 서비스를 위한 메타데이터 표준화와 모델
링에 관한 연구 / 연구책임자: 김홍기; 공동연구원: 서정욱, 최현
정; 연구원: 장석일, 김미향, 김기숙; 보조연구원: 송승재, 이유진.
-- 서울: 서울대학교 중앙도서관, 2007
p. ; cm. -- (연구보고; 2007-1)

권말부록으로 "국내외 학위논문 서비스 및 KERIS 온톨로지 기반
서비스 사례" 수록
참고문헌 수록
ISBN 978-89-956410-5-7 94020 : 비매품
ISBN 978-89-985410-0-2(세트)

024.3-KDC4
025.3-DDC21

CIP2007001305

차 례

| | |
|---|----|
| 제1장 서론 | 1 |
| 제2장 온톨로지 표준화 현황 | 3 |
| 2.1 토픽맵의 표준화 | 3 |
| 2.1.1 토픽맵의 개요 | 3 |
| 2.1.2 토픽맵의 표준화 현황 | 5 |
| 2.2 RDF/OWL 표준화 | 12 |
| 2.2.1 RDF/OWL의 개요 | 12 |
| 2.2.2 RDF/OWL의 표준화 현황 | 15 |
| 제3장 온톨로지 토픽맵 기술을 사용한 모델링 방법론 | 23 |
| 3.1 토픽맵 온톨로지의 목적과 범위 | 24 |
| 3.2 분석 | 27 |
| 3.3 주요 개념 추출 및 정제 | 36 |
| 3.4 토픽맵 온톨로지 디자인 | 38 |
| 3.5 토픽맵 온톨로지의 정제 | 45 |
| 3.6 토픽맵 온톨로지 디자인 구현 | 49 |
| 3.7 인스턴스 생성 및 테스트 | 50 |
| 제4장 학위논문 온톨로지 구현방안 연구 | 56 |
| 4.1 토픽맵 기반 학위논문 콘텐츠 메타데이터 분석 연구 | 56 |
| 4.1.1 메타데이터 항목 추출 기반 기술조사 | 56 |
| 4.1.2 학위논문 콘텐츠 메타데이터 항목 추출 | 66 |
| 4.2 토픽맵 기반 학위논문 서비스 방법론 적용 | 70 |
| 4.2.1 메타데이터들의 상호연관성 분석을 통한 관계성 정의 | 70 |
| 4.2.2 학위논문 서비스 토픽맵 디자인 및 정제 | 73 |
| 4.2.3 학위논문 서비스 토픽맵 디자인 구현 | 79 |
| 4.3 토픽맵 기반 학위논문 서비스를 위한 기초DB 구축 | 83 |
| 4.3.1 토픽맵 기반 학위논문 서비스 기초DB 구축 방안 | 83 |
| 4.3.2 토픽맵 기반 학위논문 서비스 모델링 | 99 |

| | |
|---|------------|
| 4.4 토픽맵 기반 학위논문 서비스 의의 및 활용 방안 | 104 |
| 제5장 온톨로지 기반 디지털 도서관 고도화 전략 | 106 |
| 5.1 정보관리와 시맨틱 웹 | 106 |
| 5.2 시맨틱 디지털 도서관의 개념 | 107 |
| 5.3 시맨틱 디지털 도서관의 기능 | 108 |
| 5.4 시맨틱 디지털 도서관의 사례 | 109 |
| 5.5 시맨틱 디지털 도서관 구축을 위한 제언 | 112 |
| 제6장 결론 | 115 |
| 참고문헌 | 117 |
| [부록] 국내외 학위논문 서비스 및 KERIS 온톨로지 기반 서비스 사례 | 122 |
| 1. 국외 학위논문 서비스 사례 | 122 |
| 1.1 ProQuest Company & UMI | 122 |
| 1.2 NDLTD | 124 |
| 1.3 MIT | 128 |
| 1.4 UC Berkeley | 133 |
| 1.5 Harvard Univ. | 139 |
| 1.6 Cambridge Univ. | 144 |
| 1.7 Stanford Univ. | 148 |
| 1.2 Univ. of Tokyo | 150 |
| 2. 국내 학위논문 서비스 사례 | 151 |
| 2.1 KAIST | 151 |
| 2.2 포항공대 | 155 |
| 2.3 부산대 | 159 |
| 3. KERIS 온톨로지 기반 서비스 사례 | 162 |

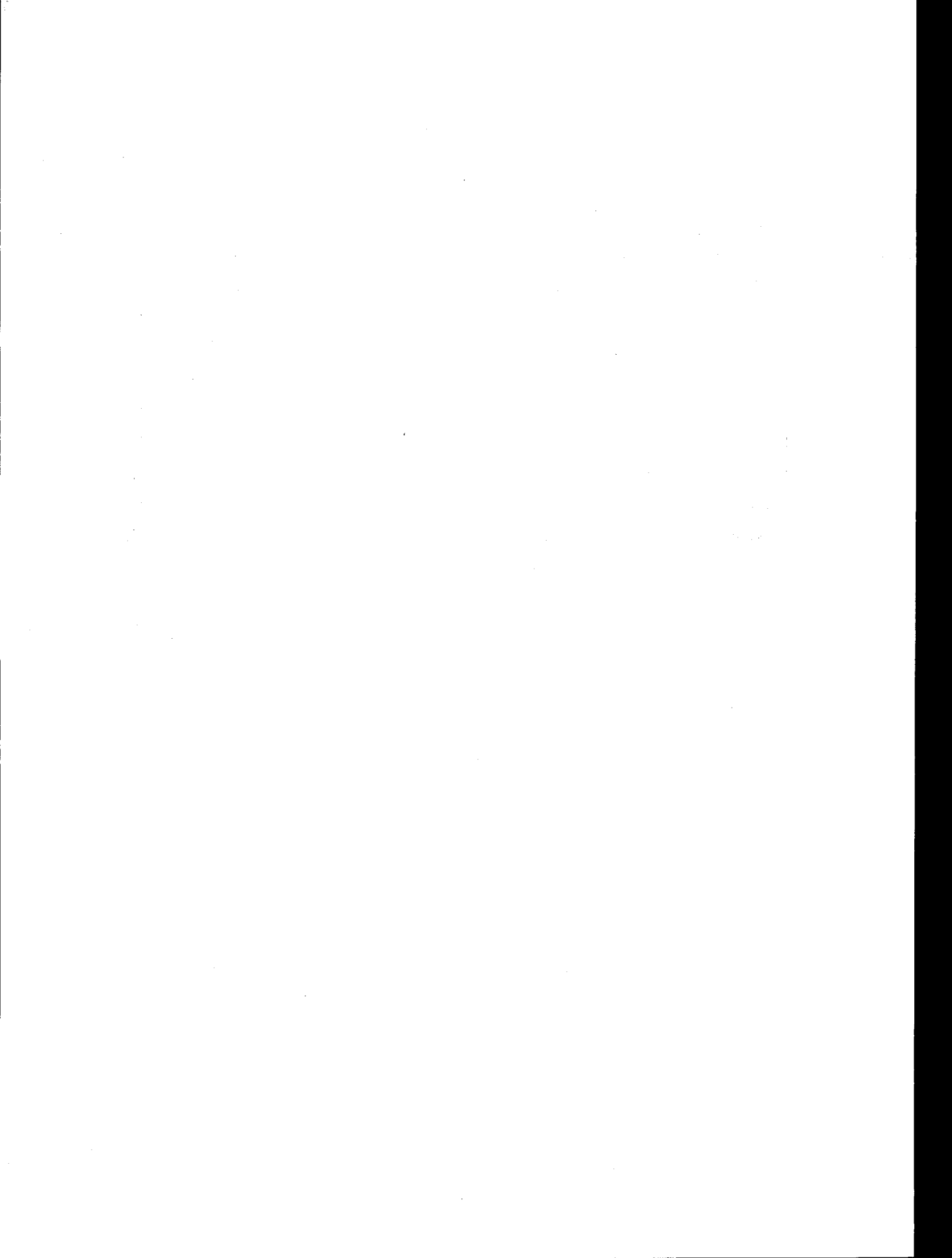
표 차례

| | |
|--|----|
| [표 1] W3C 1단계 표준화 목록 | 17 |
| [표 2] W3C 2단계 표준화 목록 | 18 |
| [표 3] 분류스키마와 토픽맵의 매핑요소 | 53 |
| [표 4] 관계형 데이터베이스와 토픽맵의 매핑요소 | 53 |
| [표 5] 더블린 코어의 구성 요소 | 57 |
| [표 6] 더블린 코어 요소와 MODS 성분 | 63 |
| [표 7] FRBR의 1집단 개체 | 66 |
| [표 8] 학위논문 콘텐츠 메타데이터 요소 | 68 |
| [표 9] 학위논문 서비스 개념의 분석 | 74 |
| [표 10] 토픽맵 요소의 DC 구성요소를 사용한 PSI 정의 | 77 |
| [표 11] 서울대학교 기초DB 구축 대상 | 83 |
| [표 12] MARCXML to 토픽맵 매핑 | 86 |
| [표 13] 전거레코드의 로컬 태그의 정의 제시 사례 | 89 |
| [표 14] 전거 작성시 필수 참고정보원 | 90 |
| [표 15] 전거레코드의 로컬 태그 사용의 사례 제시 | 91 |
| [표 16] 학과명 이력조사를 위한 참고정보원 | 94 |
| [표 17] 핵심어 조사를 위한 참고정보원 | 98 |

그림 차례

| | |
|---|-----|
| [그림 1] 토픽맵 표준 간의 관계 | 6 |
| [그림 2] TMQL과 SQL의 운용 | 7 |
| [그림 3] W3C 표준화 추진 현황 | 16 |
| [그림 4] 워킹 그룹별 표준개발 진행 현황 | 16 |
| [그림 5] Topic의 대칭 관계 예시 | 41 |
| [그림 6] n-ary Topic의 변환 도식 | 42 |
| [그림 7] 타입 계층구조 그림 예시 | 44 |
| [그림 8] 토픽 이름 구조 예시 | 47 |
| [그림 9] 학위논문 검색결과 상세화면 | 66 |
| [그림 10] 학위논문 MARC 데이터 사례 | 67 |
| [그림 11] 학위논문 국문초록, 영문초록, 목차 서비스 | 69 |
| [그림 12] 학위논문 원문서비스 | 70 |
| [그림 13] 학위논문 각 주제간의 관계성 | 71 |
| [그림 14] Topic Map ORM(Object Role Modeling) | 76 |
| [그림 15] 인물정보에 참조되는 전거레코드 사례 | 80 |
| [그림 16] 학위논문 서비스 토픽맵 모델링 | 81 |
| [그림 17] MARC 데이터의 XML 반출화면 | 85 |
| [그림 18] 온토폴리의 학위논문 Topic Type 설정 화면 | 85 |
| [그림 19] 저자, 지도교수 전거데이터 XML 반출 화면 | 87 |
| [그림 20] 인물정보를 매핑하여 등록한 예 | 88 |
| [그림 21] 전공학과리스트 열람 화면 | 92 |
| [그림 22] 학과명 Topic Type 설정화면 | 95 |
| [그림 23] 전공학과에 대한 토픽맵 서비스 화면 예시 | 96 |
| [그림 24] 온토폴리의 핵심어 등록설정 화면 예시 | 97 |
| [그림 25] 학위논문을 위한 기초DB구축 업무 구성도 | 99 |
| [그림 26] 2단계 토픽맵 기반 학위논문 서비스 구성 | 100 |
| [그림 27] 온토폴리 Topic Type 열람 화면 | 101 |
| [그림 28] 옴니게이트의 학위논문 토픽맵 요소 열람 | 101 |

| | |
|--|-----|
| [그림 29] Topic의 Occurrence 입력 화면 | 102 |
| [그림 30] Topic의 Association 선택 입력 화면 | 103 |
| [그림 31] 토픽맵 학위논문 서비스 화면 예시 | 103 |
| [그림 32] 토픽맵 vizigater | 104 |
| [그림 33] RDF의 예 | 107 |
| [그림 34] 도서관의 진화 | 108 |
| [그림 35] JeromeDL의 온톨로지 | 110 |
| [그림 36] MarcOnt Ontology의 주요개념 | 110 |
| [그림 37] FOAF의 예 | 111 |
| [그림 38] FOAF를 이용한 Social Semantic Collaborative Filtering 개념 | 111 |
| [그림 39] JeromeDL의 이용자 인터페이스 | 112 |
| [그림 40] 웹, Web 2.0, Semantic Web2.0의 특징 비교 | 113 |
| [그림 41] Social 시맨틱 디지털 도서관의 방향 | 113 |



제1장 서론

21세기 지식정보화 사회의 가장 큰 특징은 지식정보의 생산과 수요가 급격히 증가하고 지식정보의 유통 채널이 다양화 되었다는 것이다. 근대사회 이후 지금까지 지식정보의 저장과 보급에 있어서 가장 중심적 역할을 한 도서관은 혁신적 변화를 하지 않는다면 정보유통 패러다임의 변화와 더불어 심각한 정체성의 혼란에 빠지게 될 것이다. 특히 대학 도서관은 학문간 융합과 엄청난 양으로 생산되고 있는 연구 결과물의 유포에 중심적 역할을 지속적으로 하기 위하여 어떤 변화를 해야 하는지 심각하게 반성을 해야 할 시점이라 할 수 있다. 최근 디지털 도서관에 대한 새로운 시도는 도서관의 자기혁신적인 성찰로 볼 수 있다.

본 연구 과제는 대학 도서관이 디지털 도서관으로서의 서비스를 다양화 할 수 있는 방안에 대하여 연구하는 것을 목적으로 한다. 특히 학생들의 수요가 많은 학위 논문 서비스의 방법을 메타데이터 표준과 모델링의 관점에서 고찰하려고 한다. 최근 메타데이터 표준은 정보자원 간의 의미적 상호운용과 의미적 검색이 가능할 수 있도록 하는 기술을 지향하고 있다. 그 중심에 온톨로지라는 새로운 데이터베이스 기술이 있다. 기존의 관계형 데이터베이스 보다 의미에 대한 표현이 뛰어나며 정보자원 간의 관계를 다양하게 표현할 수 있도록 하는 기술이 온톨로지이다. 최근 차세대 웹이라 일컬어지는 시맨틱 웹의 표준 마크업 언어인 RDF/OWL과 ISO(국제표준화기구)의 표준 언어인 토픽맵(Topic Maps)은 의미적 메타데이터를 표현하는 가장 중요한 언어로 자리 잡고 있다.

서울대학교 도서관은 전자도서관 구축 사업을 통하여 도서관 보유 자원 및 서울대 내 가치 있는 학술정보를 디지털화하고 있다. 이렇게 구축된 다양한 콘텐츠는 유형·주제 또는 콘텐츠를 기술한 메타데이터에 따라 구축, 보존, 서비스되고 있다. 이 중 석·박사 학위논문은 매년 새로운 자료가 생산되며 이용도와 중요도가 높은 학술적 가치가 뛰어난 콘텐츠이다. 또한 학위논문은 다량의 학술지식 단위 정보 요소를 가지고 있어서 지도교수, 학과명, 핵심어(키워드) 등을 통한 주제접근이 용이하여 관련 주제로 접근할 수 있는 연계성이 우수하다는 장점을 가지고 있다. 이러한 학위논문의 우수성에 대응하여 다양한 이용자 접근점을 제공하고, 이를 토대로

한 자원간 통합 검색 서비스를 제공하기 위하여 온톨로지를 기반으로 한 학위논문 서비스 방법을 제안하고자 본 연구를 진행하게 되었다.

본 연구를 위해서는 온톨로지 기반의 학위논문 서비스를 성공적으로 수행하기 위한 메타데이터 항목 정의 및 정제가 필요하며 표준화 작업이 선행되어야 한다. 이런 메타데이터 표준화를 위해 토픽맵 및 RDF/OWL의 표준화 현황 분석 작업을 수행해야 하며, 이를 통해 메타데이터 항목 주제간의 의미적 연관성과 상호운용성을 확보하기 위한 모델링을 구축하고, 토픽맵을 사용한 모델링 방향을 제시하여 온톨로지 기반의 학위논문 검색시스템 구축에 대한 가능성을 살펴보고자 한다. 또한 실제로 시스템과 연동하여 운영할 수 있는 서비스 내용을 모델링한 후 평가 및 사전 검증을 거쳐 실 서비스에 적용할 수 있는 발판으로 삼는다. 더 나아가서는 이를 토대로 도서관이 구현해야 할 시맨틱 웹 기반의 도서관의 역할과 모습을 그려보고자 한다.

제2장 온톨로지 표준화 현황¹⁾

2.1 토픽맵의 표준화

2.1.1 토픽맵의 개요

토픽맵은 ISO가 제정한 표준으로 정보 자원들을 상호 연관성 있게 연결하고 조직하여 지식 구조를 기술할 수 있도록 하는데 목적이 있다. 이것은 대용량의 비구조화된 정보를 효율적으로 검색하기 위한 해결 방안으로 제안되었다.

- 토픽맵의 핵심 개념은 Topic(토픽), Association(연계), Occurrence(어커런스)로 구성되어 있다. 토픽맵의 핵심개념들을 간단하게 소개하면 다음과 같다.
- Topic : 토픽맵에서 Topic은 가장 일반적이면서도 포괄적으로 정의되어 있다. 즉, 토픽맵에서 지원하는 Topic은 사람, 개체(entity), 개념 등을 나타내며, 실제 존재유무와 상관없이 어떠한 것도 Topic으로 사용이 가능하다. 토픽맵 안의 Topic은 주제가 되는 지식단위를 하나의 인스턴스로 표현한다. 지식을 표현하는 Topic이 생성 된다는 것은 이 주제가 Topic을 통하여 기계가 이해하고 처리할 수 있는 객체(object)로 전환되는 것을 의미한다. 주제가 Topic으로 전환된 후에는 이 Topic을 다른 Topic과 많은 association을 생성할 수 있다.
- Topic Type(토픽타입) : Topic은 하나 이상의 Topic Type 즉, Topic 클래스(class)의 인스턴스(instance)일 수 있다. 예를 들어, “대구시”라는 Topic은 “시”라는 Topic Type의 인스턴스이며, “대한민국”이라는 Topic은 “나라”라는 Topic Type의 인스턴스이다. Topic과 Topic Type 간의 관계는 객체지향형 설계(object-oriented design)의 클래스-인스턴스 관계와 비슷하다. Topic Type도 토픽맵 안에서는 Topic으로 정의 된다. 즉, “나라”, “시”를 Topic Type으로 정의하고 타입의 인스턴스 요소로서 “대구시”, “대한민국” 등의 Topic을 생성한다. Topic Type은 정의되는 의미나 사용에 따라서 Role Type으로도 함께 등록할 수 있다.

1) 온톨로지 표준화 현황은 한국 정보문화진흥원에서 발간한 “국가지식정보 온톨로지 표준개발” 보고서의 [온톨로지 표준화 현황조사 및 분석]의 내용을 인용하였다.

- Occurrences(어커런스) : Occurrence는 Topic을 정보 자원과 연결시키는 역할을 한다. 토픽맵이 Topic에 관한 정보를 제공할 때, 각 Topic에 관한 Occurrence로 Topic에 관한 어떤 정보가 존재하는지를 확인할 수 있다. Occurrence의 값은 Topic과 자원을 연결하는 URI(Uniform Resource Identifier) 값을 지니거나, 문자열, number, date형의 값을 취한다. URI 값을 취할 때는 URL을 이용하여 자원과 Topic을 연결하며 주로 external occurrence의 값인 경우가 다수이다. Occurrence를 입력할 때에는 Occurrence Type을 정의하고, Topic에서 이 요소를 가져와 값을 입력하게 된다. 그러므로 Occurrence Type을 이용해 자원을 그룹화할 수 있다는 것이다. Occurrence Type의 예로는 URL값을 가지는 이미지, 삽화, 홈페이지 등과 문자열 값으로 나타내는 신문 기사, 비평, 정의, 설명, 초록 등이 있다.
- Association(연계): Association은 토픽맵 안에 정의된 Topic 간의 관계를 설정하는데 사용된다. 이를 통해 토픽맵 안에서 Topic이 갖는 문맥(context) 정보를 제공한다. Topic간의 관계 설정은 지식을 모델링 하는데 필수적인 요소이다. 하나의 관계가 연결할 수 있는 Topic의 수는 무한정이다. 귀납적 관계(자신과 자신과의 관계), 두 Topic간의 관계(binary), 세 Topic간의 관계(ternary) 등 어플리케이션의 필요에 따라 연결되는 Topic의 수가 다르게 설정 될 수 있다. 토픽맵 안에 정의된 Association은 방향성을 갖지 않는다. Association은 Topic사이에 존재할 수 있는 관계의 의미를 기술하는 것이다. 토픽맵에서 Association를 설정할 때, Topic간의 관계에서의 역할을 명명하여 이 부분을 분명케 한다.
- Class hierarchy(클래스 계층구조): Topic, Occurrence, Association의 인스턴스이며, 토픽맵에서 클래스 또는 Topic으로 정의된다. 클래스, Topic 또한 특정 클래스의 인스턴스로 정의될 수 있다. 결국, Topic은 클래스이며 또한 인스턴스로 정의될 수 있다는 것과, 특정 Topic이 클래스로 정의되기 위해서는 최소한 하나의 클래스 인스턴스가 존재해야 함을 의미한다. 이 표현의 의미는, “나라”라는 토픽 클래스가 되려면, 최소한 하나의 나라가 이 토픽의

인스턴스로 정의되어야 한다는 것이다. 토픽맵 제작시스템이 토픽 저자에게 새로운 토픽이 생성될 때마다 현존하는 Topic 클래스의 리스트를 제공한다면 큰 유익을 줄 것이다. 그러나 이러한 시스템은 클래스로 사용되고 있는 토픽 클래스만을 보여줄 수 있다. 즉, 클래스로 사용될 수 있으나 현재 사용되고 있지 않는 토픽은 저자에게 제공되지 않는다는 뜻이다. 이러한 토픽맵 시스템에 관련한 표준으로 사용될 것이다.

- 현재 가장 널리 사용되고 있는 LTM(Linear Topic Maps) 구문구조의 장점과 단점을 상세히 분석하고, 새 국제표준 구문구조에서 채택할만한 부분에 정하고, 전체 프레임워크 정립한 후에 기존 구문구조의 장점을 살리는 방향으로 정리할 필요가 있다.
- LTM처럼 많이 사용되고 있지는 않지만 호주의 Bond 대학 교수인 Robert Barta가 개발한 AsTMa의 구문구조도 상세하게 분석하여 장점을 제대로 정리한다.
- 새로 설계될 토픽맵 간결구문구조인 CTM(Compact syntax for Topic Maps)이 토픽맵 질의언어(Topic Map Query Language) TMQL과 토픽맵 제한언어(Topic Map Constraint Language) TMCL의 모든 요구사항을 충족시킬 수 있는지 면밀한 검토가 요구된다.
- CTM 구문구조는 현 XTM 구문구조를 보완할 것이고 다음과 같은 역할을 수행할 것으로 기대된다. (현재 제정 중인 ISO 13250 시리즈 중 Part 6로 예정 중) 수작업으로 토픽맵을 생성할 때 용이하고 간결한 구문구조를 제공하며, 사람이 용이하게 해독할 수 있는 토픽맵 구문구조를 제공할 것이다. 또한 TMCL과 TMQL을 위한 보편적이면서 가벼운 구문구조를 제공할 것이다.

2.1.2 토픽맵의 표준화 현황

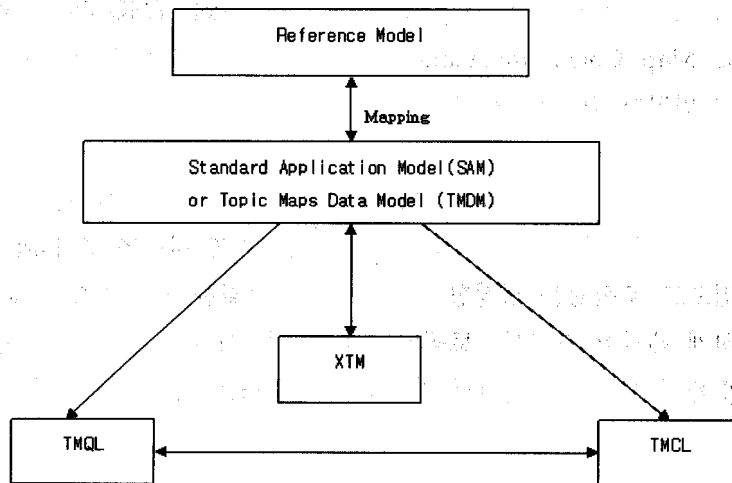
토픽맵은 ISO가 제정한 표준으로, 정보 자원간의 상호 연관성을 도출, 연결 및

조직하여 효과적인 지식 구조를 기술할 수 있도록 하는데 목적이 있다. 토픽맵은 대용량의 구조화되지 않은 정보를 효율적으로 검색하기 위한 해결 방안으로 제안되었다.

처음 ISO/IEC 13250에서 제안한 토픽맵 표준 스펙은 SGML(Standard Generalized Markup Language) 구조와 HyTM언어였으나, 2001년에 Topicmaps.org에서 개발한 URI와 XML을 기반한 XTM(XML Topic Maps)을 통합하면서 지금은 XTM 1.0과 HyTM(HyTime Topic Maps)이 토픽맵 구문을 제공하는 표준으로 제안되어 있다. 토픽맵에 관련한 표준은 여러 부분으로 나누어져 있으며, 여기에서는 이들 표준들 중에 가장 중요한 5가지에 대하여 기술하기로 한다.

- 토픽맵 질의 언어(Query Language)
- 토픽맵 제한 언어(Constraint Language)
- 토픽맵 표준 어플리케이션 모델(Standard Application Model)
- 토픽맵 참조 모델(Reference Model)
- XTM(XML Topic Maps)

위의 5개 표준간의 관계를 도식화하면 다음과 같다.



[그림 1] 토픽맵 표준 간의 관계

1) 토픽맵 질의 언어 (TMQL, ISO 18048: Topic Maps Query Language)

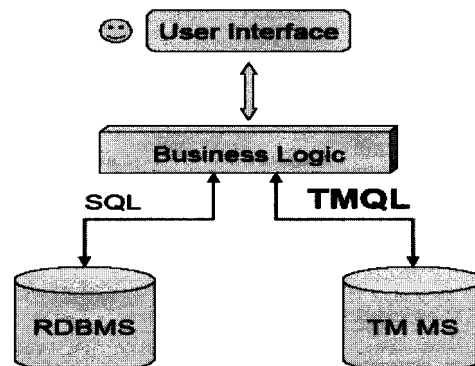
TMQL은 토픽맵을 위한 질의 언어 표준으로 현재 미완성이며 TMQL의 기능에 관한 필요 요건(ISO/IEC JTC 1/SC34 N249)만이 완성되어 있다. TMQL은 RDBMS(Relational Database Management System)의 SQL이나 XML관리 시스템의 XML Query와 같은 기능을 할 것이며 토픽맵 관리 시스템에 저장된 토픽맵의 생성/수정에 관련한 표준화된 질의를 제공할 것이다.

TMQL은 두 부분으로 구성되어 있으며 Part 1은 토픽맵에 대한 질의방법을 정의하고, Part 2는 토픽맵을 생성하고 갱신하는 방법을 정의한다. 또한 인터넷과 같은 분산 환경 하에 있는 토픽맵 응용프로그램에 대한 부분도 다룬다.

TMQL 기능요건을 요약하면 다음과 같다.

- 정보 추출과 생성을 위한 토픽맵 정보에 대한 작업
- 토픽맵 데이터 모델에 기반(SAM)
- 간결하며 인간가독형인 구문체계
- 추상적 질의 대수기능 지원(Abtract query algebra)
- 다양한 플랫폼과 응용프로그램 지원

현재 토픽맵 표준 질의 언어의 후보자로 제안된 것으로는 Ontopia의 Tolog, Empolis의 eTMQL, Ann의 LTM(Linear Topic Map) 기반 strawman과 XPath, XML Query 등이 있다. 아래의 그림은 TMQL이 3-tier 구조에서 SQL과 운용되는 모습이다.



[그림 2] TMQL과 SQL의 운용

2) 토픽맵 제한 언어 (TMCL, ISO 19756: Topic Maps Constraint Language)

TMCL은 토픽맵 스키마와 SAM에 의해 정의된 토픽맵 모델에 대한 제한 언어 표준이다. 이 제한 언어는 정식 제한 언어, 조작 의미(operational semantics) 및 구문 체계를 제공할 것이다. TMCL은 현재 미완성이며, TMCL 기능에 관한 필요 요건 (ISO/IEC JTC 1/SC34 N226)만이 완성되어 있다.

TMCL은 다음의 목적을 갖는다.

- 토픽맵 클래스의 구조와 의미에 관한 문서화
- 특정 도메인과 연계된 토픽맵 어플리케이션과 하나 이상의 도메인에 걸쳐 정의된 토픽맵 어플리케이션에 기초를 제공
- 특정 토픽맵 안에서의 일관성은 물론 여러 토픽맵 사이의 일관성을 보장하기 위한 검증 방법 제공
- 응용프로그램의 토픽맵 생성 및 관리를 위해 보다 직관적인 사용자 인터페이스 제공
- 토픽맵 모델링과 토픽맵 정보 입력의 이중화

TMCL은 토픽맵 표준에 상응하는 토픽맵 클래스에 대한 제한사항을 표현하는 방법을 제공한다. TMCL의 목적은 특정 토픽맵이 토픽맵 안에서 또는 토픽맵 클래스 안에서 정해진 규칙을 준수한다는 것을 표현할 수 있는 언어를 제공하는 것이다. 또한 TMCL은 토픽맵 저장의 최적화와 스키마 정보에 기반을 둔 토픽맵 질의 언어의 (TMQL) 질의 최적화를 용이하게 한다. 또한 사용자가 토픽맵을 생성하거나 유지하는데 있어 편리한 사용자 인터페이스를 제공하는데 도움을 줄 수도 있다.

TMCL은 토픽맵 표준 어플리케이션 모델 SAM(Standard Application Model)에 기반을 두어 기술되고, 어떠한 특정 저장 형식에 국한되어 기술되지 않는다. TMCL은 XTM, HyTM(HyTime Topic Maps), LTM(Linear Topic Maps), AsTMa=를 지원하며, 이들 언어들 모두 SAM으로 매핑이 가능하다.

TMCL 제한 사항이란, “주어진 특정 기준 (즉, 선택자, Selector)을 만족하는 정보 항목은 반드시 두 번째 기준인 한정사항(restriction)을 만족해야 한다.” 라는 요구 사항을 의미한다. 특정 토픽맵에 주어진 한정사항은, 토픽맵 구성체의 의미

를 제한하거나, 토픽맵 모델 안에 정의되지 않은 수적 제한(cardinality)과 데이터 타입을 정의할 수도 있다. 또한 TMCL 제한 사항은 타입핑(typing) 유형을 명시함으로써 토픽맵 모델을 확장 할 수도 있다.

“TMCL 스키마”는 TMCL을 이용하여 정의된 제한 사항의 집합으로 정의된다. TMCL 스키마는 특정 어플리케이션에 허용된 입력 정보를 정의하는 제한 사항에 관한 어휘를 정의하는 목적으로 쓰일 수 있다.

“타입핑 유형 시스템”은 하나의 토픽맵이나 통합된 토픽맵 집합 안에서 적용되는 타입들의 정의들을 지칭한다.

“스키마 자가 분석(schema introspection)”은 스키마 정의를 반입하여 (혹은, 새로 생성하여) 조작하거나, 필요한 정보를 접근 하고, 다시 반출할 수 있는 기능을 말한다.

“제한 구상화(constraint reification)”는 토픽맵 안에서 제한 사항을 토픽으로 표현할 수 있는 기능을 말한다. (이것은 “시스템” 토픽맵이 될 수 있고, 이 경우, 제한 사항이 적용되는 “일반적” 토픽맵과는 구별된다.) “제한 구상화”는 제한 사항에 관한 여러 특징의 지정을 가능케 한다. 또한 제한 사항의 타입을 정의할 수도 있다.

3) 토픽맵 데이터 모델(TMDM ISO 1325 or SAM(Standard Application Model))

2003년 5월 런던에서 개최된 SC34총회 및 WG회의는 N396을 ISO 13250-2: Topic Maps - Data Model의 첫 working draft로 채택하였다. N396의 공식적인 명칭은 "The Standard Application Model for Topic Maps Model"이다. SAM은 XML 정보 집합 형식과 UML을 이용하여 토픽맵의 표준 데이터 모델 및 관련 의미를 산문체로 기술한다.

SAM은 토픽맵 구조에 관한 정의는 물론, 토픽맵 통합, 중복 제거 등의 작업, 그리고 공적 주제(published subjects)를 정의한다. 토픽맵 구문체계(XTM, HyTM)는 SAM을 기반으로 재정립 될 것이며, TMQL과 TMCL은 SAM을 기반으로 할 것이다.

SAM은 정보 항목 타입(Information item types)과 속성의 정의를 통해 토픽맵 구조를 정의한다. 정의된 정보 항목 타입에는 위치(locator) 항목, 소스 위치, 토픽

맵 항목, 토픽 항목, 토픽 이름 항목, 별칭(variant name) 항목, 어커런스 항목, 관계 항목, 관계 역할 항목이 있다.

SAM은 또한 통합이 가능한 요소들과 이들의 통합 절차를 기술한다. 통합이 가능한 요소로는 Topic, 토픽 이름, 별칭, Occurrence 항목, Association 항목, Association 역할 항목, 위치 항목이 있다.

SAM에 기반한 새로운 버전의 XTM과 HyTM 구문체계는, XTM이나 HyTM구문으로 쓰여진 문서의 SAM 모델로의 변환을 기술할 것이다. 이를 통해 SAM은 이 두 구문체계의 참조 모델이 될 것이며, 이들 구문 체계간의 비교는 이들이 생성하는 SAM 모델들을 비교함으로써 가능케 된다.

또한, SAM은 새로운 토픽맵 구문체계가 개발되었을 때, 이들 체계와 표준 토픽맵 구문 체계를 연계할 수 있게 한다. 또한, SAM을 통해 토픽맵 이외의 구문체계 (예: NewsML)로부터의 매핑도 가능하게 될 것이다.

4) 토픽맵 참조 모델 (Topic Maps Reference Model)

온톨로지 및 기타 어휘의 목적은 인간으로 하여금 관심 있는 사항을 특정 도메인이나 그룹 안에서 의미를 효과적으로 전달할 수 있게 하는 것이다. 그러나 특정 도메인이나 그룹에서 제작된 지식을 그룹 밖에서도 유용하게 쓰는 방법은 전혀 다른 문제이다. 이것에 대한 해답은 이러한 어휘나 지식들이 상호 참조될 수 있는 모델을 구성하는 것이다. 토픽맵 참조 모델은 토픽맵 분야에서 이러한 질문에 대한 해답으로 제안된 것이다.

2003년 5월 런던에서 개최된 SC34총회 및 WG회의는 토픽맵 참조 모델(N393)을 'Topic Maps - Reference Model'의 첫 working draft로 채택하였다.

토픽맵 참조모델의 기본개념은 다음 두 가지에 대한 구분에서 출발한다.

- 모든 지식이 포함하고 있어야 하는 기본적인 구조적 특성. 이러한 특성을 통해 모든 종류의 지식을 유용하고 세밀하게 통합 할 수 있다.
- 위의 특성 이외에, 지식이 포함할 수 있는 특성. 이러한 특성은 상황에 따라, 다양한 방식으로 제한될 수 있는 특성이다.

토픽맵 참조 모델은 모든 지식에 공통으로 존재하는 최소한의 구조적 특성의 집합은 결국 시맨틱 네트워크 구조에 대한 제한 사항으로 표현될 수 있다고 본다. 참조 모델은 이러한 시맨틱 네트워크를 토픽맵을 표현하는 추상적 “그래프”라고 정의한다. 이 그래프 모델은 이름과 Occurrence를 Topic과 같은 수준의 노드(node)로 나타내며, 관계를 이용해 Topic과 연계시킨다. 참조 모델은 SAM보다 더욱 간단한 모델 구성체(construct)로 이루어져 있으며 확장성이 높다.

토픽맵 참조 모델은 다른 지식 표현 체계(예: RDF)와 토픽맵의 관계를 설명하는 기반을 제공하여 다른 지식 표현 체계와의 상호 운용을 용이하게 한다.

5) XTM(XML Topic Maps) 1.1

XTM은 초기에 Michel Biezunski와 Steven R. Newcomb이 표준화 작업을 수행하였고, 현 규격은 TopicMaps.Org란 독립적 컨소시엄의 생산물이다. 토픽맵 패러다임의 기원은 Davenport Group에서 처음으로 소개하였던 1993년이다. 이후 IDEAlliance에서 HyTime의 응용에 대한 규약으로 개발되었고, 이후 독자적으로 개발되어 공표되었다.

2000년 초반, 토픽맵은 처음으로 ISO 국제 표준인 ISO/IEC 13250:2000으로서 완전히 공식화된다. 그 이후 바로, 이 토픽맵의 World Wide Web에서의 응용성을 개발하기 위해 TopicMaps.Org가 구성되었고, 정보의 검색과 관리능력을 개선하기 위한 많은 노력이 진행되고 있다.

XTM 1.1의 특징은 다음과 같다.

- 토픽맵을 위한 XML 구문체계
- XTM 1.0과 호환가능하나 부분적으로 확장
- DTD, RELAX-NG 스키마와 W3C XML 스키마를 제공

XTM 디자인 목표는 다음과 같다.

- XTM은 인터넷에 바로 사용가능하여야 한다
- XTM은 다양한 어플리케이션을 지원할 수 있어야 한다
- XTM은 XML, XLink, 그리고 ISO 13250과 호환성이 있어야 한다
- XTM 문서를 처리하는 프로그램을 작성하기 쉬어야 한다

- XTM에서 선택적 특성의 수는 이상적으로는 0이고 가장 최소화되어야 한다
- XTM 문서는 인간이 읽기 쉽고 논리적으로 명확하여야 한다
- XTM 디자인은 공식적이고 간결하여야 한다
- XTM 문서는 생성하기 쉬워야 한다
- XTM 마크업에서 간결함은 최소한으로 지켜져야 한다

토픽맵은 지식관리에 매우 새롭고 효과적인 패러다임을 제공하는 기술이다. 정교하게 개발된 토픽맵은 지식을 유통하는 기관이나 사용자 모두에게 새로운 관리, 사용에 대한 경험을 제공한다. 그러나 앞서 기술한 것과 같이, 토픽맵 표준은 현재 진화하고 있고, 경험이 축적됨에 따라 토픽맵 구현방법에 있어 앞으로도 새로운 접근 방식과 프로세스가 개발될 것이다.

2.2 RDF/OWL 표준화

2.2.1 RDF/OWL의 개요

1) RDF/OWL 소개

W3C의 시맨틱 웹 활동이 제공하는 중요 표준인 RDF(Resource Description Framework), RDF 스키마(RDF Schema) 및 OWL(Web Ontology Language)은 정보의 표현과 교환에 관한 것이다.

현재 시맨틱 웹 표준은 두개의 층(layer)으로 이루어져 있다. 첫 번째 층인 데이터 층은 RDF로 이루어져 있으며, RDF는 기계가독형 정보의 교환과 통합을 위한 데이터 표현 방식을 제공한다. 두 번째 층인 온톨로지 층은 RDF 데이터로 사용되는 어휘를 기술하는 것과 관계한다. 온톨로지 층은 간단한 RDF 스키마 언어와 보다 상세한 웹 온톨로지 언어인 OWL을 지원한다.

| | |
|--------|--------------|
| 온톨로지 층 | RDF스키마 / OWL |
| 데이터 층 | RDF |

RDF/OWL의 강점

- 정보 통합 및 공유 어휘의 재사용 지원
- 구조화되지 않은 데이터 처리

- 데이터 모델링과 구문체계의 분리
- 확장성
- 세밀한 구문체계에 기초한 추론과 분류 기능
- 그래프 구조 모델링을 통한 표현방식의 융통성

RDF/OWL의 약점

- 문서 검증 능력 및 성능
- 학습하는데 많은 시간 소요
- 자원 표현 기능의 한계 (정 자원에 부여된 여러 특성을 연계시키는데 한계가 있음)

RDF/OWL은 여러 곳에 분산된 데이터를 통합하는 분야나 시스템을 통합하여 상호운용성을 극대화하기 위해 공유된 어휘를 공표하여야 하는 분야에 그 장점이 가장 잘 부각되고 있다.

2) RDF 표준

RDF(Resource Description Framework)는 W3C에서 제정한 표준으로서 자원에 대한 메타데이터를 기술하는 XML 기반의 표준이다. RDF는 자원의 기술단위로써 자원을 나타내는 주어(subject), 속성을 나타내는 술어(predicate), 속성의 값을 나타내는 목적어(object)의 세 부분으로 이루어진 triple (즉, RDF 선언(statement))을 사용한다. 자원은 기술하고자 하는 정보 자원을 의미하고, 속성은 자원의 특성을 표현하는데, 속성을 통해서 자원과 속성값 사이의 관계를 나타낸다. 예를 들어, "The coffee has the color black" 기술하고자 할 때, RDF에서는 "coffee" 자원으로, "has the color" 속성으로, "black" 속성값으로 표현된다. RDF triple은 자원에 대한 보다 정교한 표현을 가능케 하고, 속성을 통해 정보 자원에 대한 관계 설정을 자유롭게 할 수 있다.

RDF에서 자원들은 URI(Uniform Resource Identifier)로 식별될 수 있다. 어떤 자원들은 무명(unnamed) 자원으로 식별이 불가능하다. 속성도 자원으로 구분되어서 URI를 통하여 식별될 수 있다. 속성값은 자원이거나 값을 가지는 문자열

(Unicode string)로 정의된다.

RDF는 그 자체만으로는 속성에 대해 설명할 방법을 제공하지 않을 뿐 아니라, 속성간의 관계나 다른 자원간의 관계를 표현할 방법 또한 제공하지 않는다. RDF 어휘 기술 언어(Vocabulary Description Language)인 RDF 스키마는 이러한 기능을 수행한다. 즉 RDF 스키마는 클래스, 속성, 자원을 기술하는데 사용할 수 있는 클래스와 속성을 정의한다.

RDF 관련 표준은 아래와 같고, 모두 공식 표준으로 채택되었다.

- RDF/XML 구문체계 (RDF/XML Syntax Specification): RDF를 XML 구문체계로 정의한 문서
- RDF Vocabulary Description Language 1.0: RDF Schema: RDF를 이용하여 RDF 어휘를 정의하는 방법을 기술한 문서
- RDF Primer: RDF 입문자를 위한 기초 입문서
- Resource Description Framework(RDF): Concepts and Abstract Syntax: RDF의 기초를 이루는 개념적 구문체계를 기술하고, 이 구문체계를 RDF의 공식 의미 체계(semantics)와 연계하는 목적에서 기술된 문서
- RDF Semantics: RDF와 RDF 스키마의 공식 의미체계와 이에 병행하는 추론 규칙을 기술한 문서
- RDF Test Cases: RDF 테스트 용례를 기술한 문서

3) OWL 표준

시맨틱 웹의 실현을 위해, 데이터 표현 단계인 RDF를 보조하기 위한 온톨로지 언어의 개발이 필요하게 되었다. 현재 인터넷 커뮤니티에서 가장 주목하고 있는 온톨로지 언어는 OWL이며, OWL의 출현은 RDF 스키마가 제공하는 기본적인 기능을 넘어서, 기계가 자원에 대한 추론(reasoning)을 가능하게 하도록 하는데 그 목적이 있다. OWL은 RDF와 RDF스키마를 기반으로, 속성과 클래스의 기술에 있어 더욱 풍부한 어휘를 제공한다.

온톨로지 언어인 OWL의 어플리케이션 분야는 다음과 같다.

- 인터넷 포털: 분류 법칙을 이용한 검색 기능 확장

- 멀티미디어: 멀티미디어 내용에 기반을 둔 검색
- 기업 웹 사이트 관리: 데이터와 문서의 자동 텍사노미 생성, 기업부서나 기업 간의 합병에 따른 데이터 매핑
- 지능형 에이전트: 사용자 선택사항 및 관심사항 표현, 웹사이트 간의 콘텐츠 매핑
- 웹 서비스 및 유비쿼터스 컴퓨팅: 웹서비스 발견 및 구성, 저작권 및 접근 관리

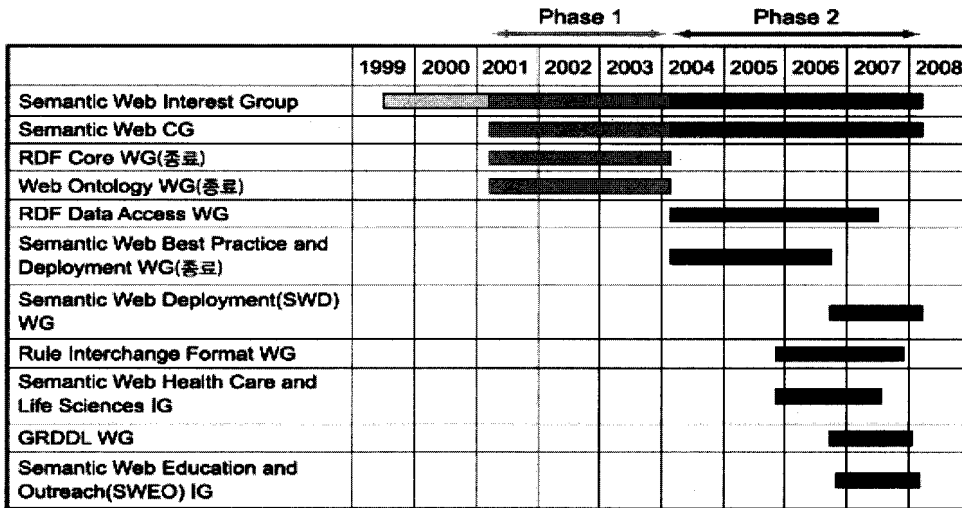
OWL은 RDF스키마를 확장하여 RDF스키마 클래스간의 복잡한 관계나, 클래스 및 속성에 관한 정교한 제한 사항의 표현을 가능케 한다. 예를 들어, 1)클래스의 속성 수나 타입을 제한할 수 있고, 2) 특정 속성을 포함하는 정보가 특정 클래스에 속함을 추론할 수 있으며, 3) 특정 클래스의 모든 멤버가 특정 속성을 지님을 결정할 수도 있고, 웹에 존재하는 문서들에 정의된 클래스간의 상호 관계 또한 표현할 수 있다.

OWL과 관련된 표준 문서는 현재 6종이 발표되었다.

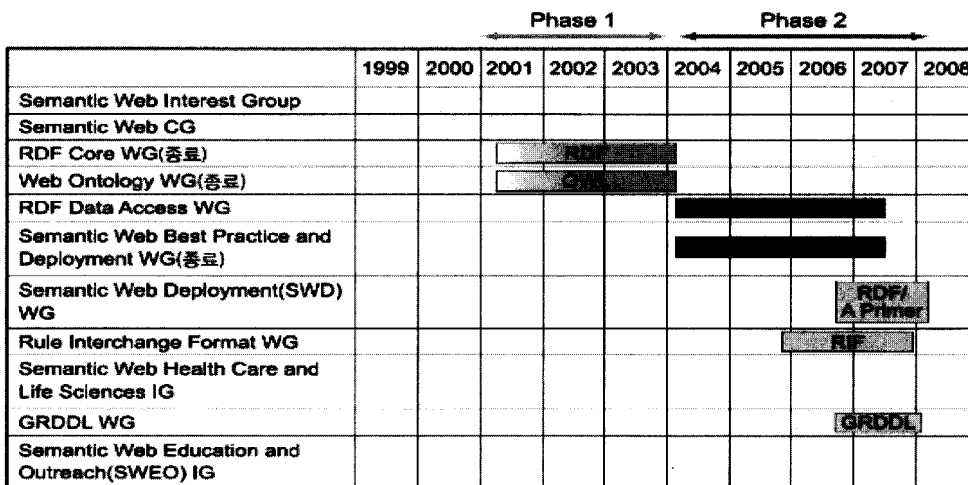
- User cases and requirements 문서
- Overview 문서: OWL의 특성에 관한 개략적 소개 및 사용 방법 소개
- Guide: OWL 특성을 이용한 많은 예를 상세히 설명
- Reference 문서: 모든 OWL 특성에 관한 참고 문서
- Test case 문서와 test suite: 표준 준수 테스트를 위한 100여개의 샘플 테스트 제공
- Semantics of OWL 및 OWL에서 RDF로의 매핑 관련 문서

2.2.2 RDF/OWL의 표준화 현황

1) W3C 표준화 추진 현황



[그림 3] W3C 표준화 추진 현황



[그림 4] 워킹 그룹별 표준개발 진행 현황

[그림 3]에서 보는 바와 같이 표준화 1단계는 OWL과 RDF언어에 대한 표준화에 초점을 맞추어 진행되었고, 이를 위해 RDF Core 워킹그룹과 Web Ontology 워킹그룹을 구성하여 작업을 하였다. 그리고 그 결과 2004년 2월 총 12건의 권고 표준을 확정/공표하였다.

1단계 표준화를 마친 후 바로 2004년 2월부터 시맨틱 웹 2단계 표준화 작업이 시작되었고, 2단계의 주요 목표는 실질적인 응용의 발굴과 개발 경험의 공유에 초점을 맞추고 있다.

[그림 4]는 워킹 그룹별 표준 개발 진행 현황을 나타낸 그림이다.

2) W3C 1단계 표준화

[표 1] W3C 1단계 표준화 목록

| Title | Date |
|--|-----------|
| RDF Primer | 2004.2.10 |
| RDF Semantics | |
| RDF Test Cases | |
| RDF Vocabulary Description | |
| Language 1.0 : RDF Schema | |
| RDF/XML Syntax Specification | |
| Resource Description Framework(RDF):Concepts and Abstract Syntax | |
| OWL Web Ontology Language Guide | |
| OWL Web Ontology Language Overview | |
| OWL Web Ontology Language Reference | |
| OWL Web Ontology Language Semantics and Abstract Syntax | |
| OWL Web Ontology Language Test Cases | |
| OWL Web Ontology Language Use Cases and Requirements | |

3) W3C 2단계 표준화

현재 2단계 표준화 작업은 RDF와 OWL을 이용한 응용 및 구현에 초점이 맞춰져 있다.

[표 2] W3C 2단계 표준화 목록

| Title | Date | WorkingGroup |
|---|---------|--------------------|
| RDF Data Access Use Cases and Requirements | 2005.3 | RDF Data Access WG |
| SPARQL Protocol for RDF | 2006.4 | RDF Data Access WG |
| SPARQL Protocol for RDF Using WSDL 1.1 | 2005.1 | RDF Data Access WG |
| SPARQL Query Language for RDF | 2006.10 | RDF Data Access WG |
| SPARQL Query Results XML Format | 2006.4 | RDF Data Access WG |
| Gleaning Resource Descriptions from Dialects of Language(GRDDL) | 2005.5 | GRDDL WG |
| RIF Use Cases and Requirements | 2006.7 | RIF WG |
| Semantic Annotations for WSDL | 2006.7 | SAWSDL WG |
| Image Annotation on the Semantic Web | 2006.3 | SWBPD WG |
| RDF/OWL Representation of WordNet | 2006.6 | SWBPD WG |
| Quick Guide to Publishing a Thesaurus on the Semantic Web | 2005.11 | SWBPD WG |
| Best Practice Recipes for Publishing RDF Vocabularies | 2006.3 | SWD WG |
| RDF/A Primer 1.0 ; Embedding RDF in XHTML | 2006.3 | SWD WG |
| SKOS Core Vocabulary Specification | 2005.11 | SWD WG |
| SKOS Core Guide | 2005.11 | SWD WG |

4) W3C 워킹그룹 별 역할

가) RDF Data Access 워킹그룹

2004년 2월에 그룹을 만들어 2004년 10월에 SPARQL 디자인 초안을 작성하고, 현재 진행 중에 있으며 2007년 5월까지 프로젝트를 마무리할 계획을 가지고 있다.

RDF Data Access 워킹그룹은 RDQL, XQuery, RuleML, N3와 같은 다양한 질의 언어를 기반으로 시맨틱 웹을 위한 표준 질의 언어인 SPARQL(Standard Protocol and RDF Query Language) 개발을 목표로 한다.

진행중인 표준화 목록

- RDF Data Access Use Cases and Requirements
- SPARQL Protocol for RDF
- SPARQL Protocol for RDF Using WSDL 1.1
- SPARQL Query Language for RDF
- SPARQL Query Results XML Format

나) SWBPD(Semantic Web Best Practice and Deployment) 워킹그룹

2004년 3월에 처음으로 회의를 시작하였고, 2006년 1월에 종료 되었다.

시맨틱 웹 개발 경험들을 공유할 수 있도록 하기 위한 각종 개발 가이드라인 제공을 목표로 하고 있으며, 이를 위해 산하에 여러 개의 태스크 포스를 두고 있다.

태스크 포스

- ADTF - Applications and Demos
- HTML - Embedding RDF in HTML
- MM - Multimedia Annotation
- OEP - Ontology Engineering and Patterns
 손쉬운 온톨로지 개발을 위한 가이드라인 제공을 목표
- PORT - Porting Thesaurii to RDF and OWL
- RDFTM - RDF/Topic Maps Interoperability
- SE - Software Engineering
- VM - Vocabulary Management
- WordNET - 워드넷을 RDF/OWL로 변환하기 위한 태스크 포스

진행중인 표준화

- Image Annotation on the Semantic Web
- RDF/OWL Representation of WordNet
- Quick Guide to Publishing a Thesaurus on the Semantic Web

다) SWD(Semantic Web Deployment) 워킹그룹

2006년에 시작하여 2006년 4월 말까지 약 20개월의 기간을 가지고 업무를 진행하고 있다.

기존 SWB WG의 활동 중 Deployment에 관한 이슈만을 추가적으로 작업하기 위한 목적으로 만들어졌다. 이 중에서도 OWL 사용, RDF와 HTML 문서 통합 등과 같은 실제적인 RDF 개발과 도입에 초점을 맞춘 표준 개발을 목표로 하고 있다.

SWD 워킹그룹의 작업을 나열하면 다음과 같다.

- Simple Knowledge Organization System(SKOS) 권고안 작성
- 웹 프로토콜을 사용하는 RDF 어휘와 OWL 온톨로지 publish하기 위한 가장 적합한 recipes을 만든다
- RDF schema와 OWL 온톨로지 저자들을 위한 기술적 가이드라인을 제공한다
- 어플리케이션간에 시맨틱 통합을 달성하는데 OWL을 사용하는 기술적 보고서를 작성한다
- HTML문서에서 RDF를 접목시키는 방법을 서술한 기술적 보고서를 작성한다

진행중인 표준화

- Best Practice Recipes for Publishing RDF Vocabularies
- RDF/ A Primer 1.0 ; Embedding RDF in XHTML
- SKOS Core Vocabulary Specification
- SKOS Core Guide

5) 워킹그룹 표준화 내용

가) RDF Data Access Use Cases and Requirements

RDF 정보를 질의하기 위한 표준화된 언어가 없고 내부 RDF 서버나 원격으로 접속하기 위한 접근 표준 프로토콜도 없다. 그러나 많은 개발자들이 많은 RDF 질의 언어를 자체적으로 만들고 있다.

이 문서는 Use case, 요구사항을 기술하고 RDF 질의 언어와 데이터 접근 프로토콜에 대한 사항을 기술함에 목적이 있다. 이 문서는 RDF 질의 언어와 데이터 접근 프로토콜을 활용하여 웹 출판, 개인 정보관리, 운송, 관광과 같은 영역에서 시맨틱 웹 어플리케이션을 유용하게 사용할 수 있도록 하는 방법을 제시한다.

Use case는 RDF 질의 언어나 접근 프로토콜 작성에 있어 중요한 특징이나 공통의 모티베이션(motivation)을 따른다.

나) SPARQL Protocol for RDF

이 문서는 (질의 클라이언트부터 질의 프로세서까지 SPARQL을 전달하는 의미의) SPARQL Protocol을 기술한다.

SPARQL 프로토콜은 SPARQL Query Language for RDF와 호환성을 가지도록

설계 되었다. SPARQL 프로토콜은 두 가지 방법으로 기술되었다. 첫째는, 어떤 명확한 실행의 독립적 인터페이스 구현, 혹은 다른 프로토콜과의 묶임으로서 기술되었다. 둘째는, 인터페이스와 HTTP, SOAP과 묶임으로서 기술되었다. WSDL과 W3C XML 스키마 문서와 잘 연립되어진 이 문서는 SPARQL 쿼리 서비스와 의뢰인의 도구를 개발하는 소프트웨어 개발자를 위한 의도를 우선적으로 가진다.

다) SPARQL Protocol for RDF Using WSDL 1.1

RDF를 위한 SPARQL Protocol은 HTTP나 SOAP와 관련된 인터페이스, 타입, 결점, 운영과 관련한 WSDL 2.0 문서를 통해 추상적으로 기술되어 있다.

RDF Data Access 워킹그룹은 추가적인 SPARQL 프로토콜의 WSDL 1.1 버전을 발간하기로 결정했다. 워킹그룹의 목적은 SPARQL 프로토콜에 의해 WSDL 2.0에서 동등한 표준으로 정의하는 것과 동등하게, WSDL 1.1 문서에서 비형식의 인터페이스, 타입, 결점, 운영을 정의하는 것이다.

라) SPARQL Query Language for RDF

RDF 그래프는 subject, predicate, object로 구성되어져 있다. RDF 그래프는 RDF Concepts and Abstract Syntax [CONCEPTS]에 정의되어 있으며, RDF 그래프를 구성하는 이 세 가지 요소는 다양한 소스로부터 도출될 수 있다. 예를 들어, RDF 문서에서 직접 도출될 수도 있으며, 다른 RDF 트리플로부터 추론될 수도 있다. 또는 다른 형식(예컨대 XML 또는 관계형 데이터 베이스)에 보관되는 자료의 RDF로부터 표현될 수도 있다.

SPARQL은 RDF 그래프로부터 정보를 얻기 위한 질의 언어이다. 이는 다음과 같은 기능을 제공한다.

- extract information in the form of URIs, blank nodes and literals
- extract RDF subgraphs
- construct new RDF graphs based on information in the queried graphs

마) RDF/OWL Representation of WordNet

WordNet[Fellbaum, 1998]은 자연언어와 정보검색에 크게 사용된 검색 소스이다. 또한 최근의 WordNet은 Semantic Web research community에 적용되어졌다.

WordNet은 주로 어노테이션과 문화유산, 제품 카탈로그, 포토 메타데이터와 같은 다른 영역에서의 검색에 사용되었다. WordNet은 온톨로지 정렬 도구에서 배경지식과 다른 어플리케이션으로 FOAF²⁾ 스키마와 같은 다른 어휘에서 사용되었다. 현재 RDF(S)나 OWL의 변환을 위해 존재한다.

바) Best Practice Recipes for Publishing RDF Vocabularies

이 문서는 RDFS 또는 OWL 어휘 발행을 위한 제일 적합한 recipes 또는 웹에 온톨로지를 기술한다. 각 recipe의 특징은 명확하게 어휘 또는 온톨로지 제작자가 그들의 특정한 상황의 필요에 가장 적합한 방식을 선택할 수 있게 기술되어 있다. 각각의 recipes는 비록 다른 환경에서 적용되어진 원리를 포함할지라도 아파치 HTTP 사용을 위한 예가 포함되어 있다. 그 recipes는 모두 현재의 웹 구조에 맞게 설계되어 있다.

2) FOAF(Friend-of-a-Friend) : JeromeDL에서 이용한 인적 네트워크 메타데이터로서 제5장 시맨틱 디지털도서관에서 설명되고 있다.

제3장 온톨로지 토픽맵 기술을 사용한 모델링 방법론³⁾

이 장에서는 다양한 어플리케이션에 적용 가능한 토픽맵 온톨로지의 개발 방법 및 절차를 기술한다. 온톨로지 디자인 및 구현과 관련하여 현재 다양한 온톨로지 개발 방법론들이 존재하고 있다. 그러나 기존의 방법론들은 특정한 규모를 중심으로(일반적으로, 작은 범위의 온톨로지 또는 대용량 온톨로지) 기술되어 있거나, RDF/OWL을 중심으로 기술되어 있기 때문에, 이러한 개발 방법론을 토픽맵 온톨로지 개발 방법론으로서 변형 없이 적용할 수는 없다. 따라서 이 장에서는 토픽맵 모델에 기반을 둔 온톨로지 디자인 및 구현 절차를 구체적으로 기술하고자 한다.

토픽맵 모델에 따라 온톨로지를 디자인하고자 하는 온톨로지 디자이너와 토픽맵 온톨로지에 기반을 둔 어플리케이션을 개발하고자 하는 어플리케이션 개발자를 대상으로 하고 있다.

온톨로지 설계자는 이 장에서 설명하는 지침 및 개발 절차를 이용해, 온톨로지의 목적과 범위를 명확하게 이해하고 표현하고자하는 지식 개체들과 그 사이의 관계를 상세하게 분석함으로써 높은 재사용성을 갖는 토픽맵 온톨로지를 개발할 수 있을 것이다.

어플리케이션 개발자는 이 장에서 설명하는 온톨로지 개발 절차 및 방법을 이해하여 온톨로지 설계자에 의해 구현된 온톨로지에 대한 이해를 높임으로써, 보다 충실하게 온톨로지를 어플리케이션에 반영함으로써 효과적인 개발 작업을 수행할 수 있을 것이다.

이 장에서는 토픽맵 온톨로지 개발 방법 및 절차와 관련된 모든 사항들을 그 범위로 한다. 온톨로지와 토픽맵에 대한 일반적 개념은 구체적으로 다루지 않으며, 오직 토픽맵 모델 기반의 온톨로지 설계와 개발 방법 및 절차를 중요 내용으로 다룬다.

3) 온톨로지 토픽맵 기술을 사용한 모델링 방법론은 정보문화진흥원에서 나온 “국가지식정보 온톨로지 표준개발” 보고서의 [토픽맵 온톨로지 개발 방법론]의 내용을 인용하였다.

3.1 토픽맵 온톨로지의 목적과 범위

3.1.1 목적

다른 온톨로지 개발 방법론들과 마찬가지로 토픽맵 온톨로지의 개발 절차에서도 온톨로지 설계를 수행하고 실제적으로 구현하기 전에 선행해야만 하는 많은 종류의 사전 조사와 준비가 필요하며, 중요한 결정 사항들 또한 다양하게 존재한다.

토픽맵 온톨로지를 설계하고 개발하기 전에 반드시 다뤄야 하는 중요한 사항들 중 하나는 토픽맵 온톨로지의 목적과 범위를 명확하게 이해하는 것이다. 즉, 온톨로지가 최종적으로 달성하고자 하는 것이 무엇인지를 명확히 파악하고, 온톨로지의 주 사용자 대상 및 온톨로지가 응용될 어플리케이션의 종류에 대해 자세하게 고찰해야 한다.

• 목적 정의

토픽맵 온톨로지의 도메인(domain)을 명확하게 설정하기 위해서는 온톨로지 설계에 필요한 기초 자료들을 광범위하게 수집하는 것이 큰 도움이 된다. 온톨로지의 목표 도메인 내에서 광범위하게 수집된 자료들은 토픽맵 온톨로지의 핵심을 이루는 주요 개념들은 무엇이 될 것인지, 토픽맵이 얼마나 자세하게 기술될 수 있는지, 어떤 방향의 접근 방식을 취할 것인지 등의 사항들을 결정하는데 큰 도움이 된다.

또한 기존에 존재하고 있는 정보원을 분석하는 것도 매우 중요하다. 이미 존재하는 정보원들을 분석함으로써 온톨로지 내에 기술해야 할 주요 개념들을 더욱 명확하게 파악할 수 있고, 경우에 따라서는 온톨로지 구축과정의 많은 부분을 자동화할 수 있는 효과도 얻을 수 있다.

• 목적 정의를 위한 실천 방안

토픽맵 온톨로지의 목적과 범위를 정확하게 이해하기 위한 다양한 실천 방법들이 있을 수 있다. 가능한 모든 관련 문서들을 확인해 보는 것, 전화 또는 미팅을 통한 고객과의 만남, 온톨로지 개발 및 관련 어플리케이션 개발과 관련된 모든 사람들을 대상으로 하는 워크샵 개최, 제3자 인터뷰 등 상황에 따라 다양한 방법들을 모색해 볼 수 있다.

어떤 활동을 어디서부터 시작해야 할지 모를 때에는 이미 존재하는 다양한 종류의 문서를 읽어보는 것이 큰 도움이 된다. 문서들을 살펴보는 과정에서 프로젝트 대상 기관의 담당자와도 자연스러운 인터뷰가 이루어질 수 있으며, 이 과정에서 프로젝트에 대한 새로운 아이디어를 얻을 수도 있고 프로젝트에 대한 대략적인 범위도 가늠해 볼 수 있다.

온톨로지 설계자 주도하에 모든 프로젝트 참여자를 대상으로 하는 워크숍을 개최하는 것도 좋은 방법이다. 실제로 온톨로지 설계의 시작단계에서 가장 일반적인 방법이 워크숍을 여는 것이며, 온톨로지 설계자는 워크숍 준비 과정에서 온톨로지 설계를 위해 수집했던 기초자료들을 활용하여 워크숍의 핵심 이슈들을 도출해 내고, 워크숍이 진행되는 동안 프로젝트 참여자를 대상으로 활발한 브레인스토밍의 기회를 제공함으로써 온톨로지 설계에 도움을 줄 수 있는 많은 질문들을 이끌어 낼 수 있다.

제3자와의 인터뷰는 주로 정보원 소유자와의 인터뷰를 의미하며 이를 통해서도 온톨로지 디자인을 위한 많은 도움을 얻을 수 있다.

3.1.2 범위

온톨로지 설계와 구현 전에 행하는 모든 절차 및 활동들은 온톨로지의 도메인과 범위를 결정하는 것과 밀접하게 연관되어 있다. 온톨로지의 도메인을 결정한다는 것은 '설계하려는 온톨로지는 무엇에 대한 것인가?', '무엇이 모델링될 것인가?' 라는 중요한 질문에 명확한 답을 제시한다는 것을 뜻하며, 온톨로지가 기술할 범위의 한계와 정밀도를 파악하는 것이다.

• 범위 및 정밀도의 결정

범위를 결정한다는 것은 온톨로지를 어떤 시각에서 작성할 것인지와 관련되어 있다. 온톨로지의 정밀도를 결정한다는 것은 매우 어려운 문제이지만, 온톨로지의 재사용성과 온톨로지 구축에 드는 비용을 결정한다는 점에서 매우 중요한 문제이다.

실제로 기 구축된 온톨로지들의 사례에서 볼 수 있듯이, 온톨로지의 정밀도와 재사용성은 온톨로지 구축비용 및 시간과 반비례 관계에 놓여 있다. 낮은 정밀도의 온톨로지는 적은 노력으로도 작성할 수 있기 때문에 적은 구축비용이 소요되

나, 온톨로지의 단순성으로 인해 향후의 재사용성은 매우 낮다. 반면 '의학 온톨로지'와 같은 정밀한 온톨로지는 작성하는데 긴 시일과 막대한 노력이 필요하며 많은 비용이 소요되지만, 온톨로지의 정밀함으로 인해 향후의 재사용성은 매우 크다고 볼 수 있다.

• 토픽맵 온톨로지 기반 어플리케이션의 종류 및 범위

대부분의 토픽맵 온톨로지 설계 및 개발 단계는 어플리케이션을 구현하는 프로젝트의 한 부분으로 포함되어 있을 가능성이 높다. 때문에 개발되는 토픽맵은 프로젝트를 통해 개발하게 될 어플리케이션의 종류 및 범위에 영향을 받을 수밖에 없다.

프로젝트를 통해 최종적으로 개발되는 토픽맵 온톨로지 기반의 어플리케이션은 크게 두 가지 형태로 나눌 수 있다. 하나는 개별 조직의 업무 흐름도의 일부분으로 포함되는 어플리케이션으로 이를 통해 해당 업무 담당자들이 유용한 작업들을 수행하도록 돕는다. 다른 하나는 이용자들이 정보를 검색하고 정보와 지식을 관리하는데 도움을 주도록 하는 정보검색 어플리케이션이다.

명확하게 구분 지을 수 없는 어플리케이션의 형태를 굳이 구분지어 생각하는 이유는 어플리케이션의 형태가 어느 쪽에 가까운지에 따라 토픽맵 모델링에 차이가 생길 수 있기 때문이다. 본 문서에서는 정보 검색 어플리케이션에 초점을 맞추어 토픽맵 모델링에 대해 기술하지만, 전체적인 토픽맵 개발 방법론에는 큰 차이점이 없다.

• 범위에 따른 온톨로지 분류체계

자료를 효과적으로 검색하기 위한 다양한 분류방법이 존재한다. 분류에 있어서 중요한 것은 어떤 상황에서 어떤 분류를 사용할 것인가라는 선택에 있다. 상황에 적합한 최선의 분류가 한 개만 존재하는 것이 아니라 상황에 따라 혹은 문제의 복잡도에 따라 적절한 분류를 사용하는 것이 중요하다.

실세계에서 가장 흔하게 접할 수 있는 분류는 텍사노미이다. 텍사노미란, 개념들 간의 상하관계를 계층구조로 표현한 것으로서 도서관 분야, 인터넷 포탈, 지식베이스 등에서 가장 광범위하게 쓰이는 분류구조라고 할 수 있다. 하나의 개념에 복수의 술어를 허용하는 시소러스도 광범위하게 쓰이는 분류구조이며, 이보다 복잡한 형태인 정형화된 태그 분류법도 대량의 지식구조를 표현하는데 많이 사용되

고 있다.

여기서 여러 가지 분류체계에 대해 열거하는 이유는 기술하려는 영역에 포함된 모든 항목을 완전히 온톨로지 구조로 기술한다는 것이 시간적, 비용적 낭비를 초래할 수도 있기 때문이다. 특정 도메인 내의 전체 지식을 완전한 온톨로지의 형태로 기술할 필요는 없으며, 규모가 작고 단순할 경우에는 텍사노미나 시소러스를 이용하여 단순한 형태의 온톨로지를 구축하는 것이 보다 효율적인 방안일 수도 있다.

완전하고 정교한 온톨로지를 개발해야 할 때도 텍사노미나 시소러스와 같은 기존의 분류개념들을 활용할 수 있다. 즉, 기술하려는 영역 내에서도 명료한 주제와 불명료한 주제들이 복합적으로 존재하기 때문에, 자세하고 명료한 주제에 대해서는 온톨로지 구조를 적용하여 개념과 관계들을 상세하게 기술하도록 하고, 일반적이고 불명료한 주제들에 대해서는 시소러스나 텍사노미를 사용하는 복합적인 방법을 취하는 것이 비용 절감 측면에서 효과적일 수 있다.

3.2 분석

목적과 범위의 정의를 통해 온톨로지가 추구하는 목표와 범위가 정해지고 검토해야 할 사항들에 대한 정리가 끝났다면 실제적인 검토와 분석이 이루어져야 한다.

분석 단계에서 가장 중요한 사항은 프로젝트의 목표와 범위에 따라 요구사항을 철저하게 분석하고 기존의 데이터 소스에 대한 충분한 이해에 도달하는 것이다. 데이터 소스들의 구조와 내용들을 자세히 검토하면서 온톨로지로 변환하기에 가치 있고 적절한 데이터 소스인지를 확인해야만 한다. 데이터 소스의 정확한 이해를 위해 데이터 소스 소유자와의 직접적인 커뮤니케이션을 권장하지만, 소유자들이 항상 협조적이지는 않다는 사실을 고려해야 한다.

데이터 소스의 분석을 진행하면서 고려해야 할 사항은 데이터 소스들이 변환되어 그 결과물로서 설계될 온톨로지가 목표한 어플리케이션의 요구사항을 충족하고 결과적으로 이용자의 요구사항도 충족시킬 수 있을 것인가라는 사실이다. 목표로 하고 있는 온톨로지를 설계하기에는 데이터 소스들이 부적합하거나 부족하다고 판단된다면 이에 대한 다른 방안을 고려하고 해결 방법을 모색해야 한다.

3.2.1 요구사항 분석

프로젝트 요구사항 분석은 온톨로지 디자인의 전 과정에 걸쳐서 이루어지는 모든 결정사항에 영향을 미치는 중요한 요소이므로 이 과정은 매우 세밀하게 수행되어야 한다. 특히, 프로젝트를 통해 개발된 어플리케이션을 실제 사용하게 되는 이용자들의 요구사항은 매우 세심하게 분석되고 고려되어야 한다.

아래에 기술된 네 범주의 질문 리스트를 적용하여 프로젝트의 요구사항을 분석해 봄으로써 온톨로지 개발 시에 어떠한 점에 중점을 두고 어떤 요구사항에 우선순위를 두어 결정을 내려야 하는지에 대한 지침 및 상세한 요구사항 리스트를 작성할 수 있다.

1) 설계

- 어플리케이션 도메인의 영역은 어디까지인가?

이 질문이 중요한 이유는 토픽맵 온톨로지에 기술될 지식의 '넓이'와 '깊이'에는 한계가 없기 때문에 관리 가능할 정도로 제한할 필요가 있다. 정의된 영역은 주제가 얼마만큼 상세하게 기술될 것인지, 얼마만큼 넓게 기술될 것인지를 결정하고, 토픽맵 온톨로지의 설계와 규모를 제어하게 될 것이다. 향후 도메인 영역이 어떤 방향으로 어떻게 변화해 갈 것인지를 고려하는 것 또한 중요하다.

- 주 사용자층은 누구이며, 그들이 어플리케이션에 기대하는 것은 무엇인가?
특별한 요구사항을 가진 중요 사용자 집단이 있는가?

모든 어플리케이션은 실제 이용자의 요구사항과 기대를 충족시켜 줄 수 있어야 하며, 이는 토픽맵 온톨로지에도 적용된다. 이것은 토픽맵 온톨로지 디자인의 모든 단계에서 고려되어야 하는 사항이기도 하다. 만약 토픽맵 온톨로지가 다양한 이용자 집단을 대상으로 하는 어플리케이션에 활용되어야 한다면, 이들 중 가장 중요한 이용자 집단을 중심으로 해서 토픽맵 온톨로지를 기술하는 것이 좋다.

- 비용과 시간에 관한 제약사항은 무엇인가? 이용 가능한 인적 자원에는 무엇이 있는가?

토픽맵 모델은 매우 낮은 비용으로 온톨로지 개발을 시작할 수 있도록 해주며 차후 여건에 따라 온톨로지를 전개시켜 나갈 수 있는 유연함을 제공하여 주므로,

온톨로지 개발 초기단계부터 너무 상세하게 기술하려 할 필요는 없다. 상세하게 기술한다는 것은 그만큼 큰 비용과 많은 시간의 투여를 의미한다. 가능하다면 이해하기 쉽고 단순하게 온톨로지 디자인을 시작하되 이용자의 요구사항에서 벗어나지 않도록 하는 것이 중요하다.

- 토픽맵의 구조에 대한 제약사항이 필요한가?

토픽맵 온톨로지에 적용된 트리(tree) 계층구조는 온톨로지를 더욱 더 강력하게 할 뿐만 아니라 이해하기 쉽도록 하고, 다양한 형태의 관계들을 활용하여 디자인된 토픽맵 모델을 보다 쉽게 관리할 수 있도록 한다는 것은 기구축된 여러 온톨로지들을 통해 이미 증명되었다.

온톨로지의 복잡성을 제거하기 위하여 계층구조의 단계 또는 깊이를 제한할 필요가 있는 경우, 계층구조의 형태(주제 카테고리, 텍사노미, 시소러스, 웹 사이트 맵 등)에 따라 토픽맵에 어떤 형태의 계층구조를 적용할 것인지를 세심하게 고려해야 한다. 이러한 계층구조의 결정은 향후 토픽맵 디자이너들과 토픽맵 이용자들이 수행하게 될 모든 작업들에 영향을 미치게 된다.

- 이용자에게 얼마만큼 자세하게 토픽맵 온톨로지의 요소들을 보여 줄 것인가?

일부 토픽맵 기반의 어플리케이션들은 이용자에게 온톨로지의 모든 상세 요소들을 디스플레이해 준다(모든 Topic Type, Topic, Topic Name, Occurrence, Association 등). 반면 다른 어플리케이션들은 자세한 사항들은 감추고 이용자에게 친숙한 개체들만을 보여줌으로써 토픽맵 모델을 최대한 단순하게 보여주기도 한다.

선택은 어플리케이션을 이용하는 이용자 집단의 수준에 달려있다. 어플리케이션을 이용하는 이용자 집단이 복잡한 질의를 이용하여 최대한 세밀하게 정보 검색을 수행하거나 토픽맵 모델을 이해하고 있는 전문가 집단이라면, 가능한 자세하게 온톨로지 요소들을 노출하는 것이 도움이 되지만, 반대의 경우에는 이용자들이 쉽게 이해할 수 있는 요소들만을 보여주는 것이 더 올바른 선택이 될 것이다.

2) 범위(scope)

- 다양한 관점이 필요한가?

범위라는 개념은 단일한 토픽맵 온톨로지를 다양한 관점에서 볼 수 있도록 해주며, 토픽맵의 일부분만을 보여주거나 특정 요소를 기준으로 필터링하고 복잡한 요소들을 감추는 등의 기능을 제공한다. 관점이라는 것은 이용자의 흥미, 언어, 접근 제어 등 다양한 요소들을 의미하며, 토픽맵 온톨로지 개발 단계의 모든 단계에서 범위를 추가하거나 제거할 수 있다.

- 두 개 이상의 언어를 사용하여 토픽맵을 기술해야 하는가?

범위를 활용해 토픽 이름을 다양한 종류의 언어로 기술할 수 있다. 일반적으로 두 개 이상의 언어를 사용하는 어플리케이션을 개발하는 것은 많은 노력과 시간이 요구되는 작업이지만, 토픽맵 온톨로지 개발 단계에서 범위를 적용하여 다양한 언어를 기술한다면 토픽맵 온톨로지의 재사용성을 크게 증가시킬 수 있을 뿐만 아니라 어플리케이션 개발 단계에서 소요되는 비용 및 시간을 획기적으로 단축시킬 수 있다.

3) 구현

- 토픽맵 온톨로지는 어떤 과정을 거쳐 구현될 것인가?

토픽맵 온톨로지는 전 과정에서 토픽맵 디자이너에 의해 수작업으로 디자인되고 구현될 수도 있지만, 자동화된 과정을 통해 구현될 수도 있다. 수작업에 의해 구현된 온톨로지가 질적인 면에서 높은 정확도와 품질을 지니게 되는 것은 일반적인 사실이지만, 온톨로지의 규모와 소요 시간, 구축비용, 이미 활용 가능한 소스의 존재 유무 등을 고려하여 수작업 및 자동화의 범위를 결정해야만 한다.

- 레거시(legacy) 데이터가 필요한가?

여기서 레거시 데이터란 토픽맵 온톨로지로 변환 가능한 텍사노미, 시소러스, 온톨로지 등 이미 존재하는 데이터들을 의미한다. 프로젝트 수행 과정에서 이러한 레거시 데이터들이 이미 존재하고 있을 가능성은 매우 높으므로 레거시 데이터에 대한 분석은 토픽맵 온톨로지 개발의 좋은 출발점이 될 수 있다. 중요한 요소들이 있다면, 가능한 한 추출하여 토픽맵 온톨로지로 변환하도록 한다.

- 토픽맵과 실제 정보 자원들은 얼마나 긴밀하게 연결되어야 할 것인가?

수작업을 통한 Occurrence의 생성 또는 유지는 매우 많은 비용이 소요되는 작업이다. 자동 카테고리/분류 도구를 사용하여 Occurrence를 생성, 유지하는 방안을 고려하도록 한다.

4) 편집

- 온톨로지 구현에 누가 참여할 것인가? 그들이 갖추어야 할 기술은?

온톨로지를 디자인하는 것도 중요하지만, 실제 온톨로지 내용의 대부분을 구성하게 될 인스턴스를 입력하는 것 또한 매우 중요한 작업이다. 온톨로지 디자이너에 의해 설계된 온톨로지 구조에 실제 인스턴스를 입력하는 작업은 토픽맵 편집자들에 의해 이루어지며, 토픽맵 편집자들은 온톨로지가 기술하고 있는 주제 분야에 대해 풍부한 지식을 갖추고 있어야 하고, 온톨로지를 실제로 이용하게 될 사용자 집단의 특성을 잘 이해하고 파악하고 있어야 한다.

- 토픽맵 온톨로지의 개발에 투입될 편집자의 규모는?

분석단계에서 프로젝트의 성격과 규모에 따라 참여하는 편집자의 수를 고려하여야 한다. 온톨로지 개발에 온톨로지 디자이너 외에 둘 이상의 편집자들이 투입된다면, 온톨로지를 작성하는 과정에서(혹은 인스턴스 작성 과정에서) 온톨로지의 디자인 또는 스키마가 변경되지 않도록 하는 것이 매우 중요하다. 온톨로지 작성 과정 중에 스키마가 변경된다면 온톨로지 구조 전체에 매우 심각한 문제가 발생할 수 있기 때문이다.

- 편집자들에 대한 접근 제어 및 관리가 필요한가?

온톨로지에 기술되는 사항들 중에는 외부에 노출되어서는 안 되는 기밀 사항들이 포함될 수도 있다. 이런 경우, 토픽맵 편집자들에 대해서도 토픽맵 온톨로지의 특정 영역에 대한 접근 제어를 고려해야만 한다.

- 토픽맵 온톨로지의 유지보수는 어떻게 이루어질 것인가?

더 이상 필요하지 않거나 유효하지 않은 정보들은 주기적으로 제거하고, 새로운 정보들은 즉시 추가될 필요가 있다. 온톨로지의 업데이트 및 삭제 주기와 자동적으

로 이루어질 것인지 아니면 편집자에 의해 정해진 주기에 일어날 것인지를 결정해야 한다.

- 향후 통합에 대한 가능성은 없는가?

온톨로지가 향후 통합될 가능성은 없는지를 고려해 보는 것도 분석과정에서 수행해야 하는 중요한 요소 중 하나이다. 향후의 통합을 고려하여 온톨로지를 개발해야 한다면 표준을 더욱 엄격하게 준수하도록 해야 한다. 표준을 더욱 엄격하게 준수한다는 것은 공적 주제(PS : Published Subjects, 이하 PS)를 적극적으로 활용한다는 것을 의미하며, 온톨로지 디자이너는 이미 존재하고 있는 PS를 광범위하게 조사하고, 이를 토대로 온톨로지 편집자들이 가능하면 PS를 사용하여 온톨로지를 편집할 수 있도록 이끌어야 한다. 앞으로 작성하게 될 온톨로지와 관련하여 새로운 PS를 구축하는 것도 좋은 방법이 될 수 있다.

3.2.2 기존 시스템 및 자료 분석

일반적으로 시스템을 완전히 새로 작성하는 것보다는 기존에 존재하는 것을 최대한 재활용하여 이용하는 것이 효율적인 방법이다. 토픽맵 온톨로지를 작성할 때도 온톨로지의 모든 요소들을 새로이 작성하는 것보다는 작성하려는 분야에 대해 기 구축된 온톨로지가 존재할 경우 이를 최대한 재활용하고, 이미 존재하는 정보 자원, 저장소 또는 레거시 시스템으로부터 활용할 수 있는 모든 요소들을 추출하여 토픽맵 온톨로지로 변환하고 이용하는 것이 시간과 비용을 절약할 수 있는 효율적인 방법일 것이다.

이미 존재하는 정보 자원들, 저장소, 레거시 시스템들을 가능한 한 세밀하게 조사하여 토픽맵 온톨로지에 재사용할 수 있는 요소들을 추출해 내야 한다. 이러한 과정에서 동일한 분야에 대해 이미 구축된 스키마나 온톨로지를 발견할 수도 있을 것이다.

토픽맵 온톨로지에서도 재활용할 수 있는 요소들을 추출하기 위해 상세하게 분석해 봐야 할 정보 소스들로는 크게 구조화된 데이터 소스와 비구조화된 데이터 소스로 나누어 살펴볼 수 있다.

1) 구조화된 데이터 소스

구조화된 정보원이란 어떤 형식을 통해서든지 정제과정을 거쳐 이용자가 사용하기 쉬운 형태로 가공된 정보 자원들의 집합을 의미한다. 구체적으로, 이미 존재하는 온톨로지나 데이터베이스의 스키마 정보, 스키마에 대해 상세히 기술한 문서들, 데이터 반출물 또는 샘플 데이터 등이 모두 구조화된 데이터 소스에 포함되며, 구조화된 정보는 이미 데이터 소스들이 가지고 있는 중요한 요소들과 요소들 사이의 관계를 표현하고 있기 때문에 이를 입수하고 분석하는 것이 중요하다.

한 가지 유의해야 할 점은 스키마와 문서에 대한 지나친 신뢰를 지양해야 한다는 것이다. 스키마나 문서를 통해 많은 것들이 도출되지만, 스키마에 어떤 필드가 존재한다고 해서 그에 대한 값이 항상 존재하는 것은 아니라는 것을 알아야만 한다. 즉 데이터 소스가 가치 있는 값을 가지고 있는지를 파악하기 위해서는 실제 데이터를 살펴보는 것이 가장 확실한 방법이다.

- 모든 형태의 온톨로지

구현된 형태와 상관없이 구축하려는 주제와 관련된 모든 종류의 기구축된 온톨로지를 의미한다. 이는 책자 형태일 수도 있고, RDF/OWL로 구현된 온톨로지일 수도 있다.

- 주제 분류표, 텍사노미, 시소러스 등

어느 분야에서든 정보를 효율적으로 분류하려는 노력이 있어 왔으며, 표준화된 다양한 주제 분류와 텍사노미, 시소러스 등이 존재한다.

- 데이터베이스 스키마

현재 대부분의 정보 자원들이 데이터베이스에 저장되어 있으며, 관계형 데이터베이스의 스키마나 객체지향적 데이터베이스의 스키마에 대한 분석을 통해 온톨로지 디자인에 관한 많은 아이디어를 얻을 수 있다.

- 메타데이터 용어집

메타데이터 용어집을 통해 온톨로지 개체의 명칭에 대한 아이디어를 얻을 수 있다.

- 목차, 인덱스, 시소러스, 데이터 사전

토픽맵 온톨로지의 기술 영역에 관한 서적, 시소러스, 사전 등을 살펴봄으로써 그 분야에 대한 풍부한 지식을 얻을 수 있을 뿐만 아니라 온톨로지의 중요한 클래스들(Topic Type), 클래스간의 계층구조 및 Association에 대한 유용한 정보를 얻을 수 있다.

- 문서 구조(DTD, XML Schema)

분야에 따라 많은 문서 구조들이 존재하며, 웹상에서도 상세한 정보를 손쉽게 검색해 낼 수 있다. 대부분의 이러한 문서 구조들은 공용 포맷인 XML로 기술되어 있어서 추출 및 변환이 용이하다는 점 외에도, 정교한 구조 및 기술 규칙을 포함하고 있다는 점에서 많은 도움을 줄 수 있다.

위에 나열한 정보 자원들의 추출 및 활용에 대해서는 뒷부분에서 더욱 자세하게 다루도록 한다.

2) 비구조화된 데이터 소스

현재 대부분의 데이터 소스들이 비 구조화된 형태로 존재하며, 그 형식 및 형태도 매우 다양하게 존재한다. 수집한 데이터 소스들의 언어, 포맷, 문자 인코딩 등에 대한 사항들을 정확히 파악하여 가장 효과적으로 변환하는 방법을 찾아야 한다. 어떤 메타데이터들을 자동으로 추출할 수 있을 것인지, 키워드를 추출할 방법은 없는지 등의 의문을 가지고 데이터 소스를 분석하면 비 구조화된 데이터 소스 속에서도 가치 있는 정보들을 발견할 수 있다.

3) 데이터 소스 분석 유의사항

데이터 소스를 분석할 때는 아래의 몇 가지 항목들에 유의하면 데이터 소스 변환 시 장애가 될 요소들을 사전에 많이 제거할 수 있다.

첫 번째는, 데이터 셋의 ID 또는 Key값에 관한 것이다. 하나의 데이터 셋에 존재하는 ID 값이 다른 데이터 셋들과 어떻게 연결되는지를 정확하게 파악하고 그 관계와 구조를 이해하도록 한다. 하나의 데이터 소스에서 다른 데이터 셋과 연결 관계를 설정해야 함에도 불구하고 특별한 ID나 Key 값이 존재하지 않는다면 데

이터 셋들 간의 관계를 설정할 수 있는 다른 방법을 찾아보아야 한다.

두 번째는, 데이터 값의 포맷에 대한 것이다. 특히, 날짜 또는 시간 형식의 데이터 포맷에 유의해야 한다.

세 번째는, HTML 코드가 값으로 포함되어 있는 경우이다. 실제로 많은 데이터베이스 내에 HTML로 코딩된 데이터들이 값으로 존재하고 있으며, 데이터 변환 시에 가장 많은 문제를 야기한다. 데이터 변환 시에 전체를 변환할 것인지, 일부 분만을 추출하여 변환할 것인지를 결정해야 한다.

마지막으로 유념할 것은, 데이터가 항상 옳지는 않다는 것이다. 예를 들어, 파일 확장자가 .xml로 끝나는 모든 문서들이 well-formed 문서가 아닌 것처럼, 데이터베이스 내에도 드러나지 않는 많은 문제가 내포되어 있는 경우가 있다. 참조 무결성 제약이 없는 데이터베이스의 경우 관계에 많은 오류가 있을 수 있으며, 일반적인 모델로 구현된 데이터베이스라도 그 구조를 파악하기 어려운 경우도 많다. 데이터 소스에 따라서 같은 의미의 칼럼이라도 칼럼이나 테이블 이름은 다르게 쓰는 경우가 많기 때문에 이러한 부분들도 정확하게 파악해야 한다.

3.2.3 브레인스토밍

기술하려는 도메인 영역 내의 모든 주제들은 토픽맵 온톨로지에서 토픽으로 기술되어 컴퓨터로 처리될 수 있어야 하기 때문에, 주제에 대한 브레인스토밍은 좋은 출발점이 될 수 있다.

어떤 것이 Topic Type이 될 것인가? Association Type, Occurrence Type, 인스턴스 등에 대한 세밀한 분석 없이 모든 주제들의 리스트를 작성하도록 한다. 요구사항 분석 및 기존 시스템의 분석을 통해 얻은 다양한 데이터 소스들로부터 온톨로지를 구성하게 될 핵심적인 주제 및 단어들을 추출하여 나열해 보도록 하고, 실제 이용자들이 온톨로지를 어떠한 시각에서 바라 볼 것인지를 고려하여 다양한 관점에서 주제 분야를 바라보고 생각나는 모든 개념들과 주제들을 리스트로 정리하도록 한다.

1) 브레인스토밍 예시

브레인스토밍은 매우 창의적인 작업이긴 하지만 구조화된 단계를 거치는 체계적인 작업은 아니기 때문에, 브레인스토밍을 통해 얻은 개념 및 주제 리스트들을

다음 단계인 온톨로지 디자인 단계에서 효과적으로 활용하기 위해서는 실제적인 디자인 과정 이전에 정제하는 작업이 요구된다.

- 모든 주제가 서로 상이하며 동의어가 아님을 확인한다. 동의어인 경우 주어진 주제의 대체이름으로 표시한다.
- 주제 리스트를 보고 Topic, Occurrence, Association 및 Role에 관한 클래스를 선정한다. 리스트 내에서 Topic, Occurrence, Association에 대한 인스턴스를 발견할 시에는 관계된 클래스의 예시로 표시한다.
- 관점, 컨텍스트 정보, 접근 권리를 정의하는 범위로 사용될 주제를 선정한다.

- 브레인스토밍 예 (판소리)

브레인스토밍이 된 주제: 판소리에 관한 주제 리스트

강산제, 더늠, 동초제, 동편제, 명창, 춘향가, 심청가, 발성법, 병창, 사람, 사설집, 서편제, 작곡가, 장단, 중고제 판소리, 판소리 종류, 유명 지역, 유파의 대표 명창, 주요 대목, 창의 유명 명창, 판소리 대표장단, 판소리 작곡가, 계보, 공연정보, 기사, 내용, 논문, 비평, 사진, 제작연도, 출생연도, 홈페이지, 활동연도, 강정숙, 권삼득, 성우향, 성창순, 유공열의 이별가, 설령제, 자진모리, 중모리, 휘모리, 진양조, 고수, 송영주, 출생지, 부산, 전남, 전북 등등

3.3 주요 개념 추출 및 정제

브레인스토밍 과정을 거쳐 작성된 주제와 개념들은 온톨로지의 핵심을 이루게 될 주제와 용어들을 대부분 포함하고 있는 가치 있는 결과물임에는 분명하지만, 불필요한 정보를 포함하고 있을 가능성도 무시할 수 없다. 그러므로 실제 토픽맵 온톨로지 디자인에 들어가기 전에 리스트를 구조화하고 정제할 필요가 있다.

3.3.1 중복제거

브레인스토밍을 통해 나열한 주제들이 서로 다른 주제들인지, 중복되는 것은 없는지 확인할 필요성이 있다. 실제로 브레인스토밍을 거치는 과정에서 같은 뜻을 가진 용어들이 중복되어 기술되는 경우가 흔하기 때문에, 이를 확인하여 대표

적인 단어로 주제를 기술할 필요가 있다. 중복되어 기술된 유의어 등은 Topic의 이름을 결정하는 과정에서 활용할 수 있도록 정리하도록 한다.

3.3.2 클래스와 인스턴스의 구분

나열된 주제들 중에서 클래스(또는 타입)와 인스턴스에 포함될 항목들을 구분한다. 토픽맵에서 클래스란 타입을 의미하며, Topic Type, Association Type, Occurrence Type, Role Type 등이 이에 속한다. 타입의 정의는 다음과 같이 요약할 수 있다.

- 무엇이든 상관없이 일반성을 내포하는 개체들의 집합
- 적어도 하나 이상의 속성을 공유하는 개별 개체들의 집합
- 내재적 또는 외재적으로 정의된 카테고리 각각의 카테고리는 이에 속하는 인스턴스 집합을 가짐

토픽맵 온톨로지 모델링에서의 타입이란 위 세 개 중 마지막의 개념에 더 가까운 것으로, 하나의 카테고리에 여러 개의 인스턴스가 포함될 수 있다고 할 때, 그 카테고리를 '타입'이라고 한다.

타입을 발견하는 가장 좋은 방법은 토픽맵내의 개념 중 가장 범용적이고 일반적인 것을 살펴보는 것이다. 이러한 개념들은 온톨로지 내에서 최상위에 위치하며 다른 개념들을 하위요소로서 포함하게 된다. 예를 들어, 브레인스토밍 후 다음과 같은 주요 개념들을 도출하였다고 가정하자.

- 개념의 나열 : Lars Marius Garshol, Boeing, Norway, Michel Biezunski, France, Ontopia, Google, Inc., USA, Steve Newcomb, Steve Pepper...

이러한 개념들은 다음과 같이 분류될 수 있다.

- Person : Lars marius Garshol, Steve Newcomb, Michel Biezunski, Steve Pepper
- Company : Boeing, Ontopia, Google, Inc.
- Country : Norway, France, USA

Person, Company, Country처럼 인스턴스 요소들을 가지는 카테고리들을 타입이라고 하며, 만약 어떤 인스턴스도 가지고 있지 않다면 그것은 타입이 아니다. 즉 개념 또는 개체에 그것에 귀속되는 인스턴스가 존재하지 않는다면, 그것은 타입이 아니다. 타입으로 정의되지 않은 모든 개체들은 인스턴스 항목으로 분류된다.

3.4 토픽맵 온톨로지 디자인

분석의 마지막 단계인 브레인스토밍 과정과 개념 추출 및 정제 단계를 거치면 실제적인 온톨로지를 디자인할 수 있는 기초가 마련된다. 이전 단계를 통해 얻은 주제 리스트를 가지고 개별 타입들을 보다 상세하게 정의하고 구조화시키도록 한다.

여기서 주의해야 할 점은 이전 단계, 즉 주요 개념 추출 및 정제 과정을 거쳐 분류한 타입 및 인스턴스 구분은 확정된 것이 아니라는 점이다. 특정 주제가 Topic Type이 될 것인지 또는 Association Type, Occurrence Type, 또는 단순한 인스턴스가 될 것인지에 대한 결정은 온톨로지 디자인 단계 전반에 걸쳐 더욱 명확해 질 수 있으며, 변동될 수도 있는 사항이기 때문이다.

3.4.1 Topic type 디자인

추후 토픽맵 온톨로지를 편집하는 과정에서 토픽맵 편집자들이 인스턴스 개체들을 복잡한 판단과정 없이 정의된 Topic Type에 포함시킬 수 있다면, 그 타입은 잘 정의된 타입이라고 할 수 있다. 즉, 각각의 인스턴스들이 어떤 타입에 포함될지를 판단할 때 복잡한 사고과정 없이 직관적으로 판단할 수 있다는 것은 그 타입이 그만큼 명료하고 정확하게 정의되었음을 의미하는 것이다.

토픽 타입을 정의하는 과정에서 일반 데이터 리소스와 Topic 인스턴스, Topic Type을 구분하는 것은 매우 중요하다. 오직 Topic으로 구현된 정보 자원만이 Name, Occurrence, 다른 Topic과의 Association 등의 속성을 가질 수 있다는 사실을 항상 인식해야만 하며, 이는 일반 데이터 리소스와 Topic 인스턴스를 구분할 수 있는 중요한 기준으로 작용한다. 만약 어떤 정보 자원이 Name이나

Occurrence, Association 등의 속성을 필요로 하지 않는다면 그 정보 자원은 순수한 인스턴스로 표현된다.

3.4.2 Occurrence type 디자인

토픽맵 내에서 Occurrence는 개별 Topic들의 세부 속성을 기술하거나 Topic과 외부 자원과의 관계를 기술하는데 사용된다. 흔히 Topic에 대한 세부 속성을 기술하는 것을 Topic에 대한 internal occurrence(내부 어커런스)를 기술한다고 하며, 외부 자원과의 연결 표현에 대해서는 external occurrence(외부 어커런스)를 기술한다고 한다. 여기서 외부 자원은 해당 Topic과 연관된 문서, 이미지, 웹사이트, 웹페이지 등 모든 형태의 자원들을 의미하며, Topic과 연관된 대부분의 Occurrence는 external occurrence를 의미한다.

1) Occurrence와 Topic

모든 외부 정보들이 Occurrence로 표현되어야 하는 것은 아니다. 예를 들어, 특정 외부 정보에 대한 구체적인 정보(제목, 저자, 출판일 등)가 필요한 경우, 이 외부 자원이 토픽맵 내에서 external occurrence로 기술되어 있다면 그 자원에 대한 상기의 정보를 얻기 어렵다. 이와 같은 경우에는 이 외부 정보를 대표하는 Topic을 생성하여 표현하는 것이 바람직하다.

상세하게 기술할 필요가 있는 외부 자원들을 별도의 Topic으로 정의하면, 정보 자원간의 관계를 활용해 기술할 수 있고, 정보자원의 URL은 Topic에 대한 주제 위치자(subject locator)가 된다.

2) 속성과 Topic

Topic과 관련한 속성 중에는 특정한 값을 갖는 경우가 있다. 예를 들어 person 이라는 Topic을 기술하는 sex라는 속성은 male 또는 female이라는 두 가지의 값을 갖는다. 또 다른 예로 '작업흐름 상태'라는 Topic은 initiated, draft, accepted, published, deleted라는 다섯 가지 상태의 값을 갖는다.

이와 같은 경우, 속성을 Occurrence Type으로 정의하고 각각의 값을 어커런스로 기술하기 보다는, 개별 값들을 Topic으로 정의하고 속성을 Topic Type으로 정

의하여 Association을 통해 그 연관 관계를 기술하는 것이 더욱 효과적이다. 즉, sex라는 Occurrence Type을 정의하여 male과 female이라는 값으로 기술하기 보다는, sex를 Topic Type으로 정의하고 male과 female을 sex의 인스턴스로 정의하여 person과 Association을 통해 기술하는 것이 바람직하다.

3.4.3 Association Type 디자인

Association Type은 용어 리스트에서 손쉽게 찾아낼 수 있으므로, Association Type을 정의할 때 주의를 기울여야 할 것은 리스트에서 표현되지 않은 관계를 도출해 내는 것이다. 토픽맵 온톨로지가 풍부한 의미와 관계를 표현할 수 있기 위해서는, 정의된 Topic Type들 사이에 존재할 수 있는 여러 형태의 관계들을 도출해내고 이들을 Association Type으로 정의해야 한다.

1) Association Type의 명명

Topic Type의 경우와 같이 Association Type의 이름은 명사형을 사용한다. 동사형으로 타입을 기술하면, 동사형이 내포하는 연계의 방향성 때문에, 온톨로지 편집 과정에서 Association을 잘못 기술할 가능성이 있다.

영문으로 온톨로지를 기술할 경우 Association을 동사형으로 기술하는 것이 적절하고 보다 명확할 수 있지만, 언어적 차이로 인해 한국어로 기술할 때는 명사형으로 기술하는 것이 권장된다. 다만 동사형으로 기술하는 것이 적절하다고 판단될 경우에는 온톨로지의 특성에 따라 선택적으로 기술하도록 한다.

2) 관계 Role(역할)의 명명

Topic Type의 경우와 같이 관계 Role의 이름은 명사를 사용하며, 이 명사는 해당 Role이 인스턴스가 되는 Topic 클래스의 하위클래스나 상위클래스를 나타낸다. 예를 들어, "회사" (호텔의 상위클래스)/ "위치"(도시의 상위클래스)로 표시할 수 있다.

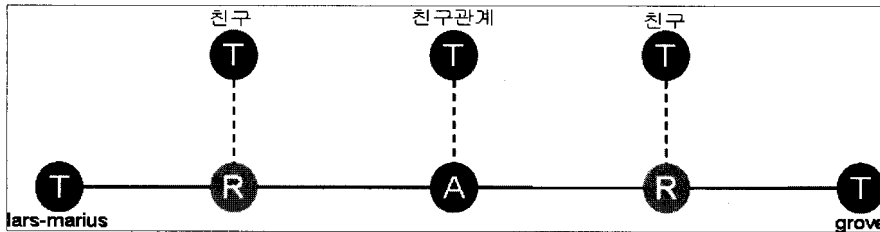
3) 관계 Role Type

Topic과 Topic 사이의 연관 관계를 기술할 때는, Association을 중심으로 각각의 Topic이 어떤 역할을 하는지를 명확하게 기술해야 하며, 모든 Association은

관계 Role과 함께 기술된다. 관계 Role 또한 Topic이다.

4) 대칭 관계

두 Topic 사이의 관계가 대칭관계라는 것은, 서로 관계를 맺고 있는 개별 Topic들의 어떤 방향에서도 Association 및 레이블(label)이 동일하다는 것을 의미한다. 이와 같은 관계의 경우 Association에서의 Role은 항상 동일하다. 아래 그림은 대칭 관계에 대한 예시이다.



[그림 5] Topic의 대칭 관계 예시

5) Association에서 Role의 개수

토픽맵 모델에서 Association이 포함할 수 있는 Role의 수에는 제한이 없다. 일반적으로 두 Topic이 Association을 통해 연결될 때, 관계를 명확히 기술하기 위해 두 개의 Role을 사용하지만 때로는 그 이상의 Role이 필요할 때도 있다. 오직 하나의 Role만을 사용하여 관계가 기술될 수도 있으며, 이 경우를 일진(unary)관계라고 한다. 일진관계는 yes/no 상태를 나타낼 때 주로 사용한다.

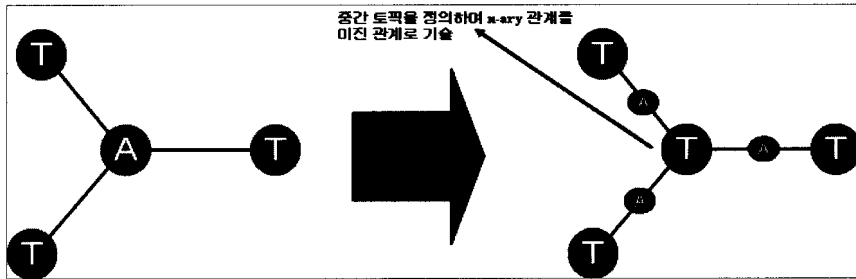
6) Association Role 정의에서의 원칙

한 Association에 포함되는 Role의 수는 가능하면 적게 유지하는 것이 좋다. 가장 일반적인 관계가 두 Topic 사이의 이진관계(binary association)이며 이것이 가장 바람직한 것처럼, 두 Topic 사이에서 각각의 Topic에 대한 Role 또한 Association 당 두 개의 Role을 갖도록 하는 것이 가장 바람직하다.

하나의 Role이 동일한 Association에 대해 중복되어 사용된다는 것은, Association이 컬렉션 또는 그룹을 의미하는 것이므로, 컬렉션 및 그룹을 위한 Topic을 생성하는 것이 가장 최적의 방법이라고 할 수 있을 것이다.

Association에 따라 3개 이상의 Topic을 연결하는 “n-ary” 관계가 존재할 수도

있지만, 가능하면 n-ary 관계도 이진 관계로 변환하여 표현하는 것이 좋다. n-ary 관계의 변환은 Topic들 사이를 매개하는 중간 Topic을 정의하고, 매개 Topic과 Topic들 간의 관계를 분리하여 정의하는 것으로 이를 그림으로 나타내면 다음과 같다.



[그림 6] n-ary Topic의 변환 도식

7) Association Type과 Role Type

Association Type에서 사용된 Role Type들은 Association Type의 정의를 기술하는 일부분이다. 일반적으로 동일한 Topic Type에 대한 모든 관계는 동일한 Role Type을 통해 기술되어야만 한다. 이것에 대한 예외 사항은 n-ary 관계이다.

대부분의 Association들이 Topic들 사이의 관계를 기술하면서 반복적으로 사용되며, 일부 Role Type은 너무 광범위하게 쓰여서 다양한 Association Type에 복수로 사용될 수도 있다.

3.4.4 범위(scope) 집합 정의

토픽맵 온톨로지에서 범위는 Topic의 속성인 Name, Occurrence, 그리고 Association을 특정한 시각에 적합하게 표현되도록 함으로써 온톨로지의 의미를 이용자에게 보다 풍부하게 보여주는 역할을 한다. 범위는 온톨로지에 대한 다양한 관점을 정의하고, 네임 스페이스를 통한 이름 중심의 토픽맵 통합에 큰 영향을 미치기 때문에, 온톨로지 디자인 과정에서 다양한 관점을 분석하고 그에 따라 범위를 정의하는 것이 매우 중요하다.

3.4.5 타입 계층구조(Type hierarchies) 정의

타입 계층구조를 이해하기 위해서는 수퍼타입(supertype)이라는 개념을 먼저 이해하여야 한다. 수퍼타입과 서브타입이라는 개념은 토픽맵 온톨로지 내에서 Topic Type간의 계층관계를 기술하는데 자주 사용되는 개념이다.

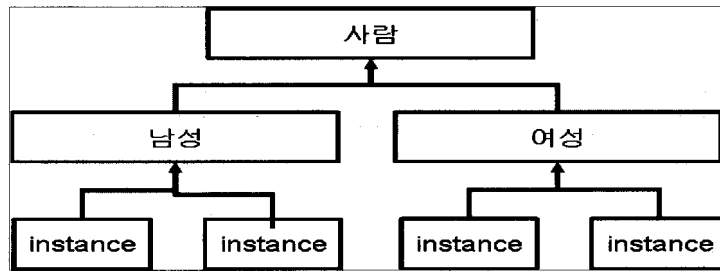
예를 들어 '사람', '남성', '여성'이라는 Topic Type들이 있을 때, '모든 남성은 사람이다.', '모든 여성은 사람이다.', '모든 사람이 남성은 아니다.', '모든 사람이 여성은 아니다.' 등은 이 세 가지 Topic Type들이 서로 관계가 있음을 보여주며, 이 경우 '사람'이라는 Topic Type은 '남성'과 '여성'의 수퍼타입이 된다. 이처럼 타입간의 관계를 최대한 명확하게 해주는 것은 토픽맵 온톨로지 모델링 시에 가장 권장되는 것이기도 하다.

1) 타입 계층구조의 생성 이유

타입 계층구조를 생성하는 이유는 Topic 간 관계의 명확화, 검색 시의 효율성, 제약 사항의 손쉬운 적용 등 세 가지를 꼽을 수 있다. 계층구조가 많은 타입들을 포함하고 있을 경우, 수퍼타입과 서브타입 관계의 추가는 Topic들 간의 관계를 더욱 명확하게 표현할 수 있도록 한다.

검색의 효율성에 있어서는, 예를 들어, '사람'이라는 Topic Type에 속한 모든 인스턴스를 검색하고자 할 때, 모든 Topic 인스턴스들이 '사람'의 하위 타입인 '남성'과 '여성'이라는 토픽 타입의 인스턴스로서 직접 정의되어 있기 때문에 '남성'과 '여성'타입 각각에 대한 인스턴스를 구해야 한다고 생각할 수도 있다. 그러나 '사람'이라는 토픽 타입은 '남성'과 '여성'의 수퍼타입으로 정의되어 있기 때문에 '사람'이라는 토픽 타입의 인스턴스를 구하는 것만으로도 원하는 결과를 얻을 수 있다. 이것은 온톨로지가 복잡해지고 계층구조가 복잡해질수록 더욱 큰 강점으로 작용하게 된다.

제약사항을 손쉽게 적용할 수 있다는 것도, 예를 들어, '남성'과 '여성'이라는 Topic Type 모두가 '출생일'이라는 Occurrence를 포함하도록 하고 싶을 때는 각각의 타입에 '출생일'이라는 Occurrence를 추가시키는 것이 아니라, 수퍼타입인 '사람'에 Occurrence를 추가하는 것으로 간단히 문제를 해결할 수 있다. 이 역시 온톨로지가 복잡해지고 커질수록 큰 강점으로 작용하게 된다.



[그림 7] 타입 계층구조 그림 예시

2) 타입 계층구조의 기술 방법

Topic Type들 간의 계층구조, 즉 슈퍼타입과 서브타입의 관계도 Association Type을 통해 기술된다. 계층구조를 기술하는 관계와 역할은 XTM 1.0 PSI(Published Subject Identity, 이하 PSI)를 통해 정의되어 있으며, PSI는 다음과 같다.

<http://www.topicmaps.org/xtm/1.0/core.xtm#superclass-subclass>

<http://www.topicmaps.org/xtm/1.0/core.xtm#superclass>

<http://www.topicmaps.org/xtm/1.0/core.xtm#subclass>

PSI를 사용해 기술된 계층구조는 XTM 표준에 의거한 모든 소프트웨어에서 인식된다.

3) 타입 계층구조의 작동 방식

계층구조를 잘못 사용할 경우 온톨로지의 구조적인 결함을 야기할 뿐만 아니라 최종적으로 어플리케이션 이용자에게까지 영향을 미칠 수 있다. 예를 들어, 어플리케이션에서 tolog와 같은 질의어를 사용하여 계층구조에 대한 질의를 수행할 경우, 잘못된 계층구조는 잘못된 결과를 출력하게 된다.

타입 계층구조에서 주의할 경우의 예를 들면, A라는 Topic이 B라는 Topic의 슈퍼클래스라고 할 때, A와 B는 모두 클래스여야만 하며, C라는 Topic이 B의 인스턴스라면 이것은 또한 A의 인스턴스이기도 하다는 것이다. D라는 토픽이 B의 서브클래스라면 D는 A의 서브클래스이기도 하다.

토픽맵 요소들의 가장 중요한 특징 중의 하나는 Topic이 클래스(또는 타입)임과 동시에 인스턴스일 수 있다는 것이다.

3.4.6 주제식별자(Subject identity) 정의

제작중인 온톨로지를 미래에 다른 토픽맵 온톨로지와 통합할 가능성이 있거나 통합하고자 할 계획이 있을 경우, 주제 식별자를 표준에 맞춰 작성하는 것이 요구된다. 표준에 따라 작성한다는 것은 공적 주제(PS : Published Subject)를 사용하여 주제 식별자를 기술한다는 것을 의미한다.

현재 설계하는 토픽맵 온톨로지와 동일한 주제 영역을 다루는 PS가 이미 존재하고 있을 경우, 그것을 재사용하는 것이 바람직하다. PSI를 활용하여 각 Topic에 대한 식별자를 기술하면 추후 별다른 노력 없이 토픽맵 온톨로지 사이의 통합(merge)을 수행할 수 있다.

활용할 수 있는 PSI가 없을 경우, 새로운 PSI를 만들어 웹을 통해 공표하는 것도 바람직한 방법이다. 이는 주제 영역에 대한 식별자를 자유롭게 정의할 수 있음을 의미하며, 다른 온톨로지 디자이너들이 이를 재사용할 수 있도록 함으로서 토픽맵 전반에 관한 보급과 발전에도 기여하는 것이기 때문이다.

3.4.7 속성(properties) 정의

각각의 Topic들은 속성과 속성 값으로 기술되며, 이 때 속성이란 Occurrence Type을 의미하고, 속성 값은 실제 값으로 입력되는 데이터를 의미한다.

3.5 토픽맵 온톨로지의 정제

이 과정에서 이루어져야 하는 핵심적인 사항은 네 가지 정도로 요약할 수 있다.

- Topic Name에 관한 모델링
- 정의한 Topic Type들에 관련한 제약사항의 세밀한 기술
- 정의한 Association Type들에 관련한 제약사항의 세밀한 기술
- 정의한 Occurrence Type들에 관련한 제약사항의 세밀한 기술

3.5.1 Topic Name

1) 무범위(unscooped) 이름

만약 Topic에 특별히 범위가 적용되지 않은(unscooped) 무범위 이름이 주어졌다면, 이 이름이 주어진 Topic의 기본 이름이다. 각각의 Topic에 대해 이름을 정의할 때 무범위 이름이 두 개 이상 정의되지 않도록 주의해야 한다. Topic에 정해진 기본 이름이 없거나 정할 수 없다면, 범위가 적용된(scoped) 이름만을 정의할 수도 있다.

2) 이름 충돌

토픽맵 온톨로지 내에서 두 개의 서로 다른 Topic이 같은 이름을 가지는 것이 허용되기 때문에 이름 충돌이 발생할 수 있다. 온톨로지를 기술하는데 있어서는 문제가 없다고 하더라도 실제 어플리케이션을 사용하는 사용자에게는 혼란이 야기될 수 있다.

중복된 이름에 의한 혼란을 해결할 수 있는 하나의 방법은, 토픽맵 온톨로지 전체에 적용되는 이름 명명 규칙을 정의하는 것이다. 예를 들어 Paris라는 단어에 대해 Paris(city), Paris(Greek Hero) 식으로 이름을 명명하도록 한다면 이름 충돌 문제를 해결할 수 있을 것이다. 또 다른 방법은 사용자 인터페이스 상에 불확실성 및 중복을 제거하는 로직을 추가하는 것이지만, 이는 어플리케이션 부분에 복잡한 코드 작성이 필요할 뿐만 아니라 설계한 로직이 가능한 모든 이름 중복 상황에 대해 완벽하게 작동한다는 것도 확신할 수 없기 때문에, 명확한 이름 명명 규칙을 정하여 기술하는 것이 가장 효과적인 방법이라고 할 수 있다.

3) 다양한 형태의 이름

토픽맵에서는 하나의 Topic에 대해 다양한 형식의 이름이 가능하도록 하는 variant name이라는 개념이 존재한다. 예를 들어, 검색 시 효율적인 정렬을 위한 정렬 이름을 지정할 수도 있으며, 복수의 표현 및 발음, 다른 언어로 표현된 이름 등을 기술할 수 있다. 하나의 Topic에 대해 이러한 다양한 이름을 기술할 수 있다는 것은 큰 장점이다. 예를 들어, 모든 Topic Name에 여러 언어로 표현된 이름들을 준다면, 이를 바탕으로 이들 언어로 표현된 어플리케이션을 손쉽게 제작할 수 있다. 정렬이름을 따로 정해줌으로써 검색 시 토픽들이 올바른 순서대로

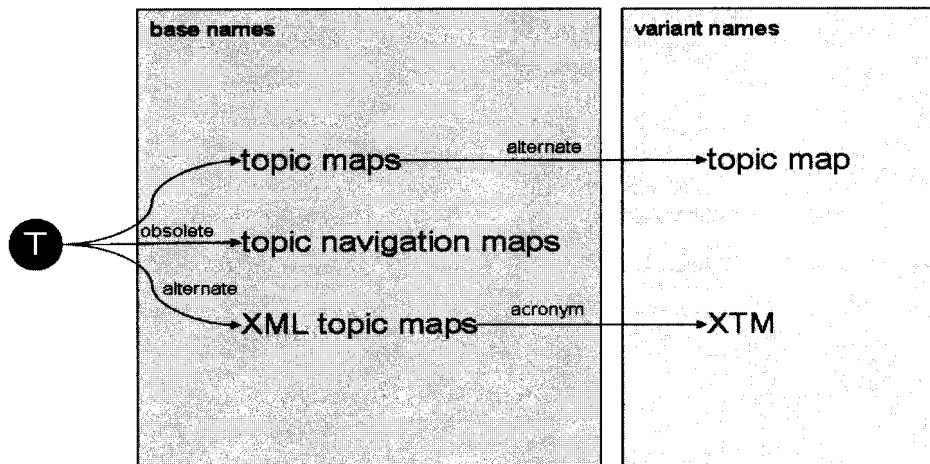
정렬되도록 할 수도 있으며, Topic Name에 대해 흔히 잘못 사용되는 이름들을 기술함으로써 이용자가 잘못된 이름으로 검색을 수행해도 찾고자 하는 정보를 정확히 찾을 수 있도록 할 수도 있다.

4) 이름과 검색

토픽맵에 대한 full-text 검색의 결과는 Topic Name과 가장 일치하는 순서로 나타나므로, 검색의 정확도를 높이기 위해서는 가능하면 각각의 Topic들에 대해 다양한 이름을 기술하는 것이 좋다.

예를 들어, 이용자가 '토픽맵'에 대해서 검색하고자 할 경우, 대부분의 이용자가 '토픽맵'이라는 단어를 검색할 것이지만, 'topic map', 'topicmap', 'topic navigation map', 'xtn', 'XML Topic Map' 등의 단어를 이용하여 검색하는 이용자들을 위하여 이들 모두를 다 기술해 준다면, 검색의 시간도 단축시킬 수 있을 뿐만 아니라 정확도도 크게 향상시킬 수 있을 것이다.

하지만 모든 어플리케이션에서 이름의 다양한 표현이 요구되는 것은 아니다. 지식 관리나 정보 검색과 관련해서는 한 Topic에 대한 다양한 이름의 표현이 상당히 유용하지만, 제품 설정과 비즈니스 프로세스 모델링 관련해서는 그렇지 못한 경우가 있다.



[그림 8] 토픽 이름 구조 예시

3.5.2 타입 제약사항 정의

Topic Type, Association Type, Occurrence Type에 제약사항을 정의한다는 것은 각각의 타입에 속하게 될 인스턴스들이 어떻게 기술될 것인지를 정의한다는 것이다. 현재는 토픽맵 온톨로지를 기술할 때 제약사항을 구체적으로 기술하지 않는 것이 일반적이지만, 모든 타입을 기술할 때에는 해당 타입의 제약조건이 함께 기술되는 것이 바람직하다. 제약사항은 타입의 정의를 보다 명확하게 함으로써 온톨로지 구조 자체의 정확성을 높이는 효과가 있으며 온톨로지 개발과 관련된 문서화 작업을 수월하게 하기 때문이다.

토픽맵 온톨로지의 제약사항을 상세하게 기술할 수 있는 TMCL이 있으나 아직 표준화되지 않았으므로, TMCL 또는 자체적인 기술 지침 등을 마련하여 타입과 제약사항을 명확하게 기술할 수 있도록 한다.

1) Topic Type 제약사항 정의

- 범위에 따라 적용되는 Topic Type 이름
- 범위에 따라 적용되는 Topic Name의 개수
- 범위에 따라 적용되는 Occurrence Type과 카디널리티(cardinality)
- 특정한 클래스와의 관계에서 인스턴스가 맡게 될 Association Role

2) Occurrence Type 제약사항 정의

- Occurrence Type이 사용될 Topic Type들
- Occurrence Type에 적용될 범위
- 범위에 따라 적용되는 Occurrence Type 이름

3) Association Type 제약사항 정의

- Association Type에 적용된 범위
- Association Role과 카디널리티
- Role로 허용된 Topic Type의 조합
- 관계의 방향에 따른 라벨
- 범위에 따라 적용되는 Occurrence Type과 카디널리티

3.6 토픽맵 온톨로지 디자인 구현

이제는 현재까지의 디자인을 컴퓨터가 처리할 수 있는 형태의 온톨로지 디자인으로 구현해야 할 차례이다. 토픽맵 온톨로지의 디자인을 구현하는 방법은 크게 세 가지 형태로 이루어질 수 있다.

- 텍스트 편집기, XML/SGML 편집기를 이용한 XTM 코드 문서의 작성
- XTM으로 자동 변환될 수 있는 의사(pseudo) 코드의 작성
- 토픽맵 디자인 인터페이스를 제공하는 토픽맵 소프트웨어를 이용한 작성

텍스트 편집기를 이용하여 온톨로지 디자인을 XTM으로 기술한다는 것은 매우 힘든 작업일 뿐만 아니라 많은 문법상의 오류가 발생할 수 있다. pseudo 코드를 활용하여 구현하는 것은 XTM으로 기술하는 것보다 훨씬 간단하긴 하지만 작성한 코드가 올바르게 XTM으로 변환되기 위해서는 pseudo 코드를 정확하게 기술해야 하고 이를 위해서는 별도의 문법을 익혀야 하기 때문에 부담으로 작용할 수 있다. 하지만 간단하고 작은 규모의 토픽맵을 기술할 때는 효과적인 방법일 것이다.

온톨로지 디자인을 구현하는 가장 효율적인 방법은 토픽맵 디자인 인터페이스를 제공하는 토픽맵 에디터나 소프트웨어를 이용하여 온톨로지를 디자인하는 것이다. 이러한 소프트웨어는 토픽맵의 복잡한 문법적 요소들은 감추고 인터페이스를 통해 핵심적인 요소들만을 표현해 줌으로서 토픽맵 디자이너가 디자인에만 집중할 수 있도록 한다는 장점을 가지고 있다. 그리고 인터페이스를 통해 작성된 토픽맵 온톨로지는 표준 형식인 XTM으로 자동 전환되며, RDF/OWL 등과 같은 다른 문법 형식으로의 변환도 지원한다. 현재 무료로 이용할 수 있는 다양한 종류의 토픽맵 온톨로지 에디터가 존재하므로, 온톨로지 디자이너는 자신의 환경에 맞는 에디터를 구하여 작성하면 될 것이다.

디자인 구현 과정에서 XTM으로 기술할 필요는 없어졌다고는 하지만, 이것이 온톨로지 디자이너가 XTM 문법에 대해서는 알 필요가 없다는 것을 의미하는 것은 아니다. 토픽맵 온톨로지 에디터가 XTM의 모든 스펙을 지원하는 것은 아니며, 실제 구현 과정에서 XTM 파일을 직접 점검해야 하는 상황도 생길 수 있으므로 온톨로지 디자이너는 XTM 스펙에 대해서도 완전한 이해를 갖추도록 해야 한

다.

3.7 인스턴스 생성 및 테스트

토픽맵 온톨로지의 디자인까지 성공적으로 구현되었다면, 이를 토대로 온톨로지의 실제 콘텐츠를 구성하게 될 인스턴스들을 생성할 차례이다. 인스턴스들을 입력하고 실제로 네비게이션 해 보면서 온톨로지 상의 구조적 결함은 없는지, 잘못 정의된 관계는 없는지 등의 문제점들을 찾아내고 전체적인 완성도를 테스트해야 한다.

인스턴스의 생성 과정은 수작업으로 입력하는 것 그리고 기존의 데이터 소스로부터 자동으로 생성하는 것의 두 가지 경우로 생각해 볼 수 있다. 각각의 인스턴스 생성 방법이 어떻게 하면 효과적으로 이루어질 수 있는지를 알아보도록 한다. 더불어 새로운 인스턴스의 등록 및 불필요한 요소의 삭제 등 온톨로지의 유지보수와 관련된 사항도 간략하게 기술하도록 한다.

3.7.1 수동 생성

수동 생성이란 토픽맵 온톨로지의 인스턴스 요소들을 온톨로지 편집자의 수작업에 의해 생성한다는 의미이다. 이것은 많은 시간과 비용이 소요되는 작업이지만 경우에 따라서는 가장 효율적인 생성 방법일 수도 있다.

토픽맵 온톨로지의 규모가 크지 않을 경우 몇 명의 온톨로지 편집자의 작업만으로도 충분히 원하는 온톨로지를 생성할 수 있다. 또한 수작업을 통해 작성된 것이어서 온톨로지의 질적인 측면이나 정확도는 자동 생성을 통해 생성된 온톨로지보다 높다고 할 수 있다. 데이터 소스의 양이 적거나 온톨로지로 변환할 수 있는 적합한 데이터 소스가 존재하지 않는 경우에 수동 생성을 사용한다.

수동 생성에 있어서 중요한 점은 인스턴스의 생성 과정을 얼마만큼 효율화시킬 것인가에 있다. 즉, 온톨로지 편집자들이 직관적이고 쉽게 인스턴스를 생성할 수 있는 이용자 인터페이스를 제공하는 것이 수작업에 의한 인스턴스 생성에 큰 영향을 미치는 요소이다.

현재 많은 토픽맵 온톨로지 에디터들이 웹 기반 또는 폼 기반의 환경을 제공함으로써 편리한 인스턴스 생성을 지원하고 있다. 일부 상용 툴들은 프로젝트에 따

라 온톨로지 편집 인터페이스를 직접 설계할 수 있는 개발 환경을 제공하므로 이러한 툴들을 활용하여 토픽맵 편집자들이 인스턴스를 생성할 수 있는 환경을 구축할 수도 있을 것이다.

3.7.2 자동 생성

일정 규모 이상의 프로젝트의 경우 수작업을 통해 온톨로지 및 인스턴스를 생성하는 것은 막대한 시간과 비용 그리고 인적 자원의 투입을 필요로 한다. 그러므로 가능하면 인스턴스의 생성을 자동화시킬 수 있는 방안을 찾으려 해야 한다.

여기서 자동 생성에 대한 몇 가지 지침을 제공하지만, 자동 생성은 프로젝트의 규모, 작성될 토픽맵 온톨로지의 복잡성, 그리고 기존에 존재하고 있는 데이터 소스의 종류 등 여러 요소들에 크게 영향을 받기 때문에, 온톨로지 다자이너는 현재 진행하고 있는 프로젝트에 대해 이러한 요소들을 충분히 고려하여 가장 최적의 자동 생성 지침 및 방법론을 작성, 적용할 수 있도록 하여야 한다.

앞서 토픽맵 온톨로지 디자인을 위한 분석 과정에서, 이미 존재하고 있는 정보 시스템이나 정보자원들의 형태 및 구조를 세밀하게 분석해야 한다는 것을 언급하였으며, 이러한 분석은 자동화 과정을 통해 변환될 수 있는 인스턴스들을 식별하는데 큰 도움이 된다.

1) 활용 가능한 정보 자원들

토픽맵 온톨로지 디자인 단계의 전 단계인 기존 시스템 및 자료 분석 단계에서 도출된 정보자원들이 자동 생성을 위한 기초 자료로 사용될 수 있다. 구조화된 데이터는 물론 비구조화된 데이터도 자동 생성을 위해 사용될 수 있지만, 모든 데이터들은 컴퓨터로 처리 가능한 형태이어야 한다. 활용 가능한 데이터 소스들의 종류는 다음과 같다.

- RDF, RDF-Schema, DAML+OIL, OWL 등으로 기술된 온톨로지
- 주제분류, 텍사노미, 목차
- 관계형 데이터베이스, 객체지향형 데이터베이스
- 정보자원에 포함된 다양한 형태의 메타데이터(파일 속성 정보, 이메일 헤더 등)

- 인덱스, 시소러스, 데이터 사전
- 문서구조(XML DTD, XML Schema)
- 비구조화된 문서(HWP, Word, PDF 문서 등)

2) 온톨로지

이미 존재하는 온톨로지를 토픽맵 온톨로지로 변환하는 것은 자동 생성의 형태 중 가장 쉬운 작업일 것이다. 이미 존재하는 온톨로지는 이미 많은 노력이 투입되어, 현재 작성하려는 토픽맵 온톨로지에 근접할 정도로 작성되어 있을 확률이 높으며, 토픽맵 모델과 유사한 개념들, 즉, 클래스(타입), 인스턴스(Topic), 슬롯(Association 또는 Occurrence), 클래스 계층구조 등으로 기술되어 있을 수 있다.

온톨로지 변환을 수행할 변환 모듈의 개발을 고려할 수도 있다. 변환 모듈은 특정한 온톨로지를 대상으로 개발될 수도 있으며, 온톨로지 기술 언어의 종류에 따라 개발될 수도 있을 것이다.

특정한 온톨로지를 대상으로 개발된 변환 모듈은 요소들을 토픽맵 온톨로지보다 더 정확하게 변환할 수 있지만, 추후 다른 프로젝트에서의 재활용성은 떨어질 것이다. 반면 온톨로지 기술 언어의 종류에 따라 개발된 변환 모듈은 차후에도 높은 재활용성을 보이지만, 요소 추출에 대한 정확도는 떨어질 수 있다.

3) 분류 스키마(주제 분류, 텍사노미, 목차)

분류 스키마는 클래스, 클래스 코드, 클래스 이름, 클래스 상세기술, 슈퍼클래스-서브클래스 관계 등의 요소들을 사용하여 기술되어 있으며, 서로 다른 계층구조 아래에 있는 클래스 간의 연결을 기술하기도 한다. 분류 스키마에 따라 분류된 정보 자원들은 정보 자원의 메타데이터로서 클래스 코드를 포함하고 있거나 스키마에 직접적으로 링크되어 있을 수도 있다.

[표 3]은 분류 스키마의 요소들과 매핑될 수 있는 토픽맵 요소들을 정리한 것이다.

[표 3] 분류스키마와 토픽맵의 매핑요소

| 분류 스키마 | 토픽맵 |
|------------------|-----------------------------|
| 클래스 | 분류와 연관된 타입 토픽 |
| 클래스 코드 | 토픽의 주제 식별자 |
| 클래스 이름 | 토픽 이름 |
| 클래스 상세정보 | '상세 기술'을 의미하는 토픽 타입의 어커런스 |
| 수퍼클래스-서브클래스 관계 | 수퍼클래스-서브클래스 관계(association) |
| 서로 다른 계층구조 간의 연결 | '참조'를 의미하는 토픽 타입의 관계 |
| 분류된 정보 자원들 | '분류된 자원'을 의미하는 토픽 타입의 어커런스 |

4) 데이터베이스 스키마

일반적으로 현재 가장 많이 사용되고 있는 관계형 데이터베이스는 테이블, 열/행, 기본키, 외래키 등의 요소들로 구성되어 있으며, 이러한 요소들의 분석을 통해 변환 가능한 요소들을 찾아낼 수 있다.

예를 들어 A라는 테이블이 의미하는 바가 B라는 Topic Type과 일치하고, A 테이블 내의 열들이 이름, 키 등의 정보를 담고 있다면, 이를 통해 B에 속하는 인스턴스 Topic들을 자동 생성할 수 있을 것이다. 관계형 데이터베이스의 요소와 토픽맵 개체들의 일반적인 매핑은 다음과 같다.

[표 4] 관계형 데이터베이스와 토픽맵의 매핑요소

| 관계형 데이터베이스 | 토픽맵 |
|------------|------------|
| 테이블 | 토픽 타입 |
| 행 | 토픽 인스턴스 |
| 열 | 이름 또는 어커런스 |
| 키 | 토픽 아이디 |
| 외래키 | 관계 |

5) 메타데이터

메타데이터는 정보자원의 일부분으로 포함되어 있을 수도 있고, 정보자원과

는 별개로 독립적인 저장소에 기술되어 있을 수도 있다.

정보자원의 일부분으로 함께 기술되는 메타데이터에는 제목, 저자, 키워드, 버전, 생성날짜, 출판일 등의 정보가 있으며, MS 오피스 문서나 PDF 문서, 이메일 헤더, HTML 메타 태그 등에서 이러한 메타데이터 정보들을 추출할 수 있다. 이러한 메타데이터는 MS 오피스에서 제공하는 비주얼 베이직 API, PDF 또는 이메일의 스캐닝, HTML과 XML 파서를 이용한 문서 파싱 등을 통해서만 접근이 가능하다.

이렇게 메타데이터를 포함하고 있는 개별 정보 자원들은 토픽맵 내에서 Topic 인스턴스로 생성될 수 있으며, 그 문서의 제목과 같은 메타데이터는 Topic의 Name으로 변환될 수 있다. 리소스 파일 자체는 그것 자체로 external occurrence가 될 수 있고, external occurrence에 대한 경로는 주제 식별자로서 활용할 수 있으며, 다른 메타데이터 필드도 Occurrence, Association 등으로 변환할 수 있다.

정보자원과는 별개로 독립적인 저장소에 기술되는 외부 메타데이터의 가장 일반적인 예는 하드디스크 내의 파일시스템이다. 디렉토리는 Topic Type과 매핑될 수 있으며, 서브디렉토리는 서브 클래스와 매핑될 수 있다. 디렉토리 내에 포함된 파일은 Topic Type에 속하는 Topic 인스턴스로 표현할 수 있을 것이며, 파일 이름은 Topic Name으로 변환될 수 있다.

파일 이름을 제외한 디렉토리 경로는 파일에 대한 주제 식별자로, 파일 확장자는 Occurrence Type이나 Topic Type으로 변환될 수 있다.

문서관리 시스템이나 콘텐츠 관리 시스템과 같은 복잡한 저장소들은 파일 시스템에 관한 메타데이터뿐만 아니라 개별 정보자원에 대한 속성도 제공하는 것이 대부분이므로, 내부/외부에 존재하는 메타데이터를 세밀하게 분석하는 것이 요구된다.

6) 인덱스, 시소러스, 데이터 사전

인덱스, 어휘집, 시소러스 또는 데이터 사전은 기본적으로 양질의 Topic들을 포함하고 있다. 인덱스를 통해 Topic Type, 다양한 유형의 Topic Name, 타입이 정의되지 않은 관계 등 포함될 요소들의 정보를 얻을 수 있을 것이며, 어휘집을 통해서도 타입이 정의되지 않은 관계에 포함될 요소에 관한 상세한 정보를 얻을 수 있을 것이다.

데이터 사전은 해당 주제 영역의 중요 개념 정의, 유의어, 번역된 어휘 등을 풍부하게 포함하고 있다는 점에서 토픽맵으로 변환 가능한 인스턴스 정보들을 추출해 낼 수 있는 가장 완벽한 소스라고 할 수 있다.

7) 구조화된 문서(XML DTD, XML Schema)

계층구조를 통해 기술되는 XML 문서는 엘리먼트, 서브 엘리먼트, 엘리먼트 콘텐츠, 속성, 속성 값 등의 요소들을 사용하여 기술되며, 이러한 각각의 요소들로부터 타입 및 인스턴스 요소들을 추출해 낼 수 있다.

엘리먼트와 속성을 통하여 Topic Type, Occurrence Type, Association Type들과 매핑되는 요소들을 파악할 수 있으며, 엘리먼트 콘텐츠와 속성 값들을 통해 Topic 인스턴스와 Occurrence 요소들을 자동 생성할 수 있다.

엘리먼트와 서브 엘리먼트 관계는 토픽맵 온톨로지에서의 수퍼클래스-서브클래스 관계나 Occurrence들을 내포하고 있을 수도 있으므로, 이를 명확하게 파악하는 것이 중요하다.

8) 비 구조화된 문서(HWP, Word, PDF 문서 등)

비구조화된 문서는 자동 생성에 이용할 수 있는 데이터 소스들 중 가장 처리가 어려운 데이터 유형이다. 일반적으로 비구조화된 문서란 MS 워드나 PDF, 이메일 등을 의미하며, 그 중에서도 메타데이터가 풍부하게 기술되지 않아 자동화 과정에서 약간의 문서 구조 정보나 문서 내의 텍스트 외에는 참조할게 없는 데이터 유형을 의미한다.

문서 내의 텍스트를 분석하는 것은 자연어 처리의 영역이지만, 단순한 텍스트 패턴을 통해서도 유용한 결과들을 이끌어낼 수 있다. Topic을 추출하는 것은 Association을 추출하는 것보다는 쉬운 과정이다. 통계에 기반한 알고리즘적 접근이 Association을 추출하는데 많은 도움이 될 수 있다. 하지만 이러한 과정을 거쳐 추출한 Association을 다룰 때는 그 결과의 정확도와 유효성에 대해 주의를 기울여야 한다. 알고리즘 통해 완벽한 추출에 도달한다는 것은 불가능하므로, 가장 적합한 방법은 추출 결과물을 직접 살펴보는 것이다.

제4장 학위논문 온톨로지 구현방안 연구

4.1 토픽맵 기반 학위논문 콘텐츠 메타데이터 분석 연구

4.1.1 메타데이터 항목 추출 기반 기술조사

1) 더블린 코어(Dublin Core)

- 더블린 코어의 정의

더블린 코어는 자원의 신속한 검색을 목적으로 1995년 더블린(Dublin)에서 OCLC(Online Computer Library Center)와 NCSA(National Center for Supercomputing Applications) 기관이 개최한 워크숍에서 합의된 메타데이터를 말한다. 데이터의 호환성을 유지하고 네트워크상의 자원의 기술에 필요한 일련의 데이터 요소를 16개의 요소와 한정자들로 규정하고 있으며 메타데이터 표준은 광범위한 자원을 설명하기 위한 단순하면서도 효과적인 방법이다. 즉 자원의 기술(Resource Description)이 아니라 자원의 발견(Resource Discovery)에 초점을 맞추고 있다.

- 더블린 코어의 목적

데이터의 형식과 구조를 단순화하여 원문의 저자나 발행자가 메타데이터를 직접 작성하고, 네트워크 출판을 위한 저작도구의 개발자가 이 정보에 대한 템플릿을 직접 소프트웨어에 포함할 수 있도록 하며, 의미가 보편적으로 이해되고 그 쓰임이 국제적 범위가 되어야 한다. 작성된 데이터를 기초로 특정 분야에서 요구되는 상세한 수준으로 확장하여 사용하기 위한 것이다.

- 기본원칙

고유성 : 자원의 본질적인 특성을 기술요소로 한다.

확장성 : 규정된 필수데이터 요소 이외에 부차적인 내용이나 특성을 기술요소로 사용한다.

구문의 독립성 : 응용분야나 표현기법을 규정하지 않는다.

선택성 : 각 요소의 수록여부를 강제하지 않는다.

반복성 : 모든 기술요소는 반복사용 가능하다.

수정가능성 : 한정어를 사용하여 세부사항을 조정한다.

- 더블린 코어의 구성요소

Elements, Refinements, Encoding Scheme으로 구성되는데 더블린 코어 메타데이터가 ISO 표준으로 채택되었고 갈수록 많은 나라에서 상이한 기관간의 통합검색에 필요한 기본 메타데이터로 사용하고 있는 실정이다. 더블린 코어는 시맨틱 웹 기술의 핵심인 네임스페이스와 URI 체계를 적극적으로 적용하여 메타데이터를 관리하고 있는 모범 표준이다.

[표 5] 더블린 코어의 구성 요소

| 요소 | 속성 | 한국어 | 표준 | 정의 및 사용방법 |
|---------|-----------------|-----------|-----------------------------------|--|
| Title | | 표제 | | 자원에 부여된 제목이나 자원을 공식적으로 지칭하는 이름 |
| | alternative | 대체표제 | | 자원에 부과된 다른 표제명: 부표제, 번역표제 등 |
| Creator | | 생성자 | | 자원의 내용에 주된 책임을 갖고 있는 개체의 이름 |
| Subject | | 주제 | LCSH MeSH DDC LCC UDC | 자원의 내용이 지닌 주제. 자원의 주제를 표현하는 키워드, 구절, 또는 분류코드로 통제어휘나 공식적 분류 스킴에 의한 값을 사용하는 것을 권장함 |
| Date | | 날짜 | W3CDTF | |
| | created | 제작일 | | 자원의 생성일 |
| | valid | 유효일 | | 자원의 유효기간을 보통 시작일과 종료일로 표시 |
| | available | 이용가능일 | | 자원의 이용 가능일을 범위로 표시 |
| | issued | 발행일 | | 자원이 공식으로 발행된 일자 |
| | modified | 수정일 | | 자원 수정일 |
| | dateAccepted | 수용일 | | 자원 수용일 |
| | dateCopyrighted | 지적재산권 획득일 | | 자원의 지적재산권 취득일 |
| | dateSubmitted | 제출일 | | 자원 제출일 |

| | | | | |
|-------------------|-----------------------|----------|---------------|--|
| Type | | 자료유형 | | 자원내용의 유형 또는 장르. 통제어휘의 사용을 권장함 |
| Format | | 자원포맷 | | |
| | extent | 자원크기 | | 자원의 파일 크기나 실행 기간을 입력 |
| Identifier | | 식별자 | | 자원의 접근을 보장하는 고유의 식별자 |
| | bibliographicCitation | 서지정보 | | 자원의 서지정보를 입력할 수 있음 |
| Source | | 출처 | URI | 자원의 출처가 되는 원 정보의 접근을 보장하는 식별자. 값으로는 URI로 표현되어야 함 |
| Relation | | 관련정보 | | |
| | isPartOf | 전체/부분 관계 | URI | |
| | hasPart | 전체/부분 관계 | URI | |
| | isVersionOf | 버전관계 | URI | |
| | hasVersion | 버전관계 | URI | |
| | isFormatOf | 포맷변환 관계 | URI | |
| | hasFormat | 포맷변환 관계 | URI | |
| | isReferenceOf | 참조관계 | URI | |
| | references | 참조관계 | URI | |
| | isRequiredBy | 의존관계 | URI | |
| | requires | 의존관계 | URI | |
| | isReplacedBy | 대체관계 | URI | |
| | replaces | 대체관계 | URI | |
| conformsTo | 준수관계 | | 준수하고 있는 표준 내용 | |
| Coverage | | 범위 | | |
| | spatial | 공간적 범위 | TGN | 자원의 내용이 어떤 지역적 범위에 |

| | | | | |
|-----------------|----------------|--------|--------|------------------------------------|
| | | | | 대한 언급을 하고 있는 경우 |
| | temporal | 시간적 범위 | W3CDTF | 자원의 내용이 어떤 시간적 범위에 대한 언급을 하고 있는 경우 |
| Rights | | 이용조건 | | 자원의 접근을 보장하는 고유의 식별자 |
| | rightsAccess | 이용권 접근 | | 자원의 서지정보를 입력할 수 있음 |
| Language | | 언어 | | 자원의 언어에 대한 언급 |
| Audience | | 지원대상 | | |
| | educationLevel | 교육수준 | | 자원이 지향하는 교육수준 |
| | mediator | 대상자 | | 자원이 목표하는 대상 |

더블린 코어는 국제표준이면서 콘텐츠를 표현하는 많은 분야에서 표준으로 채택하여 네임스페이스로 사용하고 있는 메타데이터이다. 도서관의 서지요소와 더블린 코어의 구성요소, 하위요소 및 한정어에 매핑되는 부분을 많이 발견할 수 있으므로, 서지MARC 데이터를 XML형식으로 반출할 경우, 더블린 코어를 네임스페이스로 하여 메타데이터의 표준을 준수하고 있다. 따라서 본 연구과제의 온톨로지 기반 학위논문 서비스에서도 학위논문 콘텐츠에 대한 서지적 메타데이터 요소에 대해서는 더블린 코어의 구성요소를 도입하여 토픽맵의 PSI를 부여함을 원칙으로 한다.

2) MARC

MARC는 MACHine Readable Cataloging 즉, 기계가독형 목록으로 도서관이 전산화되기 시작하면서 도서 북카드를 대신하는 서지목록으로서 표준을 유지해 오고 있는 도서관에서의 대표적인 메타데이터 형식이다.

도서의 서지목록은 기본기입과 부출, 서명저자사항, 발행사항, 형태사항, 주기사항, 연관저록 등을 000부터 999까지의 tag로 지정하고, 각 tag에서 subfield로 서지목록에 필요한 각 항목을 기술하도록 정의되어 있다.

▪ MARC의 구성

Leader : 레코드 유형
 01X-09X tag : 숫자적 제어필드
 1XX tag : 기본표목
 20X-24X tag : 서명 및 서명관련사항
 250-29X tag : 판차, 발행 등 사항
 3XX tag : 형태사항 등
 4XX tag : 총서사항
 5XX tag : 주기사항
 6XX tag : 주제명부출표목
 70X-740 tag : 부출표목
 76X-79X tag : 연관저록
 80X-830 tag : 총서부출표목
 850-890 tag : 총서사항

LC에서 개발한 MARCII는 미국의 국가적인 표준인 USMARC Format for Bibliographic Data 뿐만 아니라 세계 각국의 MARC 포맷 개발의 기초가 되었다.

USMARC는 ANSI(American National Standards Institute) Z39.2-1979 : American National Standards for Bibliographic Information Interchange 라는 표준(1971년 최초 승인)으로 제정되었으며, 더불어 ISO(International Organization for Standardization) 2709-1981 : Documentation Format for Bibliographic Information Interchange on Magnetic Tape 라는 국제적인 표준(1973년 최초 승인)으로도 제정되었다.

서지포맷 USMARC for Bibliographic Data는 단행본, 문서 및 사본자료, 컴퓨터파일, 지도, 악보, 시각자료와 연속간행물 등에 관한 서지적 정보를 표현할 수 있도록 설계되어 있다. 또한 미국의 MARC는 서지포맷 이외에도 전거레코드를 위한 전거포맷 소장 정보와 위치를 표시하는 소장포맷이 있다. ANSI Z39.2와 ISO 2709표준은 USMARC 데이터의 상호교환용 포맷을 위한 레코드구조의 요건을 명시한 것이며, 서지적 기술의 내용에 대한 구체적인 사항들은 언급하고 있지 않다.

또한 MARC 형식에 있어서, 서지목록의 자료유형에 따라 단행본, 연속간행물, 컴퓨터파일, 음악자료, 지도자료, 비주얼자료 등으로 구분하고 각각의 MARC형식을 별도로 가지고 있었다.

전자자료가 도입되면서 그 표현형식도 다양화되어 기존의 MARC형식을 벗어나는 경우가 발생하면서 미의회도서관(Library of Congress, 이하 LC)에서 통합형 MARC형식인 MARC21을 발표하여 이를 적용하기 시작하였고, 한국도 통합형 KORMARC을 제정, 발표하였다.

WEB기반의 정보 발달과 표준화의 흐름에 부응하면서 더블린 코어 등 XML기반의 표준규칙을 적용한 메타데이터 기술스키마인 MODS(Metadata Object Description Schema)가 LC에서 소개되었다. MARC를 기반으로 한 LC는 전문가와 도서관 어플리케이션을 위한 메타데이터 객체 기술 스키마(MODS)를 다양한 목적을 위해 이용한다.

MODS는 도서 목록 자료를 위한 MARC21 체재의 부분 집합을 포함한다. MARC21에 근거한 체계에서 코어 분야의 변환은, 원래 자원 기술을 위한 요소셋으로 간단한 기록을 MARC에서 유효한 일반적인 태그를 사용하여 생성함을 의미한다. 미국의회도서관은 MARC를 MARC21로 개발하고 MARCXML 웹사이트로 가능한 소프트웨어와 함께 MODS로 개발했다.

현재 국내의 전체 도서관에서 정보 자원에 대한 기반 메타데이터 기술은 MARC 형식으로 작성되고 있다. 이에 따라 MARC은 온톨로지 기반의 시스템 구축에서도 기초DB가 되는 형식이고, MARC 내에서 메타데이터 요소를 추출하고 매핑해야할 필요성이 절실하다. 그러므로 학위논문 서비스에서도 기 구축되어 있는 MARC 데이터를 기반으로 하여, 메타데이터 요소를 분석하고 추출하여 서비스 방안을 모색하도록 한다.

3) MODS

MODS는 W3C의 XML 개요 언어를 사용하여 표현된다. XML은 문서의 잇점을 제공하고 HTML보다는 융통성을 더 허용하고 선별할 것으로 예상되고 있다. XML 개요 언어를 사용해서, MODS는 주요요소와 하위요소 및 주요요소와 그 속성을 정의한다. 엘리먼트의 내용은 주요요소에서 하위요소로 산재된 자료를 포함할 때 "혼합 내용"을 피하기 위하여 사용된다. 예를 들면, <titleInfo>가 <title>를 위한 하위요소를 포함하는 경우에, <partNumber>, <partName>, 그 후에 <titleInfo>는 더 특정한 요소 <title>, <partNumber>, <partName>을 포함하는 태그표시로 나타낸다.

MODS 문서는 MODS namespace를 나타내는 개요 선언부를 포함한다. 기록의 기술 또는 그룹 안에서 MODS namespace가 나타나기 때문에, 각 성분의 앞에 "mods" 접두어를 사용하는 것이 유용하다.

MODS는 MARC21 도서 목록 체재에서 구성을 재편성한 숫자적인 사용보다 MARC 의미론을 기술하며, 다른 메타데이터보다도 기존의 도서관 자료와 호환이 가능하다. MODS는 다음과 같이 잠재적으로 사용될 수 있었다.

- Z39.50로 차세대 체재 지정
- METS (Metadata Encoding Transmission Standard)에 연장한 스키마
- 메타데이터의 대표성
- XML 문법에 있는 원래 자원 기술 (MARC 의미론을 사용하는)
- XML에 있는 단순화된 MARC 대표성
- 전자 자원을 가진 XML 메타데이터

• MODS의 이점

MODS는 다른 메타데이터 체재를 보완하고자 한다. 몇몇 애플리케이션에서는 사용된 MARC 레코드보다 더 많은 장점을 갖고 있다.

- 구성요소 세트가 더블린 코어보다 더 풍부하다.
- 구성요소 세트는 ONIX보다 더 도서관 자료와 호환이 된다.
- 스키마는 전체 MARCXML 스키마보다 더 최종 사용자 위주이다.
- 구성요소 세트는 전체 MARC 형식보다 더 간단하다.

• MODS의 특징

- 구성요소는 일반적으로 MARC의 의미론을 상속한다.
- 데이터는 일부 재구성되는데, MARC에 있는 몇몇 데이터 요소의 경우는 MODS에서도 한 구성요소로 가져오게 될 것이다.
- MODS는 특정한 분류 코드를 사용하지 않는다.
- 몇몇 구성요소는 선택적 ID 속성을 가지고 구성요소 수준에 연결한다.

• MODS의 제한

MODS는 MARC21 서지목록형식에서 데이터의 부분을 포함한다. MARC에 근거한 체계에서 이용가능한 일반적인 태그보다 더 많이 사용할 수 있도록 만들어졌으며, 간단한 레코드도 허용한다.

그러나, 원래 MARC21 레코드에서 MODS로 변환하는데 있어 특정한 태그의 손실이나 데이터의 손실없이 완전하게 MARC21로 재변환되지는 않는다. 어떠한 경우에는 MARC21로 재전환하는 경우에, MARC가 MODS에서 일반적으로 매핑되지 않고 동일한 분야가 정확하게 정의되지 않을수도 있다. 따라서, 구성요소의 유형은 상세한 ID로 대표한다. 다른 경우에는 MARC에 있는 요소에 MODS에 있는 동등한 요소가 존재하지 않을 수 있다. 이때 MODS로 변환한 특정한 자료는 분실될 수 있다.

[표 6] 더블린 코어 요소와 MODS 성분

| 더블린 핵심요소 | 주 |
|----------|--|
| 제목 | MODS의 <titleinfo> |
| 생성자 | MODS는 <role>에서 포함된 기여자의 유형을 가진 repeated<name>에 있는 모든 이름을 둔다. 필요하면 생성자 또는 기증자 구별을 유지할 위해 사용 <name><namePart><role>creator MODS는 Type 속성에 있는 회의, 법인 개인 것과 같이 이름 허용 MODS는 <name> 하위요소 허용 : <namePart>, <displayForm>, <affiliation>, <role>, <description>, <namePart>의 DC 값. |
| 주제 | MODS에 있는 하위요소 <topic>, <geographic>, <temporal>, <name>, <titleInfo>, <hierarchicalGeographic>, <coordinates> 포함 |
| 묘사 | <note> DC값 |
| 발행인 | DC 값은 간행물 발행의 장소 포함 |
| 기증자 | 생성자 참조 |
| 날짜 | MODS에서 DC값 <dateOther>을 포함 |
| 유형 | MODS는 <typeOfResource> 에 있는 유형을 사용 <genre>에 있는 더 특정한 장르 사용 가능. authority= " dct "를 가진 <genre>를 이용하고 MODS에 의해 typeOfResource 값 사용 |
| 체제 | MODS는 <physicalDescription><form>의 DC 값 |
| 인식기 | MODS는 Type 속성 유형, 예를들면 <identifier type= " uri "를> 지정; "http://"를 시작하는 경우에 <location><url> DC 값 사용 |
| 근원 | <titleinfo><title> 또는 <locaton><url> 참조 |

| | |
|----|------------------------------------|
| 언어 | MODS의 <language> 사용 |
| 관계 | MODS에 있는 자료: relatedItem은 하위요소로 사용 |
| 적용 | MODS에 있는 <subject><geographic> 적용 |
| 권리 | |

Top Level Elements:

| | |
|---------------------|-----------------|
| titleinfo | note |
| name | subject |
| typeOfResource | classification |
| genre | relatedItem |
| originInfo | identifier |
| language | location |
| physicalDescription | accessCondition |
| abstract | part |
| tableOfContents | extension |
| targetAudience | recordInfo |

Root Elements:

- mods
- modsCollection

4) FRBR

FRBR은 Functional Requirements for Bibliographic Relationships의 약자이며, 전통적인 카드목록이나 서지레코드의 산물과 같은 것이 아니라, 이용자의 한결같은 요구와 모든 필수기능을 수행할 수 있는 관련된 데이터의 네트워크 목록이다. 즉 개념적인 모델이라 할 수 있다.

FRBR은 IFLA(International Federation of Libray Associations and Institutiones)에서 제안된 문서에서 그 개념이 시작되었고 목록의 전문분야 - 목

록의 원칙과 표준화 내의 환경변화에 대응하고 목록 코드를 개념화, 실행론, 서지 통합의 일반론 - 등에 대하여 계속적인 연구를 수행해 나가고 있다.

FRBR의 가치는 서지정보가 어떤 기능을 수행할 수 있는지에 대한 참신한 정보를 주는 보편적이고 시스템적인 모델이다. 그렇다면 목록이 이용자에게 주는 가장 중요한 가치는 무엇인가? 가장 가치있는 정보를 더 효율적으로 이용하기 위하여 어떻게 해야 할까? 등의 질문에 FRBR은 답을 줄 수 있다. 기존의 목록은 카드목록에서 MARC 구조에 따라 서지간의 관계를 나타내고 검색결과에서도 간략, 상세서비스의 연계와 소장서지정보의 나열된 구조에서 FRBR은 서지간의 연결구조를 지닌 목록구조로 개념의 변화와 차세대 목록의 기능을 명시하고 있다. 즉 FRBR을 적용하여 자료간의 번역, 개정, 축약판, 각색, 해제판 등의 관계를 파악가능케 한다.

FRBR은 서지정보가 이용자의 요구에 대응하도록 구조화된 프레임을 정의하며 ER모델링을 기반으로 서지정보와 관련된 세 개의 집단과 각 집단의 개체를 각각 정의하고, 각 개체간의 상호 관계와 일련의 속성을 정의한 모형을 정의하고 있다.

FRBR이 말하는 이용자의 4가지 요소는 검색기능의 Find 개체, 이용자가 원하는 확인기능의 Identify 개체, 이용자의 선택기능인 Select 개체, 그리고 구입, 대출 등을 통한 접근방식인 Obtain 개체를 4요소를 정의하고 있다. 여기에 제5요소는 Navigate를 추가할 수 있을 것이다.

FRBR의 제 1집단 개체는 서지레코드에 명시된 지적, 예술적 시도의 결과물로 저작, 표현형, 구현형, 개별자료로 구성된다. 제2집단 개체는 지적, 예술적 콘텐츠에 책임을 지고 있는 개체로 개인, 단체로 구성된다. 마지막 제3집단은 지적, 예술적 시도에 대한 주제를 제공하는 개체로 개념, 대상, 사건, 장소로 구성된다. 이 모든 개체는 저작과 주제관계로 연결되어 있으며 개체간의 다양한 의미적 관계를 효과적으로 표현하고 있다.

FRBR은 위와 같이 다양한 의미적 관계의 개념구조를 가지고 연관관계를 설명하고 서비스하고자 하는데 효과적인 모델이므로, 이후 정보간의 연관관계에 중점을 두는 온톨로지 기반의 서비스 구축에 개념적 기반을 제공할 수 있는 메타데이터 기술형식이라 할 수 있다.

이상에서 서지를 대상으로 하는 메타데이터의 종류에 대하여 살펴보았다. 이들 메타데이터 추출요소들을 기반으로 도서관에서 서비스하고 있는 학위논문 콘텐츠

에 대한 메타데이터 항목을 추출해 보기로 한다.

[표 7] FRBR의 1집단 개체

| 개체 | 설명 | 예시 |
|------------------------|--|------------------------|
| 저작 (work) | 지적으로 예술적으로 독창성을 지닌 창작물 다양한 버전이 존재해도 동일하게 간주되는 창작물 | 동일양식 주 제 장르의 변화 |
| 표현형 (expression) | 하나의 저작을 문자, 숫자, 기호, 이미지 등의 형식 을 통하여 실현한 것 텍스트의 개념과 대등하며, 음향, 이미지와 같은 비 언어적 내용도 포함 | 번역판 개정판 축약판 |
| 구현형 (manifestation) | 한 저작의 표현형을 물리적으로 구현한 것 출판물의 개념과 관련 있지만, 미간행된 자료도 포 괄한 일반적인 용어를 사용 | 복사전송본 재쇄본 필름 복제본 |
| 개별자료 (item) | 하나의 구현형에 사례가 되는 구체적인 자료 | 각 기관 소장자료 |

4.1.2 학위논문 콘텐츠 메타데이터 항목 추출

현재 서울대학교 SOLARSIII에서는 MARC 데이터를 기반으로한 메타데이터를 사용하고 있으며 학위논문을 검색한 결과 상세화면은 다음과 같다.

| | | | | |
|-----------|---|-------------------|------|-----------|
| ● 상세정보 | | | | |
| 자료유형 : | 학위논문(석사) | | | |
| 서명 / 저자 : | 지식서비스산업에서 서비스품질평가의 상대적 차이 : 컨설팅 서비스의 고객군별 차이를 중심으로 = Relative difference in service quality in the knowledge-based industry : an empirical study of consulting service / 강버물. | | | |
| 개인저자 : | 강버물, 1974- | | | |
| 발행사항 : | 서울 : 서울대학교 대학원, 2005. | | | |
| 형태사항 : | vi, 77 p. ; 삽도 ; 26 cm. | | | |
| 학위논문 : | 학위논문(석사)-- 서울대학교 대학원 : 경영학과 경영학전공, (지도교수 : 이유재,) | | | |
| 키워드 : | 서비스 품질, SERVQUAL, 서비스 로열티, 컨설팅 서비스, 신뢰, 지식기반 산업, 지식서비스 산 업, 전문 서비스 기업(PSF), 상호작용품질, 결과 품질, 환경 품질, 산업제 고객군, 소비자 고객군, Service Quality, SERVQUAL, Interaction Quality, Result Quality, Environmental Quality, Trust, Consulting Service, Professional Service Firms, Service Royalty, Knowledge Service Industry | | | |
| 수록매체 : | PDF (PDF: 674kb) | | | |
| ● 소장사항 | | | | |
| 번호 | 소장위치 | 청구기호▲ | 도서상태 | 반납기한/반납시간 |
| 1 | 중앙도서관 학위논문실 | TM 658 2005 G154f | 열람가능 | 불명 |

[그림 9] 학위논문 검색결과 상세화면

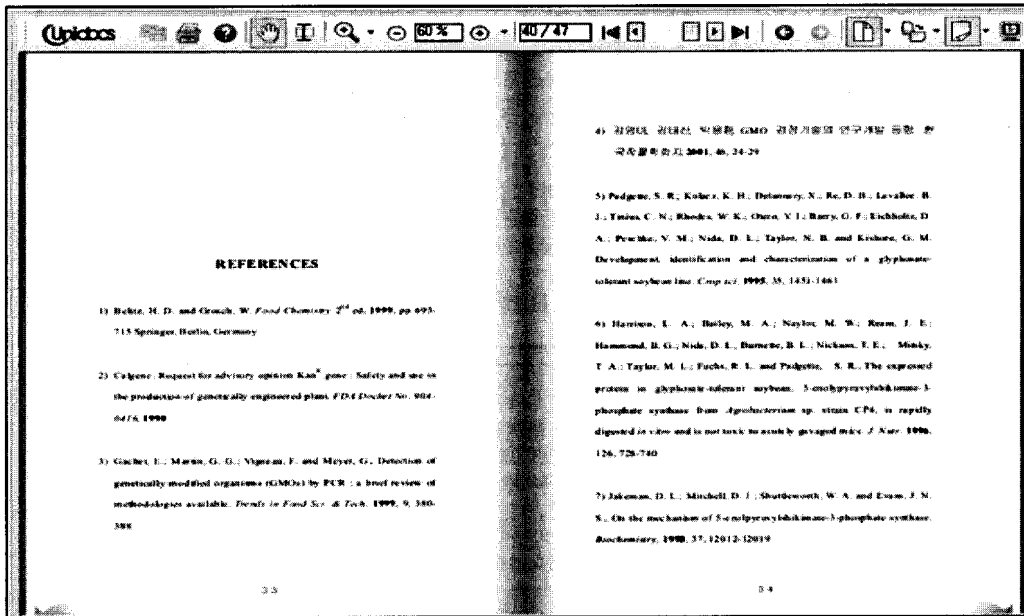
다음의 상세화면 항목은 메타데이터인 MARC 데이터를 기반으로 하여 추출된 정보이다.

| | |
|-----|--|
| LDR | 01030cam ak200229 k 4500 |
| 001 | 0001198913▲ |
| 005 | 20060710110648▲ |
| 008 | 050311s2005 ulka AK 000a kor ▲ |
| 040 | ▼a211032▲ |
| 041 | 0 ▼akor▼beng▲ |
| 082 | ▼a658▼220▲ |
| 090 | ▼a658▼b2005▼cG154j▲ |
| 100 | 1 ▼a강버들.▼d1974-▲ |
| 245 | 10 ▼a지식서비스산업에서 서비스품질평가의 상대적 차이 :▼b컨설팅 서비스의 고객군별 차이를 중심으로 = ▼xRelative difference in service quality in the knowledge-based industry : an empirical study of consulting service /▼d강버들.▲ |
| 260 | ▼a서울 :▼b서울대학교 대학원.▼c2005.▲ |
| 300 | ▼avi, 77 p. :▼b삽도 :▼c26 cm.▲ |
| 502 | 0 ▼a학위논문(석사)--▼b서울대학교 대학원 :▼c경영학과 경영학전공.▼d2005.▼e미유재.▼5AK▲ |
| 653 | ▼a서비스 품질▼aSERVQUAL▼a서비스 로열티▼a컨설팅 서비스▼a신뢰▼a지식기반 산업▼a지식서비스 산 업▼a전문 서비스 기업(PSF)▼a상호작용품질▼a결과 품질▼a환경 품질▼a산업제 고객군▼a소비제 고객군▲ |
| 653 | ▼aService Quality▼aSERVQUAL▼aInteraction Quality▼aResult Quality▼aEnvironmental Quality▼aTrust ▼aConsulting Service▼aProfessional Service Firms▼aService Royalty▼aKnowledge Service Industry▲ |
| 709 | ▼a미유재.▼d1960-▲ |
| 995 | ▼a200220672▼b2005.2▲ |

[그림 10] 학위논문 MARC 데이터 사례

위의 MARC에 대하여 설명한 바와 같이 각 태그에 해당하는 서브필드의 요소들이 상세히 기술되어 있고, 이를 기반으로 상세정보 메타데이터 요소가 [표 8]과 같이 추출되었다.

메타데이터 요소를 기초로 각 항목별 세부요소들을 추출하여 토픽맵을 기반으로 한 서비스의 구성 요소로 활용한다. 여기에 연구자 정보에 관련한 요소들은 기존에 입력된 전거레코드를 바탕으로 추가 구축하여 활용 서비스할 대상이 된다.



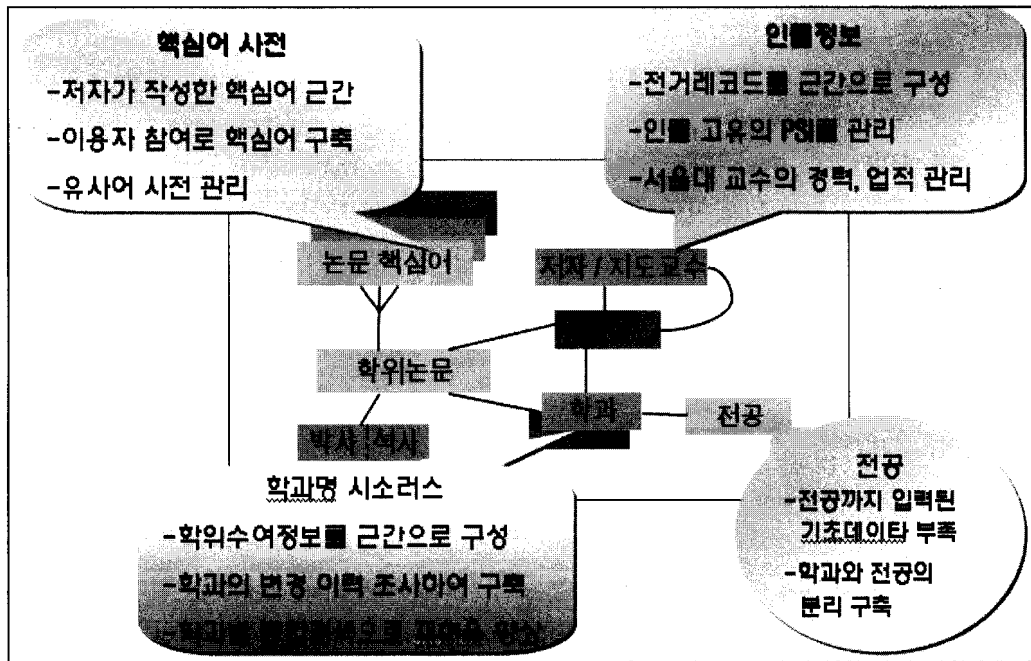
[그림 12] 학위논문 원문서비스

현재 제공되는 2차 정보에서 더 확대하여 참고문헌을 활용하면 유익한 정보가 될 수 있다. 참고문헌 데이터를 원문열람이 아니고도 별도 항목으로 열람하도록 하며, 텍스트로 구축하여 서지사항에서 바로 재검색으로 서지정보간의 연결 관계를 생성하여 서비스의 범위를 확대하면 이용자는 한가지의 검색에서 관련된 많은 논문이나 학술문헌 정보까지 이용할 수 있게 된다.

4.2 토픽맵 기반 학위논문 서비스 방법론 적용

4.2.1 메타데이터들의 상호연관성 분석을 통한 관계성 정의

앞 장에서 추출한 메타데이터 요소들의 상호 연관성을 분석해 보면서 각 주제간의 관계를 표현해 보면 다음과 같다. 우선 각 요소들 가운데 학위논문 서비스의 대표적인 요소를 학위논문 서지정보와 저자, 지도교수, 학위수여정보, 학과정보, 논문핵심어 등을 추출하여 각 주제간의 관계성을 살펴보도록 한다.



[그림 13] 학위논문 각 주제간의 관계성

위의 관계성은 하나의 학위논문에서 형성되는 관계도를 나타내고 있다.

1) 학위논문 / 학과 정보

현재 기구축되어 있는 MARC기반의 메타데이터에는 502tag 학위수여 정보에 석박사구분, 학위수여기관, 학과, 전공, 학위수여년도, 지도교수 등을 기술하고 있지만, 목록기술규칙에 의거해 인준지(또는 표제지)에 있는 그대로 기술하기 때문에 학과정보만 있을 뿐 상세한 전공까지 기술되지 않은 경우가 많으므로 전공과 학과를 계층구조로 관계성을 구성하기에 어려움이 있다. 학과와 전공은 하나의 학과에서 다수의 세부 전공분야가 나올 수 있으므로 개체간의 계층구조가 형성된다. 또한 서울대학교 뿐 아니라 대부분의 대학은 학과의 명칭이 시대의 흐름에 따라 계속적으로 변화해 오고 있다. 2000년 이후 학부제 도입으로 소속과 명칭에 대한 범주도 많이 바뀌었다. 따라서 이용자가 현재의 학과명칭으로 검색했을 때, 이전 학위논문 자료를 찾을 수가 없는 문제점을 가진다. 따라서 서비스 이용효과를 증대하기 위해서는 현재의 학과명칭을 기준으로 계속적으로 변경되어온 history 내역을 연혁 자료에서 찾는 등의 방법으로 학과시

소러스 구축이 필요한 실정이다.

2) 논문 핵심어 정보

논문핵심어 또한 하나의 논문에서 저자가 핵심어를 생성하므로 다수가 나올 수 있으며, 논문핵심어의 메타데이터 성격은 주제어라고 하기 보다는 저자가 생성하는 비통제 키워드이므로 주제어의 통일성을 갖추지 못한 폭사노미(Folksonomy)⁴⁾라 할 수 있다. 따라서 논문 핵심어에 대한 시소러스 또는 유사어 사전을 구축하여 검색의 재현율을 높일 필요가 있다.

3) 인물정보

저자의 경우는 학위논문을 작성한 저자로서의 역할이 이후에는 새로운 학위논문 저자의 지도교수로서의 역할을 할 수 있다. 즉, 동일 인물이 서울대학교 내에서 논문의 저자로서, 지도교수로서의 역할을 가질 수 있는 것이다. 서울대학교에서는 인물정보에 대한 전거DB가 구축되어 있으므로, 동일인물에 대해서는 통합관리가 이루어지고 있으며, 학위논문 서비스에서 인물정보에 대한 필수적인 사항이다. 이번 온톨로지 기반 학위논문 서비스에서는 서울대학교 교수와 학위논문 저자에 대한 인물정보 메타데이터를 좀 더 상세히 확보하고, 인물에 대한 전거통제를 이용하여 PSI를 관리하며, 저자, 지도교수, 제자 등의 역할관계를 명시함으로써 연구자협업 서비스를 위한 기반을 마련하도록 한다.

4.2.2 학위논문 서비스 토픽맵 디자인 및 정제

1) 목적과 범위

온톨로지 토픽맵 모델링의 방법론에서 초기단계는 대상 도메인의 목적과 범위

4) 폭사노미(Folksonomy) : folk(people) + order + nomos(law)의 합성어로 한글로 풀어쓴다면 "사람들에 의한 분류법" 정도로 해석할 수 있습니다. 폭사노미는 태그(tag)를 사용하여 공동으로 분류하는 방식을 말하는 신조어이다. 이 방식은 소셜 소프트웨어가 변화하며 나타나기 시작했다. 최근 온라인 folksonomies의 가장 좋은 예는 del.icio.us 같은 소셜 북마크 사이트, 북마크 공유 사이트, 사진을 공유하는 Flickr과 목표(Goal)을 공유하는 43 Things 이다. Gmail의 라벨 시스템도 (사용자들이 분류한 것을 공유할 수 없기 때문에 Folksonomy는 아니지만) tag의 사용이라는 면에서 다소 비슷하다. Folksonomy는 - 도서관(library science)에서의 분류부터 장난스런 분류까지 - 상호관계가 명확한 것은 아니다. (wikipedia 정의)

를 파악하고 요구사항 및 기존 시스템의 분석단계를 거친다. 분석을 통하여 주요 개념들을 추출하고 정제, 범위를 정의한다.

간략하게 대상 도메인 학위논문 서비스에 대한 목적, 방법 및 범위를 정리하면 다음과 같다.

가) 목적과 방법

- ① 단계별 구축에 따른 확장가능성 확보와 지식 구조(Knowledge Structure)를 체계적으로 기술하고, 서울대학교에 다양한 형태로 존재하는 실제 정보자원들과 지식을 연결하여 궁극적인 지식통합(Knowledge Integration)을 구현하기 위한 국제표준 적용 (ISO13250)
- ② 효과적인 지식자원의 조직, 필요자원의 정확한 추출, Navigation 및 XTM(XML Topic Map) Schema 구조 적용
- ③ 하나의 주제에 관련된 다양한 형태의 관련 지식을 자유롭게 연계시킬 수 있는 효율적인 방법 제시

나) 서비스 범위

- ① 토픽맵을 적용하여 학위논문의 저자, 지도교수, 학과명, 핵심어 등을 토픽화하여 다양한 접근성과 연관성을 제시한다.
- ② 한번의 검색과 결과상세화면에서 논문연구자와 연계된 모든 정보와 논문정보의 접근경로를 줄여준다.
- ③ 검색서비스 활성화를 위해 논문명, 저자명의 검색 서비스뿐 아니라, 저자가 추출한 논문 핵심어, 서울대학교의 학과명, 전공을 연계하여 접근점을 높여주고, 검색결과에서도 한 화면에서 논문정보뿐 아니라, 연구자의 연구분야, 인물정보까지 열람함으로써 서비스 효과를 높여준다.
- ④ 사용자 참여를 유도하는 핵심어 사전 구축으로 특정 핵심어에 대한 유사어, 관련어 검색으로 확대한다.
- ⑤ 논문의 참고문헌을 텍스트화하여 서지사항의 조회 및 재검색, 저자의 저작을 외부 웹정보원을 통한 검색 등 참고문헌을 활용하여 서비스를 확대한다.

단, 논문의 참고문헌 활용서비스는 텍스트화하여 구축하는데 많은 시간과 비용이 필

요하다. 예를 들면, 참고문헌 리스트 작성의 통일성문제, 기존의 PDF화일에서 참고문헌을 OCR로 텍스트화 할 경우의 정확성이나 인식문제, DB구축을 위한 인력과 비용 등 고려할 사항이 많으므로 1단계의 기초 DB 구축 범위에서는 제외한다.

2) 분석

기존시스템의 분석을 통하여 활용가능한 메타데이터 요소들을 파악해 보았고, 요구사항 분석으로 학위논문 저자에 대한 인물정보 구축에 필요한 메타데이터 요소도 살펴 보았다. 분석과 브레인스토밍을 통한 Topic type, Occurrence type, Association type 등을 정리하면 다음과 같다.

[표 9] 학위논문 서비스 개념의 분석

| | |
|------------------|--|
| Topic type | 학위논문, 핵심어 인물, 학과, 전공 |
| Occurrence type | 학위수여기관, 발행기관, 석사/박사, 형태사항, 초록, 원문, 저작물, 형식, 소장여부, 소장기관, 청구기호, 참고문헌, 게재지, 보고서형식, 주관기관, 지적재산권 정보, 대등서명, 이메일, 사진, URL, PSI, 경력, 소속, 학회, LAB |
| Association type | 저자, 지도교수, 스승/제자, 저자소속, 전공, 참여연구자, 게재, 인용, 세부전공, 논문핵심어 |

학위논문 서비스에서 타입간의 계층구조는 학과와 전공간의 계층구조가 표현될 수 있다. 단과대학과 학과간의 수직적 계층구조를 지니며, 학과와 학위논문에 기술된 세부전공에 대해서도 수직적 계층구조를 지닐 수 있다. 현재까지의 학과명칭의 변경이력을 시소러스 구축하는데 있어서도 시소러스의 다양한 관계모형, 즉 동의관계, 계층관계, 종속관계의 표현을 도입하기엔 변경이력에 대한 조사나 개념이 적용되기 어렵고, 대개는 개정된 학과명칭간의 시간적인 전후명칭에 대하여 수평적 관계구조를 기술하도록 한다.

3) 주요개념 추출 및 정제

데이터베이스의 기본 관련성을 표현할 때 UML⁵⁾형식을 많이 사용한다. 이 UML을

5) UML(Unified Modeling Language) : 요구 분석, 시스템 설계, 시스템 구현 등의 과정에서 사용되

기반으로 토픽맵을 위한 ORM(Object Role Modeling)을 도입하고자 한다. ORM은 개념레벨에서 데이터베이스를 설계하기 위한 방법론을 말한다. 데이터베이스 모델링에서 가장 어려운 점은 업무를 잘 이해하는 고객이 기술적인 면에서는 문외한이라는 것이다. ORM은 고객과 모델러 사이에 원활한 의사소통을 할 수 있도록 설계되었다. 그러므로 토픽맵의 개념레벨을 ORM diagram 형식을 도입하여 표현함으로써 모델링의 기초로도 활용하며 더불어 토픽맵의 주요 요소인 Topic, Occurrence, Association을 클래스단위에서 적용하여 타입간의 관계도를 비주얼하게 도식화할 수 있는 방법이라 할 수 있다. 아래의 그림은 토픽맵을 이용한 학위논문 서비스를 모델링의 기본 설계가 되는 Topic type, Occurrence Type, Association Type 단위에서 ORM을 이용하여 도식으로 구성해 본 것이다.

Topic type은 입체원으로 표현하여 학위논문, 핵심어, 학과명, 인명을 정의하고, 여기에 전공도 확대하면서 나올 수 있을 것이다. Occurrence Type은 타원으로 표현하고 각 Topic type에서 사용되는 정보들로 그룹화하여 정의한다. Topic Type과 Occurrence 간의 관계에서도 주로는 일대일의 관계이지만, 일대다의 관계인 경우를 Occurrence Type의 타원 앞에 점을 표기하여 다수의 값을 가지는 경우를 표현한다. 예로 인명의 경우, 서울대학교 전거레코드에서 400tag에 보라참조 사항으로 저자의 영문표기, 한자표기, 다른 이명표기를 상세히 기술하고 있으며, 이를 변환하여 토픽맵을 구현할 때 한 저자에 대한 다른 이명표기값이 다수가 올 수 있다. 이 값들을 토픽맵에서는 “이명표기”라는 occurrence로 모두 기술해 준다.

Association은 Topic Type간의 연관관계를 정의하는데, 대표관계 명칭을 정의하는 Association Name과 Topic type 간의 쌍방향 관계를 알려주는 label명칭을 요구한다. 이를 두 개의 사각형으로 연결하고 쌍방향 화살표로 표현하였으며, 사각형위에 대표하는 Association name을 기술하였다. 예를 들어 학위논문과 인물 간에는 대표적 역할로서 Association name은 “논문저자”이고, 학위논문은 인물의 “저작”이면서, 인물은 학

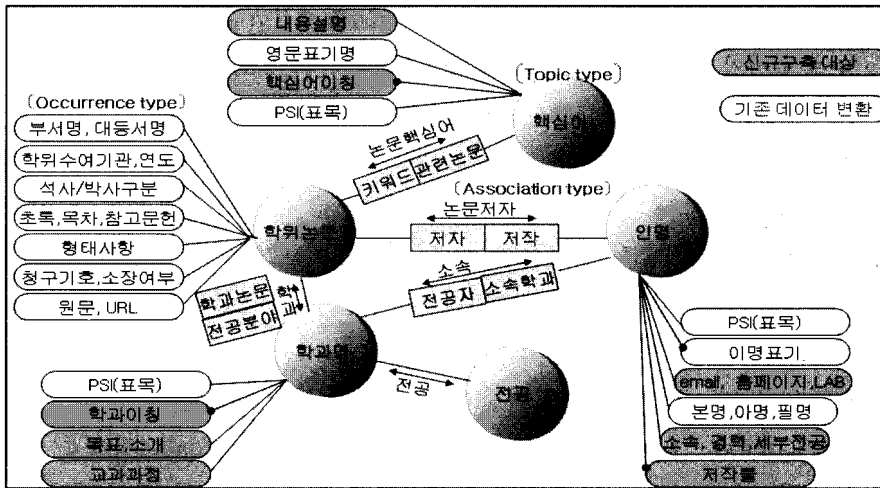
는 모델링 언어

객체 관련 표준화기구인 OMG에서 1997년 11월 객체 모델링 기술(OMT; object modeling technique), OOSE 방법론 등을 연합하여 만든 통합 모델링 언어로 객체 지향적 분석·설계 방법론의 표준 지정을 목표로 하고 있다. 요구 분석, 시스템 설계, 시스템 구현 등의 과정에서 생길 수 있는 개발자간의 의사 소통의 불일치를 해소할 수 있다. 모델링에 대한 표현력이 강하고 비교적 모순이 적은 논리적인 표기법(notation)을 가진 언어라는 장점이 있다. 따라서 개발자간의 의사 소통이 쉬워지며 생략되거나 불일치되는 모델링 구조에 대한 지적도 용이하고, 개발하려는 시스템 규모에 상관없이 모두 적용 가능하다.

위논문의 “저자”관계가 생성되는 것이다.

Role Type은 주로 Topic Type과 동시에 사용하면서 Association을 나타낼 때 역할을 정의할 수도 있지만, Role Type만을 부여하여 Topic Type간의 새로운 관계를 정의할 수도 있고, 하나의 Topic Type에서도 회귀성 관계가 생성될 때에도 이 정의를 위하여 사용할 수 있다. 학위논문서비스에서는 지도교수가 대표적인 예이다. 지도교수는 인명이라는 Topic Type에 포함되는 인물정보이면서, 논문저자의 지도교수의 역할을 가지며, 동시에 논문을 작성한 저자의 역할을 가질 수도 있다. 이 경우에 Role Type에 “지도교수”를 정의하고 인물 Topic 사이에 Association을 형성할 때 지도교수로 기술한다. 토픽맵의 장점인 주제지식(topic) 간의 다양한 연관관계의 표현이 자유로운 점이 나타날 수 있는 사례로 생각된다.

[그림 14]를 바탕으로 하여 상세한 토픽맵 디자인을 구현해 보고자 한다.



[그림 14] Topic Map ORM(Object Role Modeling)

4) PSI 부여정책

PSI는 Published Subject Identifier의 약자로 단일 주제를 식별하기 위한 것으로 그 주제의 식별자로 취급하기 위해 만든 것으로 공적주제식별자라고 명명한다. 공적주제식별자 PSI는 사람이 그 주제에 대해서 분명하게 이해할 수 있는 해석 가능한 메타데이터가 포함된 문서나 레코드로 안내를 해야 한다. PSI는 정의되고 배포된 식별자 역할을 목적으로 주제를 나타내거나 설명하는 문서를 변환하기 위한 URL이다. PSI의 주요 장점은 개념을 설명하기 쉽다는 것이다. 또한 PSI는 전

적으로 새로운 패러다임을 만들기보다는 현재 사용 중인 것을 확장한다. URL은 이미 문서를 위한 식별자로 널리 이해되고 있으며 오랫동안 많은 커뮤니티들은 단일 주제를 위한 식별자로 이를 사용해 왔다.

기존의 글로벌 식별자 체계를 갖지 않는 시스템에서는 인터넷 도메인과 주제 도메인 요소를 기반으로 한 네임스페이스 URL을 가지고 알파벳이나 숫자 코드를 전치시켜서 PSI로 변환한다.

저자와 지도교수의 인물정보 DB 구축과 관련하여 인물정보에 대한 각 토픽에 대한 PSI를 부여한다. 또한 서울대학교 내의 학과명칭이 계속적인 변화가 있어왔으며, 이에 대한 시소러스 구축이 필요하다. 현 단계 학위논문 서비스를 위한 학과명 시소러스 구축시에도 PSI를 부여하여 식별체계를 마련하고 이후 상호 연관성있게 이용될 수 있도록 한다.

Topic Type, Occurrence Type, Association Type을 정의할 때, 일부 개체는 더블린 코어(DC)의 구성요소와 매핑되는 사항이 존재한다. 이 부분은 DC의 구성요소를 사용하여 표준에 가깝게 주제식별을 정의하도록 한다. DC의 구성요소와 서브구성요소인 dcterms에서 값을 도입하여 정의하는 PSI 항목들은 아래의 표와 같다.

[표 10] 토픽맵 요소의 DC 구성요소를 사용한 PSI 정의

| 유형 | 정의요소 | PSI |
|------------------|------------|---|
| Topic type | 논문명 | http://purl.org/dc/elements/1.1/title |
| | 논문주제 | http://purl.org/dc/elements/1.1/subject |
| Occurrence type | 대등서명 | http://purl.org/dc/terms/alternative |
| | 발행기관 | http://purl.org/dc/elements/1.1/publisher |
| | 형식 | http://purl.org/dc/elements/1.1/type |
| | 범위 | http://purl.org/dc/elements/1.1/coverage |
| | identifier | http://purl.org/dc/elements/1.1/identifier |
| | date | http://purl.org/dc/elements/1.1/date |
| | 목차 | http://purl.org/dc/terms/tableOfContents |
| | 자원형식 | http://purl.org/dc/elements/1.1/format |
| | 형태사항 | http://purl.org/dc/terms/extent |
| | 초록 | http://purl.org/dc/terms/abstract |
| Association type | 지도교수 | http://purl.org/dc/elements/1.1/contributor |

인물정보와 학과명 시소러스에 대한 PSI 부여정책에 관하여 기본적인 전제를 기술하면 다음과 같다.

가) PSI 관리체계

전체 PSI의 네임스페이스 부여정책은 국제적으로 가장 일관성 있게 네임스페이스를 관리해 온 더블린 코어의 정책을 기반으로 하고, 서울대학교의 필요와 특성에 따라 PSI 정책을 확장시켜 가기로 한다.

PSI를 부여하기 위한 웹 관리체계는 "<http://psi.기관명/terms/용어명>"의 형식을 취한다. 서울대학교에서 관리하는 용어들은 "<http://psi.snu.ac.kr/terms/용어명>"의 형태로 기술될 수 있다. 용어의 관리와 사용은 별개의 것이다. PSI는 어느 기관에서든지 활용할 수 있고 그 PSI를 관리하는 기관이 그 용어의 정의와 용례를 명확히 밝히고 관리의 책임을 지는 것이다.

학위논문 콘텐츠에서 작성되는 인스턴스에 PSI를 부과하는 것은 토픽맵의 "Subject Identifier"를 활용하거나 Occurrence Type에 PSI라는 유형을 만들어서 입력함으로써, 인물정보, 학과정보 등에 대한 PSI를 공적으로 사용할 수 있도록 한다.

인스턴스를 등록할 경우, 한글 고유명사의 경우는 대체로 한글로마자표기법을 따라 홍길동의 경우 HongGilDong과 같이 영문화하여 생성하지만, 본 사업의 경우에는 기존의 전거표목의 변환이라는 과제와 일관성을 유지하기 위하여 한글 표기를 그대로 기입하도록 한다.

나) 인물정보 PSI

학위논문 서비스의 주요 요소가 되는 저자, 지도교수에 대한 인물정보 Topic과 학과명 시소러스 Topic에 대해서는 인스턴스 레벨에서 PSI를 부여하여 관리한다. PSI를 부여하는데 있어서 인명과 학과, 주제와 같은 중요한 용어들은 하위 디렉토리를 생성하여 관리하는 것을 원칙으로 한다. 예를 들면, 저자 홍길동을 Topic으로 등록하면서 PSI를 관리하고자 할 경우, <http://psi.snu.ac.kr/person/홍길동>으로 '홍길동'을 등록한다. 하지만, 이름만으로는 동명이인이 너무나 많이 발생할 수 있으므로 유일한 값의 부여를 위하여 "이름 + 생몰년"으로 부여하기로 한다. 이름+생몰년에서도 동명이인이 발생할 수 있으며, 또한 개인정보 노출이라는 문제점을 안고 있긴 하지만, 현재의 서울대학교 전거레코드의 변환을 고려하여 PSI를 부여하고 이후, 더욱 이상적인 방법이 연구된다면 도입하기로 한다. 즉 <http://psi.snu.ac.kr/person/홍길동,1931-1997>로 부여한다. 만약, 출생연도가 같은 동명이인이 있을 경우에는 연월일까지를 기술하여 유일한 값을 갖도록 한다.

인물정보에 대한 PSI 부여 정책은 현재의 전거레코드 표목과 동일한 방식을 따라 기술하지만, 이후의 새로운 부여방식의 확대를 위하여 현재 서울대 전거레코드에서 가지고 있는 contents ID를 identifier로 값을 주어, 이후의 확장이나 변환에 대한 대비를 하도록 한다.

다) 학과명칭 PSI

논문작성 시 등록된 학과명칭은 변경, 통합, 폐지 등에 의하여 현재의 명칭과 상이하여 검색 효율이 떨어진다. 그러므로 현재의 학과 체계를 중심으로 이전의 변천과정을 topic에 구현하여 학과 및 전공 검색 시에 일관된 검색서비스가 활용될 수 있도록 학과명 시소러스를 구축한다. 서울대학교 내의 현재 학과 체계와 변천과정을 전거 또는 시소러스 구축하여 검색효율을 높이고 서비스를 강화한다. 학과의 최근 표기를 PSI로 부여하고, 계속적인 변천기록을 어커런스로 기록하면서 계층구조를 형성한다.

4.2.3 학위논문 서비스 토픽맵 디자인 구현

학위논문 서비스를 위한 토픽타입과 어커런스타입, 관계타입 등에 대하여 살펴 보았고, 서비스 모델링도 도식화하여 보았다. 이제 학위논문 온톨로지 구현방안을 서비스에 맞추어 구체적으로 설계해 보도록 한다.

○ Topic Maps Schema 설계를 위한 구현

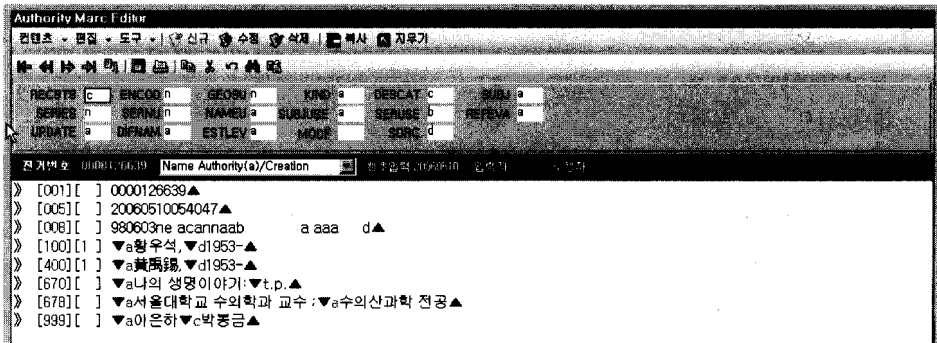
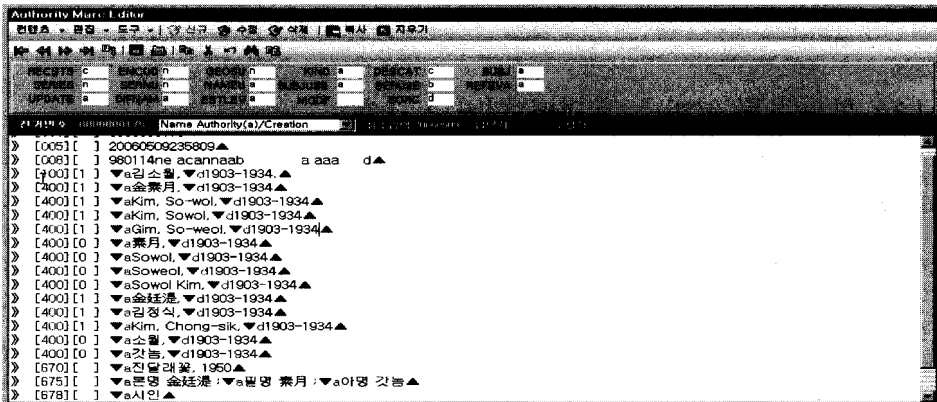
- 학위논문(토픽타입)

- . 토픽 : 기본 서명
- . 대등서명 / 부서명
- . 발행지, 발행기관, 학위수여기관
- . 발행연도, 학위수여연도
- . 학위구분, 전공
- . 형태사항, 페이지정보
- . 원문링크, 초록, 목차

- . 주기사항
- . 참고문헌

- 인물정보 (토픽타입)

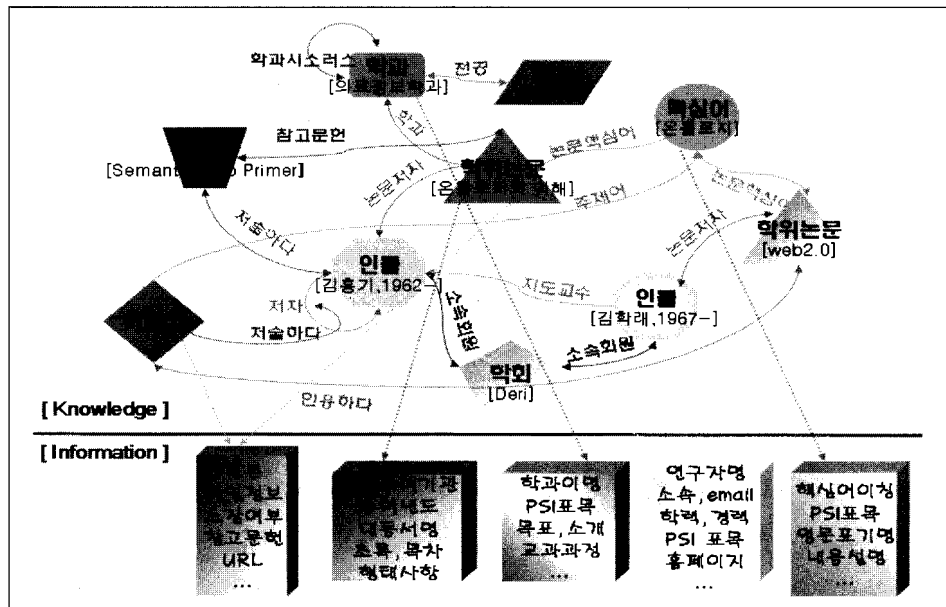
- . 토픽 : 이름+생몰년
- . 연관 : 지도교수, 학위논문의 저자, 참여활동
- . PSI 정보 (선거통제값)
- . 이메일
- . 홈페이지, LAB
- . 소속, 경력
- . 세부전공
- . 저자의 이명표기 - 본명, 아명, 필명 등
- . 저서



[그림 15] 인물정보에 참조되는 전거레코드 사례

- 학과명 시소러스
 - . 토픽 : 기본 학과명
 - . 상위토픽 : 단과대학
 - . 연관 : 이전 학과명, 이후 학과명
 - . 학과 영문표기명
 - . 설명
 - . 다른 학과명표기

- 핵심어
 - . 토픽 : 논문핵심어
 - . 연관 : 학위논문, 유사 핵심어
 - . 영문 표기명
 - . 설명
 - . 다른 표기명



[그림 16] 학위논문 서비스 토픽맵 모델링

각각의 Topic Type과 Occurrence Type, Association Type을 토픽맵 모델링으로 구현하여 지식층과 정보층으로 나누어 그 상관관계를 도식화하면 [그림 16]과 같다. Topic Type 아래에는 사례가 되는 대표적인 Topic 인스턴스를 기술하여 이해를 돕도록 한다.

이상에서 온톨로지 기반의 토픽맵을 활용한 학위논문서비스의 모델링을 제시하였다. 모델링은 시스템의 설계에 해당하는 밑그림이라 할 수 있다. 그러므로 학위논문 온톨로지 구현방안은 모델링을 기반으로 이에 필요한 인스턴스의 생성과 토픽맵 엔진을 이용한 시스템 구현으로 서비스화 하는 것이다. 인스턴스의 생성은 기존의 학위논문 MARC데이터와 전거DB, 원문구축DB 등을 활용하여 매핑변환을 하겠지만, 이와 더불어 모델링에서 기술한 인물정보 부분, 학과명 시소러스 부분, 논문핵심어 등의 주요 Topic Type은 온톨로지 기반의 서비스를 위하여 추가적인 DB구축이 요구된다. 이상적인 서비스는 훌륭한 데이터와 튼튼한 DB가 구축되어야 그 효과를 발휘할 수 있듯이 기존의 서울대학교 DB에 추가적인 기초DB를 구축함으로써 온톨로지 기반의 토픽맵 서비스를 위한 초석을 마련하는 것이다.

이러한 튼튼한 초석 위에서 2단계 사업으로 토픽맵 엔진을 도입하여 이용자 친화적이고 비주얼한 인터페이스와 이용자 참여가 활성화된 학위논문 온톨로지 시스템을 구현하는 것이다.

4.3 토픽맵 기반 학위논문 서비스를 위한 기초 DB 구축

4.3.1 토픽맵 기반 학위논문 서비스 기초 DB 구축방안

이 장의 DB구축 방안은 기본적으로 온톨로지 기반의 학위논문 서비스가 되기 위한 기초DB 구축부분에 대한 구축방안을 기술하고자 한다. 기초DB는 학위논문 지도교수 인물DB, 학과명 시소러스DB, 논문 핵심어사전 DB 영역에서 구축하고자 한다.

아래 [표 11]은 기구축된 서울대학교 학위논문을 대상으로 서울대학교 소속 전체 교수를 대상으로 학위논문 지도교수 등록 레코드 건수와 학위논문 서지레코드 내의 학위수여정보 내 등록된 학과명을 대상으로 학과간의 이력 관계를 간략 시소러스DB 구축 대상으로 한다. 논문 핵심어는 653tag에 논문 핵심어가 등록되지

많은 핵심어를 대상으로 서지레코드를 파악한 데이터 양이다.

[표 11] 서울대학교 기초DB 구축 대상

(단위 : 서지레코드 건수)

| 구분 | 신규 | 전거/시소러스/사전 | 비고 |
|------|--------|------------|------------------|
| 지도교수 | 67,836 | 2,800* | 709태그/ 502태그 \$e |
| 학과명 | 10,287 | 84,308 | 502태그 \$c |
| 핵심어 | 69,287 | 84,308 | 653태그 |

각각의 기초DB 구축 대상에 대한 구체적인 구축방안을 살펴보기로 한다.

1) 학위논문 DB구축 방안

학위논문 DB는 기존의 SOLARSIII 시스템에 구축된 DB를 그대로 활용한다. 학위논문 DB는 MARC 형식으로 저장되어 XML로 반출이 가능하다. MARC의 메타데이터 요소별로 데이터를 추출하여 토픽맵의 구조로 매핑, 인스턴스를 가져오도록 한다. 학위논문 서비스에서 학위논문은 신규 구축대상이 아닌 대부분의 요소는 학위논문 MARC와 그 외 관계형DB에서 XML형식으로 반출하여 토픽맵 기반 서비스에 활용할 수 있다. 학위논문, 인물정보, 학과명, 핵심어 등의 주요 Topic Type의 name 인스턴스와 학위논문의 Occurrence가 메타데이터 요소로 매핑된다.

[그림 17]은 MARCXML로 반출된 데이터필드 태그와 서브필드 코드의 값들을 나타내고 있다. [그림 18]은 토픽맵의 상용 편집기인 온토폴리(ontopoly)⁶⁾를 이용하여 Topic Type을 정의할 때, 해당하는 Occurrence Type과 Association Type의 요소를 가져와서 학위논문 Type의 정의를 완성하는 화면이다. [그림 17]의 MARCXML 추출 내용을 기반으로 [그림 18]의 토픽맵 구성화면을 설계해 보는

6) 온토폴리(Ontopoly)는 ontopia사에서 OKS(Ontology Knowledge Suit)에서 토픽맵을 생성할 수 있는 Tool을 제공한다. 이 OKS는 토픽맵 엔진과 토픽맵 생성 및 편집을 위한 온토폴리(Ontopoly), 그리고 토픽맵을 웹 상에서 보여주고, 시각화하여 나타내는 vizigator를 포함한 옴니게이트(Omnogator) 등으로 구성되어 있다. 여기서는 모델링과 기초DB구축방안을 위한 기본적인 토픽맵 생성과 샘플 인스턴스 등록을 위하여 편집기인 온토폴리를 사용하여 토픽맵을 구현하면서, 이해를 돕기위하여 온토폴리 화면을 바탕으로 설명하기로 한다.

예시를 나타내고 있다.

```

- <datafield tag="090" ind1="" ind2="">
  <subfield code="a">660</subfield>
  <subfield code="b">2006</subfield>
  <subfield code="c">H991c</subfield>
</datafield>
- <datafield tag="100" ind1="1" ind2="">
  <subfield code="a">황재웅,</subfield>
  <subfield code="d">1971-</subfield>
</datafield>
- <datafield tag="245" ind1="1" ind2="0">
  <subfield code="a">Characterization of biofilter packing material and
  development of novel biotrickling filters =</subfield>
  <subfield code="x">바이오필터용 담체 특성 연구 및 신개념 바이오트리클링 필터의 개
  발 /</subfield>
  <subfield code="d">黃載雄.</subfield>
</datafield>
- <datafield tag="260" ind1="" ind2="">
  <subfield code="a">서울 :</subfield>
  <subfield code="b">서울대학교 大學院,</subfield>
  <subfield code="c">2006.</subfield>
</datafield>
- <datafield tag="300" ind1="" ind2="">
  <subfield code="a">x, 141 leaves :</subfield>
  <subfield code="b">Ill. ;</subfield>
  <subfield code="c">26 cm.</subfield>
</datafield>
- <datafield tag="502" ind1="1" ind2="">
  <subfield code="a">Thesis(doctoral)--</subfield>
  <subfield code="b">서울대학교 대학원 :</subfield>
  <subfield code="c">응용화학부,</subfield>
  <subfield code="d">2006.</subfield>
  <subfield code="e">최차용.</subfield>
  <subfield code="5">AK.</subfield>
</datafield>

```

```

- <datafield tag="504" ind1="" ind2="">
  <subfield code="a">Includes bibliographical references.</subfield>
</datafield>
- <datafield tag="653" ind1="" ind2="">
  <subfield code="a">바이오 필터</subfield>
  <subfield code="a">교반형 바이오 트리클링 필터</subfield>
  <subfield code="a">회전식 드럼 바이오 트리클링 필터</subfield>
  <subfield code="a">퇴비</subfield>
  <subfield code="a">발포 폴리 우레탄.</subfield>
</datafield>
- <datafield tag="653" ind1="" ind2="">
  <subfield code="a">biofilter</subfield>
  <subfield code="a">agitating biotrickling filter</subfield>
  <subfield code="a">rotating drum biotrickling filter</subfield>
  <subfield code="a">compost</subfield>
  <subfield code="a">polyurethane.</subfield>
</datafield>
- <datafield tag="709" ind1="" ind2="">
  <subfield code="a">최차용,</subfield>
  <subfield code="d">1942-</subfield>
</datafield>
- <datafield tag="995" ind1="" ind2="">
  <subfield code="a">96402816</subfield>
  <subfield code="b">2006.2</subfield>
</datafield>
- <datafield tag="999" ind1="" ind2="">
  <subfield code="a">Unidocs</subfield>
  <subfield code="c">김선도</subfield>
</datafield>
</record>
</collection>

```

[그림 17] MARC 데이터의 XML 반출화면

Name: 학위논문

PSI: http://psi.snu.ac.kr/terms/thesis

Description:

Create instance

Delete this topic type

Delete

View instances of this type
>> 학위논문

Browse this topic in Omnigator
>> Omnigator

View this topic in Vizigator
>> Vizigate

Existing fields:

| Name | Name | Exactly one | Zero or one | Zero or more |
|--------|--------|--------------|-------------------------------------|--------------------------|
| 부서명 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 대들서명 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 석박사구분 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 학위수여기관 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 학위수여년도 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 현재사항 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 소장여부 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 참구번호 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 서지주기 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 초록 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 목록 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 참고문헌 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 원문URL | URI | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 원문화일정보 | String | Zero or one | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 논문저자 | 인물정보 | Zero or more | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 전공분야 | 학과 | Zero or more | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 핵심키워드 | 핵심어 | Zero or more | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 지도교수 | 인물정보 | Zero or more | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

Add fields:

Name:

Occurrence:
Creator

Association:
지도교수(스승)

Other:
Subject Identifier

[그림 18] 온토폴리의 학위논문 Topic Type 설정 화면

MARCXML의 각각 분리된 값들을 토픽맵에서 [표 12]와 같이 매핑하여 DB를 구축한다.

학위논문 DB는 기구축 되어 있는 분량에서 서지레코드의 수정 보완작업이 필요하다. 언어코드가 고정장에 잘못 들어간 데이터에 대하여 전체 서지데이터 중 오류를 추출한 후 수정작업을 한다. 전체 학위논문을 대상으로 학위 수여기관에 대한 오류 데이터도 추출하여 전거표목을 근거로 수정한다. 예를 들어, "서울대 대학원"으로 작성된 값은 "서울대학교 대학원" 으로 수정한다. 더불어 502태그의 학위수여정보 부분을 정비작업한다. 지시기호의 학사/석사 구분 값이나 \$2AK (서울대학교 학교코드 값) 값의 누락부분에 대한 정비작업이 포함된다.

[표 12] MARCXML to 토픽맵 매핑

| MARC XML 항목명 | 토픽맵 요소 |
|--|--|
| datafield tag=245 subfield code=a | topic type. 학위논문.name |
| datafield tag=245 subfield code=b | occurrence type. 부서명 |
| datafield tag=245 subfield code=x | occurrence type. 대등서명 |
| datafield tag=502 subfield code=a | occurrence type. 석박사 구분 |
| datafield tag=502 subfield code=b | occurrence type. 학위수여기관 |
| datafield tag=502 subfield code=d | occurrence type. 학위수여년도 |
| datafield tag=300 all | occurrence type. 형태사항 |
| datafield tag=504 all | occurrence type. 서지주기 |
| 관계DB | occurrence type. 소장여부 |
| datafield tag=090 all 과 관계DB | occurrence type. 청구기호 |
| 관계DB | occurrence type. 초록 |
| 관계DB | occurrence type. 목차 |
| 관계DB | occurrence type. 참고문헌 |
| 관계DB | occurrence type. 원문URL |
| 관계DB | occurrence type. 원문화일정보 |
| datafield tag=100 all | association type. 논문저자 topic type. 인물정보.name |
| datafield tag=502 subfield code=c | association type. 전공분야 topic type. 학과.name |
| datafield tag=650 subfield code=a | association type. 논문핵심어 topic type. 핵심어.name (한글) |
| datafield tag=502 subfield code=e datafield tag=709 all | association type. 지도교수 topic type. 인물정보.name |

2) 인물정보 DB구축 방안

인물정보의 DB구축은 서울대학교에서 구축한 전거레코드를 활용하여 인물정보 Topic과 매핑 구축하도록 한다. 전거레코드의 표목을 기준으로 하여 인물정보 Topic 의 Name으로 등록하며, 이 데이터 값을 인물정보 PSI로 관리한다. PSI는

앞서 기술한 부여정책에 따라 "<http://psi.snu.ac.kr/person/>전거표목값" 형식으로 등록한다. 전거레코드의 부출항목인 4XX, 5XX 태그 값들은 Occurrence Type의 이명표기로 매핑하고 670태그의 저자정보로 기술한 저작정보원은 Occurrence Type의 저작으로 매핑한다.

인물정보의 unique한 DB관리를 위해 동명이인 처리는 생몰년(생년월일), 전공 분야, 저작물 등의 정보를 활용하여 유지하도록 한다. 새롭게 부여되는 유일한 표 목관리는 PSI에서 관리하고, 효율적인 매핑 변환을 위하여 Topic Name 사항은 1XX 태그의 전거표목과 동일하게 등록한다.

```

<?xml version="1.0" encoding="utf-8" ?>
- <collection>
- <record>
  <leader>00304cz a2200121n 4500</leader>
  <controlfield tag="001">0000127117</controlfield>
  <controlfield tag="005">20060510054158</controlfield>
  <controlfield tag="008">980605ne acannaob a aae d</controlfield>
- <datafield tag="100" ind1="1" ind2="">
  <subfield code="a">최차용,</subfield>
  <subfield code="d">1942-</subfield>
</datafield>
- <datafield tag="400" ind1="1" ind2="">
  <subfield code="a">崔次鏞,</subfield>
  <subfield code="d">1942-</subfield>
</datafield>
- <datafield tag="670" ind1="" ind2="">
  <subfield code="a">서울대학교 공과대학 응용화학부 교수 ;</subfield>
  <subfield code="a">1942/11/09</subfield>
</datafield>
- <datafield tag="670" ind1="" ind2="">
  <subfield code="a">서울대학교 공업화학과 교수 ;</subfield>
  <subfield code="a">생물화학공학 전공</subfield>
</datafield>
- <datafield tag="999" ind1="" ind2="">
  <subfield code="a">김지숙</subfield>
  <subfield code="c">류양님</subfield>
</datafield>
</record>
</collection>

```

[그림 19] 저자, 지도교수 전거데이터 XML 반출 화면

학위논문 서비스의 인물정보 부분은 이후 온톨로지 기반 디지털 도서관 서비스의 초석이 되기 위해 신규로 추가할 사항을 상당히 포함한다. 즉 이를 기반으로 이후 사회적 협업 서비스를 구축, 활용할 수 있도록 구현한다. 그래서 인물정보 설계에 있는 메타데이터 요소인 전공분야, email, 홈페이지, LAB, 소속, 경력 등 다양한 정보들을 상세히 신규 DB구축 대상에 포함하여 구성한다.

학위논문 서비스의 현재 인물정보는 위의 그림과 같이 상당부분의 요소들에 대한 정보가 부족하다. 하지만 서울대학교 학위논문의 대상은 서울대학교 교수들로 구축범위가 제한될 수 있으므로 서울대학교 내의 교수연구업적관리시스템에서 교수이력사항을 참조DB로 사용할 수 있으며, 그 밖에도 한국한술진흥재단의 연구자

정보서비스를 활용하여 메타데이터 내용들을 보완, 구축할 수 있다.

인물정보 DB구축 입력은 토픽맵 입력틀을 이용하는 것이 아니라, 기존의 전거 레코드에 신규 서비스하는 토픽맵의 Occurrence Type에 대한 요소들을 전거 MARC 데이터 내에 로컬 태그로 추가정의해서 입력한다. 학위논문 MARC의 매핑 변환에 의한 서비스와 마찬가지로 전거MARC도 동일한 방식으로 SOLARSIII를 활용하면서 원활한 매핑과 서비스가 이루어질 수 있도록 한다. 전거MARC의 로컬 태그 사용 정의는 67X, 9XX 태그 등에 정의하여 입력한다.

| | |
|-----------|---|
| Name: | 최재웅, 1942 |
| 인물정보 PSI: | http://psi.snu.ac.kr/person/%5C(%A9,%201942 |
| 이명표기: | 崔次鏞, 1942 |
| 소속: | 서울대학교 공과대학 생물화학부 교수 |
| 저작물: | |
| 분류: | |
| 아명: | |
| 필명: | |
| 관심분야: | 생물화학공학 전공 |
| email: | |
| 홈페이지: | |
| 경력: | 서울대학교 공업화학과 교수 ; 생물화학공학 전공 |
| LAB명: | |
| LAB URL: | |
| 논문저자: | -- <none> -- |
| 소속학과: | 화학생명공학부 |
| 지도교수(스승): | -- <none> -- |
| 제자: | 최재웅, 1971 |
| 지도학위논문: | 바이오필터용 입체 촉매 연구 및 신개념 바이오리콜링 필터의 개발 |

[그림 20] 인물정보를 매핑하여 등록한 예

추가상세 기술에서 필요한 정보는 다음과 같다. 현재 678tag의 Note에 입력전 사항은 내용에 따른 분리기술이 필요한 바, 로컬 태그를 확장하여 필요 항목을 서브필드로 나누어서 분리 구축한다.

[표 13] 선거레코드의 로컬 태그의 정의 제시 사례

| | 원제 | 수정인 | |
|--|---------------------------|----------------------------------|---------------------------|
| | 로컬태그 (local tag) 확장 | 678tag \$a를 반복하여 모두 통합 기술함 | 필수요소 |
| | | 부가요소 | LAB 본명 참여학회 관심분야 |

가) 지도교수 입력

- ① 지도교수가 없는 서지를 대상으로 원문을 확인하여 지도교수를 입력한다.
- ② 709태그 및 502태그 \$e 사항이 없는 서지레코드를 대상으로 신규로 태그를 추가하며, 502태그 \$e에 입력된 사항에 대해 선거 유무를 확인한 후 선거DB 구축 대상을 확인하여 작업을 진행한다.

나) 선거 DB 구축

- ① 서울대학교의 SOLARSIII 프로그램을 이용하여 학위논문 데이터와 연동한 선거데이터를 직접 입력한다.
- ② 서울대학교 중앙도서관에서 추출한 지도교수 추출 리스트를 대상으로 작업 하며, 데이터의 오류나 누락자료를 수정, 보완한다.
- ③ 서울대학교의 지도교수에 대하여 선거의 기 구축 유무를 확인(F2)하며, 502 태그와 709태그 입력사항을 모두 확인하도록 하며, 해당 태그 및 필드가 누락되었을 경우 모두 보완하도록 한다.
- ④ 기 구축인 경우, 생몰년 등 필요한 정보를 보완하여 선거표목과 PSI값을 결

정하고, 구축된 전거 데이터를 보완 (생물년, 학과, 전공, 이메일주소, 저작사항, 홈페이지, 직번, LAB, 관련학회 등)

- ⑤ 기존 전거 DB가 없는 신규 구축인 경우, 지도교수 인명에 생물년 등 필수 정보로 전거표목을 구성하고, 학과, 전공, 이메일주소, 저작사항, 홈페이지, 직번, LAB, 관련학회 등 메타데이터 요소를 입력하여 구축한다.
- ⑥ 지도교수 전거데이터의 수정 및 보완작업에 필요한 참고정보원을 참조하며 작성하여, 서울대학교 내의 정보원 이외의 다른 정보원이 있는 경우 반드시 출처정보를 명시하도록 한다.

[표 14] 전거 작성시 필수 참고정보원

| 번호 | 참고정보원 | 비고 |
|----|-------------------|-----------|
| 1 | 학과 대표 홈페이지 | |
| 2 | 교수 홈페이지 | |
| 3 | 교수연구업적물관리시스템 인적정보 | 연구처 관리시스템 |
| 4 | 한국학술진흥재단 통합연구자정보 | 출처 명시 |
| 5 | JOINS 인물정보 | 출처 명시 |
| 6 | 서울대학교대표 홈페이지 | |

- ⑦ 전거레코드 구축시 서울대가 지정하여 사용하는 태그(679, 900 등)를 확장하여 사용하며 각 필드항목에 맞도록 해당값을 입력하도록 한다.
- ⑧ 서울대학교 중앙도서관에서 제공하는 전거구축 대상 리스트(엑셀 등)에 해당값을 수정하여 확인한 후, SOLARSIII 프로그램의 서지 데이터를 수정하며, 작업자의 1차 교열과 총괄자의 2차 교열이 완료를 거쳐 전거DB 구축을 재확인하여야 한다.

[표 14]는 전거DB 구축 시 사용할 필수 요소 및 메타데이터 요소에 대한 전거레코드의 로컬 태그를 정의한 예이다. 기초 DB구축 시에는 이후의 온톨로지 기반 서비스에서의 자동변환을 고려하여 매핑이 가능하도록 해당 태그와 서브필드를 명확하게 구분하고 나누어서 정의하고 입력하도록 한다. 그러면, 이후 토픽 맵 변

환시 Occurrence 값으로나 Topic 간의 Association 설정에서도 용이한 사용을 할 수 있을 것이다.

[표 15] 전거레코드의 로컬 태그 사용의 사례 제시

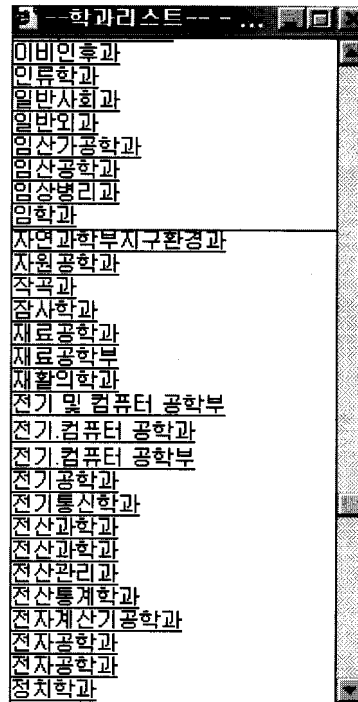
| 입력태그 | 서브필드 | 입력항목 | 비고 |
|--------|--------------------|-------------------------|--|
| 679tag | \$a | 성명+생물년 | 서브필드와 입력항목 추가 및 확장될 수 있으며, 각 필드는 반복 가능 |
| | \$b | 소속 | |
| | \$c | 학과 | |
| | \$d | 세부전공 | |
| | \$e | 직번(교번) | |
| | \$f | 이메일 | |
| | \$g | 관련학회 | |
| | \$h | 홈페이지 | |
| | \$i | LAB | |
| | \$j | 저작물(단행본) | |
| | \$k | 다른 저작물 (학위논문) | |
| \$l | 다른 저작물 (발표논문 등) | | |
| 900tag | \$a | 저자명 이명 표기 (영문 이니셜 등) | 반복가능 전거400태그 이외의 저자명 표기사항 입력 (<u>발표논문 등에 기재되는 인명 표기사항은 모두 입력함</u>) |

3) 학과명 시소러스 DB구축 방안

학위논문은 현재 학과를 주제구분하고 있기 때문에 학과를 검색하는데 있어 중요한 요소이므로 이를 정제해야 한다. 하지만, 대학의 학과는 시대의 흐름과 정책 등을 반영하기 때문에 그 명칭이나 교과과정, 내용 등은 계속 변경되어 오고 있

다. 같은 학과에서 나온 논문들도 학위수여년도에 따라서 그 전공이나 학과의 내용이 달라질 수 있다. 즉 기존의 학과명들이 이칭으로도 많이 있기 때문에 이번에 이것을 정비하고 그룹핑하여 이용자에게 효과적으로 제공해야 한다. 검색시에는 해당 학과명을 입력했을 시 학과의 이칭도 검색하시겠습니까? 하는 메시지로 보완하여 이용자가 원하는 경우 더 많은 정보로 접근할 수 있도록 유도되어야 한다.

하지만, 현재 운영하고 있는 학위논문 학과명 검색 시스템은 각 학과마다 변경된 학과명 이칭을 모두 관리하거나 기록되어 있는 자료나 데이터베이스는 어디에서도 찾을 수 없는 실정이다. 학위논문 검색에서 전공학과리스트를 열람하도록 하여 다양한 이칭표기까지도 선택하여 검색할 수 있도록 하고 있다.



[그림 21] 전공학과리스트 열람 화면

[그림 21]에서도 알 수 있듯이 학위논문 검색 서비스에서 서지레코드에 입력된 502태그 학위수여정보의 \$c전공 부분을 추출하여 가나다순 열람하는 방식이므로, 전공학과 리스트가 가나다 순으로 나타나고 있다. 여기서 “전기 및 컴퓨터 공학부”, “전기.컴퓨터 공학과”, “전기.컴퓨터 공학부” 등이 같은 학과가 계속 변경되

어 왔음을 알 수 있다. 그러나 중간에 “컴퓨터공학과“로 변경된 이력사항이 있는 경우에는 이와 같은 배열순서에서는 함께 열거되지 못하는 경우가 발생한다. 그러므로 현재의 학과리스트 열람만으로 이력을 모두 확인할 방법이 없다. 다양하게 변경된 학과명 히스토리 내역을 찾아 학과명 정보 시소러스로 구축함으로써 이용자의 주제별 접근 및 전공검색의 효율을 높이도록 한다. 시소러스 구축이라 하면, 용어 간의 상위개념, 하위개념, 동의어, 관계어, 대체어 등의 관계를 다양하게 구축하는 시스템을 말하지만, 본 연구과제의 학과명 시소러스는 서울대학교 내의 학과명에 대한 이력사항과 관계어에 대한 시소러스 구축으로 한정한다. 구축방법은 다음과 같다.

가) 학과명 시소러스 구축절차

- ① 학위논문 대상리스트에서 관련학과 및 세부전공을 그룹핑(클러스터링)하며, 이때 학위수여연도 또는 분류번호 등을 기준으로 하여 관련학과를 그룹핑하도록 한다.
- ② 그룹핑된 학과를 see also의 관계로 연결한다.
- ③ 학과명의 이력조사를 위해 학과 대표 홈페이지 및 학과 사무실, 참고정보원 (예: ‘전공과정 수강편람’ 등)을 활용하여 자료를 수집하고, 데이터상의 학과명과 일치 여부를 확인 후 학과의 이력관리 조사리스트를 모두 작성한다.
- ④ 조사한 리스트를 대상으로 학과명 시소러스를 작성하도록 하며, SOLARSIII의 CMS시스템에 업로드하여 학과명 시소러스를 생성하도록 한다.
- ⑤ 향후 입력되는 학과에 대하여 CMS시스템 내에서 see also관계를 연결 지을 수 있도록 하여 지속적인 관리가 가능하도록 한다.

[표 16] 학과명 이력조사를 위한 참고정보원

| 번호 | 참고정보원 | 비고 |
|----|--------------------------|--------------|
| 1 | 학과 대표 홈페이지 | |
| 2 | 교수 홈페이지 | |
| 3 | 서울대학교 대표 홈페이지 | |
| 4 | 전공과정 수강편람 | |
| 5 | 각 단대 및 학과별 사무실 | 전화 및 방문조사 실시 |
| 6 | 서울대학교 30, 40년, 50년사 발간자료 | |
| 7 | 교무과, 학적과, 입학본부 등의 관련 자료 | |
| 8 | 서울대학교 발행 백서 및 연혁집 등 | |

나) 학과명 시소러스 토픽맵 구축방안

학위논문 토픽맵 서비스 구현 시에는 온토폴리 입력틀에서 직접 학과명 Topic Type을 설계하고 구축하는 방법도 사용할 수 있다. [그림 22]와 같이 학과 Topic Type을 설정하여 필요한 Occurrence Type을 정의한다.

학과에 대한 Topic Type은 설정화면에서 보는 바와 같이 어커런스 타입 중에 표목PSI를 정의하고 그 값은 "<http://psi.snu.ac.kr/department/>" 하위디렉토리에 대표학과명을 등록하고 Topic Type.학과.Name 에 대표학과명을 기록한다. 각 학과마다의 홈페이지가 있으므로 이를 참조하여 학과의 목표, 소개, 교과과정, 홈페이지URL 등을 어커런스 타입으로 정의한다. 학과이칭표기에 변경이전의 학과명을 반복하여 기술하도록 한다. 혹시 학과명의 변경 이전, 이후의 명칭에 대한 순서가 중요하거나 히스토리를 필요로 한다면, 토픽맵의 Role Type을 이용하여 이전 학과명, 이후 학과명을 정의하여, 학과 Topic 요소들 간의 계층관계를 표현할 수 있다.

학과

Ontology > Topic Types > Topic Type Configuration

Supertype:

Name:

PSI:

Description:

Existing fields:

| Name | Name | Occurrence | Association |
|---------|--------|--------------|--------------------------|
| 표목(PSI) | URI | Exactly one | <input type="checkbox"/> |
| 학과이칭표기 | String | Zero or one | <input type="checkbox"/> |
| 목표 | String | Zero or more | <input type="checkbox"/> |
| 소개 | String | Zero or one | <input type="checkbox"/> |
| 교과과정 | String | Zero or one | <input type="checkbox"/> |
| 홈페이지 | URI | Zero or one | <input type="checkbox"/> |
| 세부전공 | 학위논문 | Zero or more | <input type="checkbox"/> |
| 전공자 | 인물정보 | Zero or more | <input type="checkbox"/> |
| 이후 학과명 | 학과 | Zero or more | <input type="checkbox"/> |
| 이전 학과명 | 학과 | Zero or more | <input type="checkbox"/> |

Add fields:

Name:

Occurrence:

Association:

Other:

[그림 22] 학과명 Topic Type 설정화면

학위논문의 저자(인물정보)는 전공자로서 학과와의 연계성을 가지며, 학위논문은 학과의 전공으로 연계성을 형성할 수 있다.

생성된 학과명 Topic Type에 대한 인스턴스를 등록한다. 전공학과에 대한 소개와 교과과정, 목표 등의 Occurrence와 학과에서 나온 학위논문명, 전공자, 그리고 전공이칭표기 즉, 기초DB구축에서 조사한 이전학과명, 이후학과명 등을 Association으로 등록한다. 학과 이력이 확인되는 경우는 이전학과명, 이후학과명으로 Association을 연결하여 구축하고, 확인되지 않는 다양한 학위논문 학과, 전공에 대한 표기는 학과이칭표기로 다수의 명칭을 Occurrence로 등록할 수 있다.

학과명 시소러스 기초DB 구축에서 작업된 CMS시스템 내의 값도, 토픽맵에서 설계한 타입과 매핑하여 일괄변환이 가능하다. 등록된 학과를 웹서비스 상에서 조회해 보면, 전공에 대한 전공자와 논문들, 그리고 학과명에 대한 다양한 표기방식을 한 화면에서 상세 열람이 가능하므로, 검색 접근성을 향상 시킬 수 있다. [그림 23]이 토픽맵 서비스에서 제공하는 웹서비스 화면으로 예시 데이터를 작성

한 결과이다.

| | |
|---|--|
| <p>Untyped Names (1)</p> <ul style="list-style-type: none"> • 화학생명공학부 | <p>Hierarchy (역대제출력서)</p> <ul style="list-style-type: none"> • 화학생명공학부 ◦ ... |
| <p>Associations (4)</p> <ul style="list-style-type: none"> • 이전 학과명 <ul style="list-style-type: none"> ◦ 화학공학과 • 전공자 <ul style="list-style-type: none"> ◦ 최차용, 1942 ◦ 황재용, 1971 • 학과논문 <ul style="list-style-type: none"> ◦ 바이오필터용 담체 특성 연구 및 신개념 바이오트리플링 필터의 개발 | <p>Internal Occurrences (7)</p> <ul style="list-style-type: none"> • 교과과정 <ul style="list-style-type: none"> ◦ 본 학부는 정밀화학 / 생물 화학공학 / 유기기재료(고분자반도체 재료) 등의 원천기술로부터 반응공정, 환경, 유동제어, 공정시스템 등의 실용화 기술에 이르기까지 일관된 전문지식을 학생들에게 제공하여, 기술과 경제 분야에 대한 폭넓은 시야와 풍부한 인성을 겸비한 지도자 및 세계 수준의 연구자 양성이라는 교육 목표에 중점을 두고 있다. 2004년 기준으로 학부생중 약 43%가, 그리고 거의 모든 대학원생이 장학금의 혜택과 지원을 받고 있으며, 총 수혜액은 약 24억원에 이릅니다. • 목표 <ul style="list-style-type: none"> ◦ 화학생명공학과 관련된 산업, 연구, 학계를 이끌어 나갈 인재를 배출하기 위해 다음과 같은 교육 목표를 설정한다. 1. 화학생명공학 산업, 연구, 학계의 지도자적 인력을 육성한다. 2. 엔지니어로서의 자질을 갖추기 위해 필요한 의사전달 능력을 포함한 기본 소양 및 자연과학적 기초 지식을 교육한다. • 소개 <ul style="list-style-type: none"> ◦ 본 학부는 2004년 기준으로 약 550명의 학부생과 290명의 대학원생, 그리고 총 31명의 교수진으로 이루어진 거대한 조직체이며, 서울대학교가 지향하는 대학원 연구중심교육의 학부로 발전을 꾀하고 있다. 특히 1999년부터 정부의 지원으로 학부교수 전원이 참여하는 [Brain Korea 21 항공사업단]이 발족하여 교육과 연구의 수준을 세계적인 수준으로 향상시키고 있습니다. 본 학부 졸업생들은 정유, 석유화학, 항성제품, 반도체, 제약, 시멘트, 유리, 화학, 펄프, 금속제련 등의 산업체와 국내외 대학에서 유능한 지도자로서 활약을 하고 있으며, 특히 국내 중진급 인사의 15%가 본 학부 출신이라는 점은 본 학부의 국가경제에서의 높은 기여도를 잘 반영하고 있습니다. 본 학부내에는 화학공정신기술연구소와 학부교수가 관련하는 7개의 벤처 회사가 있으며, 그외 정부로부터 지원받는 대단위 연구과제가 다수 수행되고 있다. 이를 통하여 대학 내 학문적 성과들이 산어체로 응용되는 산업화 과정을 통하여 사외에 직접 기여하고 있습니다. 화학생명공학부는 국제적인 경쟁력을 갖춘 인재를 키우기 위해 미래를 위한 도전과 기회를 제공하고 있으며, 「인류를 위한 편리하고, 건강하고, 쾌적한 미래」 구현에 정진하고 있습니다. • 학과이칭표기 <ul style="list-style-type: none"> ◦ 공업화학과 ◦ 응용생명화학부 ◦ 응용화학과 |

[그림 23] 전공학과에 대한 토픽맵 서비스 화면 예시

4) 논문 핵심어 DB구축 방안

학위논문에 있어서 저자가 핵심어로 추출한 단어는 비교적 주제로 접근할 수 있는 중요한 포인트가 되기 때문에 무시할 수 없는 요소이며 이 부분을 최대한 활용하여 이용자의 접근도를 높여야만 한다. 학위논문을 작성한 저자가 기술한 핵심어이므로, 이를 키워드 검색으로 활용할 경우, 재현율도 높아질 뿐 아니라, 다양한 접근점을 제공할 수 있다. 따라서 기 구축된 핵심어 이외에 미구축 분의 학위논문에 대해서도 핵심어를 입력하고 이를 토대로 학위논문의 핵심어 검색이 되도록 해야 한다.

보통 학위논문마다 대략 10개 이상의 핵심어가 등록되어 있고, 이는 통제된 어휘도 아닐 뿐더러 표기상에 있어서도 다양하다. 그러므로 이를 위한 보완으로 핵심어를 시소러스나 사전 등으로 정제작업이 필요하다. 이때 국립중앙도서관이나 국회도서관에서 구축한 시소러스를 활용한다든지, 상용DB에서 제공하고 있는 시소러스 등을 활용하여 핵심어 사전DB 또는 참조DB를 구축하여 이를 이용한 통

제된 핵심어 등록기능을 도입한다.

유사한 의미의 키워드들 간에 하나의 대표 Topic으로 등록하고, 핵심어 이칭표기로 동의어 등을 표기하고 학위논문에서와 같이, 영문표기도 함께 기술하여 관리한다. 또한 핵심어와 관련 학위논문 간에도 Association을 정의하여 연관정보를 얻을 수 있도록 하며, 시스템 통합차원에서 핵심어 개념에 대한 PSI를 term의 학위 디렉토리에 unique하게 부여하여 관리한다.

핵심어

Ontology > Topic Types > Topic Type Configuration

Confirm
Reset

Supertype: -- no supertype --

Name: 핵심어

PSI: http://psi.snu.ac.kr/keyword ⌕ +

Description:

Existing fields:

| Name | Name | Value | Options |
|----------|--------|--------------|---------|
| | Name | Exactly one | ⌕ ⊖ |
| 표목(PSI) | URI | One or more | ⌕ ⊖ ⊗ ⊕ |
| 핵심어이칭표기 | String | Zero or more | ⌕ ⊖ ⊗ ⊕ |
| 영문핵심어표기명 | String | Zero or more | ⌕ ⊖ ⊗ ⊕ |
| 내용설명 | String | Zero or one | ⌕ ⊖ ⊗ ⊕ |
| 관련논문 | 학위논문 | Zero or more | ⌕ ⊖ ⊗ ⊕ |

Add fields:

Name: ⊕

Occurrence: Creator ⊕

Association: 논문저작 ⊕

Other: Subject identifier ⊕

System properties:

[그림 24] 온토폴리의 핵심어 등록설정 화면 예시

검색에 있어서도 이용자가 입력한 쿼리와 동일한 것이 일치되는 것을 우선순위로 보여주고, 관련 핵심어 등으로 표현하여 이칭으로 입력된 값도 검색되도록 한다. 핵심어나 학과명은 등록이나 DB정의 등의 성격이 유사하며, 이용자가 검색한

후 오류신고, 유사어의 추가 등록 및 수정, 보완을 할 수 있도록 웹에 처리하여, 보다 정확하고 풍부한 정보를 제공할 수 있는 이용자 참여형 형태로 구축한다.

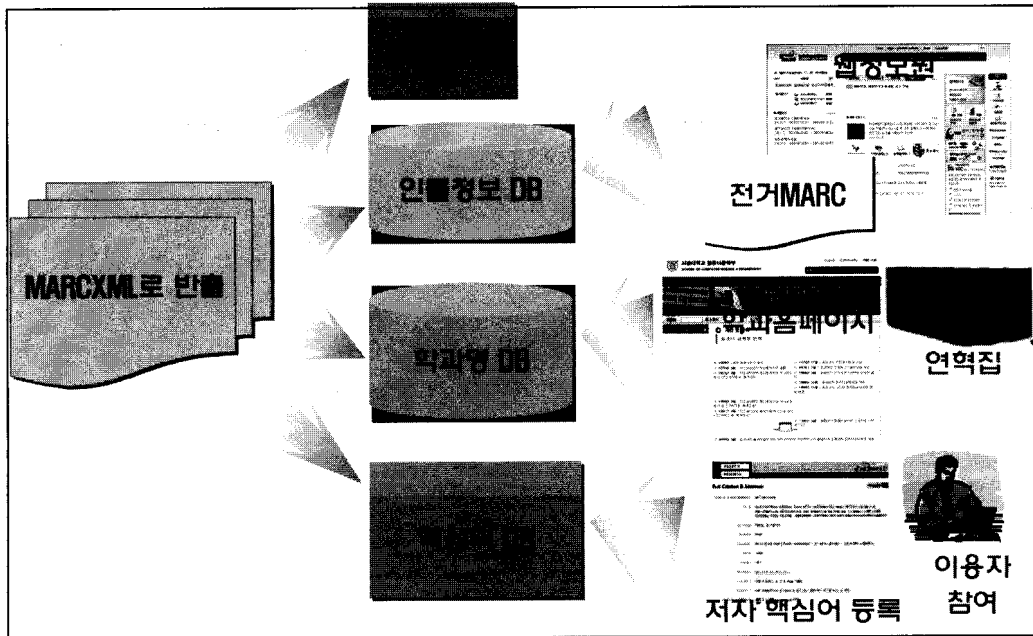
가) 논문 핵심어 구축절차

- ① 구축된 학위논문 DB 중에서 서지레코드에 653tag 비통제주제어가 누락된 데이터를 대상으로 SOLARSIII 시스템의 학위논문 서지MARC 데이터에 직접 입력한다.
- ② 핵심어를 한글, 영문, 유사어 등으로 그룹핑하여 시소러스를 작성한다.
- ③ 핵심어 시소러스를 위한 다음과 같은 참고정보원을 활용하도록 하며, 기 구축된 시소러스를 최대한 활용하며, 이때 출처정보원을 표시해 준다.
- ④ 핵심어는 한글, 영문, 유사어 등으로 사전을 구축하되, 기존에 입력한 값과 비교하여 “유사어로도 검색하시겠습니까?” 라고 메시지를 제공하며, 유사어를 클릭하면 해당 유사학파로 사전 만든 것을 이 모두 검색되도록 한다. 이 이용자 참여형 서비스를 위한 학위논문 검색시 핵심어를 추가할 수 있는 인터페이스를 제공하도록 하며, 입력된 내용은 SOLARSIII CMS시스템에서 관리자의 검증을 거쳐 등록, 관리한다.

[표 17] 핵심어 조사를 위한 참고정보원

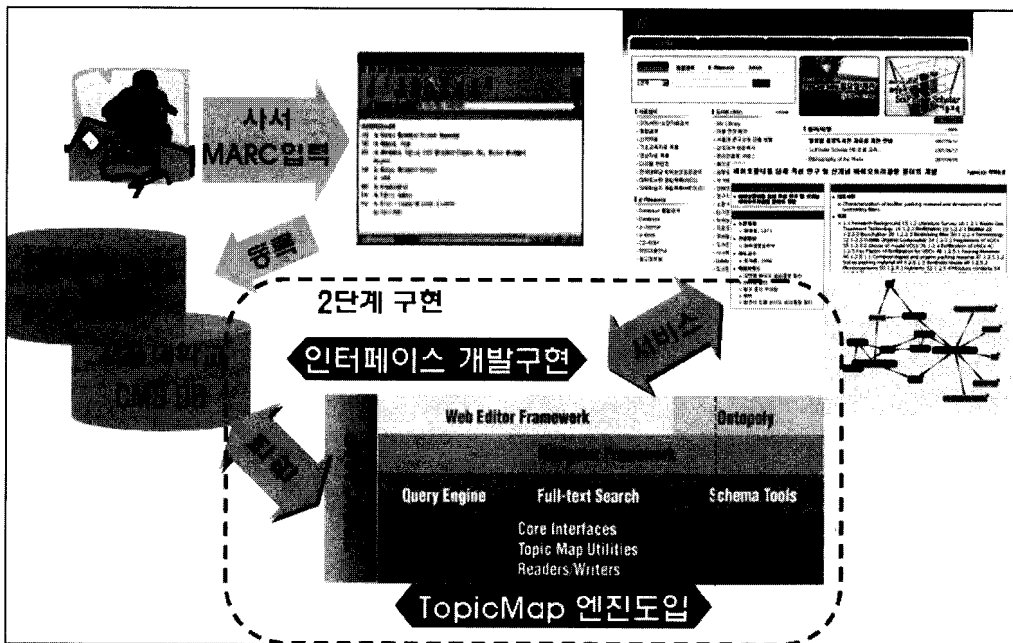
| 번호 | 참고정보원 | 비고 |
|----|--------------|-----------|
| 1 | 국회도서관 시소러스 | |
| 2 | 국립중앙도서관 시소러스 | |
| 3 | 상용 단어 사전 | |
| 4 | 각종 사전류 | 국내의 모두 포함 |

[그림 25]는 학위논문 서비스를 위한 기초DB 구축의 전체 업무 구성도를 요약해 주고 있다. 이러한 기초 DB 구축을 통하여 온톨로지 기반의 토픽맵을 구축하고 궁극적으로 시맨틱 디지털도서관 구축을 이룰 수 있는 것이라고 본다.



간의 연관관계가 확보되고, 기존의 꼬리물기 재검색이 아닌, 한 화면에서 학위논문 메타데이터 요소와 저자 인물정보, 학과 전공정보와 핵심어에 대한 연계성을 높여주는 서비스화면과 관계정보 자원간의 vizigator를 통한 네비게이션 서비스도 제공한다.

학과명, 논문 핵심어 등에서는 이용자 참여형 서비스를 강화한 서비스를 시도한다. [그림 26]은 2단계의 온톨로지 기반 토픽맵 서비스 구현을 위한 절차를 설명하고 있다.

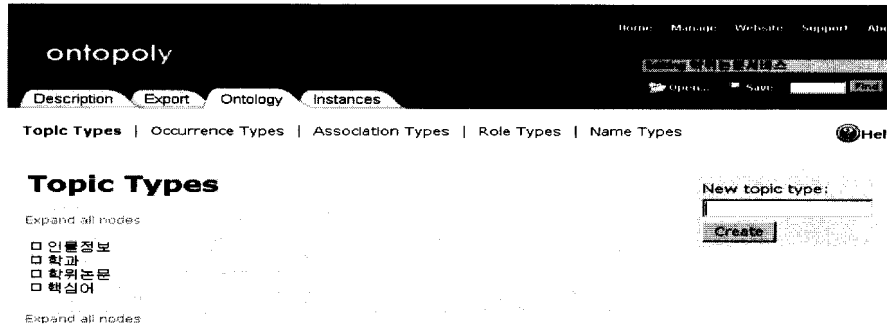


[그림 26] 2단계 토픽맵 기반 학위논문 서비스 구성

지금까지 학위논문 서비스 토픽맵 디자인을 바탕으로 필요한 기초DB구축 방안에 대하여 살펴보고, 서비스 모델링의 이해를 돕고자 토픽맵 편집기인 온토폴리를 이용하여 샘플 인스턴스도 생성해 보았다. 학위논문 온톨로지 구현에서 채택한 토픽맵을 기반으로 온토폴리에 의한 토픽맵 생성과, 생성한 토픽맵을 온토폴리아의 옴니게이터(omnigator)를 통하여 관계성과 정보층의 내용 열람을 간단한

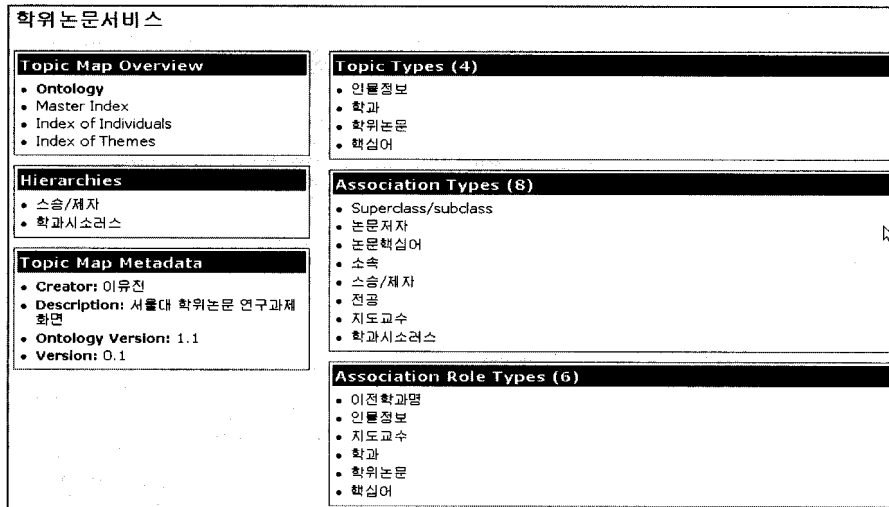
예로 살펴보도록 한다.

다음은 온토플리의 구성과 첫 입력 화면이다.

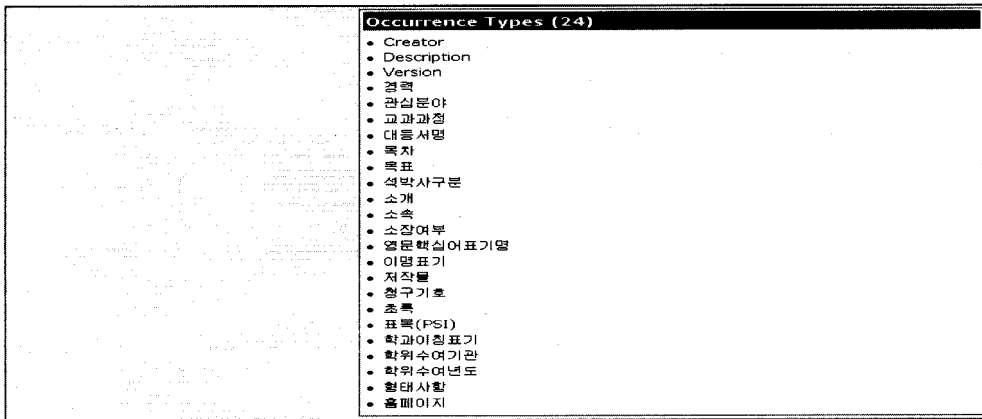


[그림 27] 온토플리 Topic Type 열람 화면

토픽맵으로 구성하면서 필요한 메타데이터의 요소들을 살펴보고, 요소들간의 관계성도 알아보았다. 이를 근거로 온토플리에서 토픽맵 디자인에 맞춘 Topic Type, Occurrence Type, 그리고 Association Type의 정의를 옴니게이트 브라우저로 살펴보면 다음과 같다.



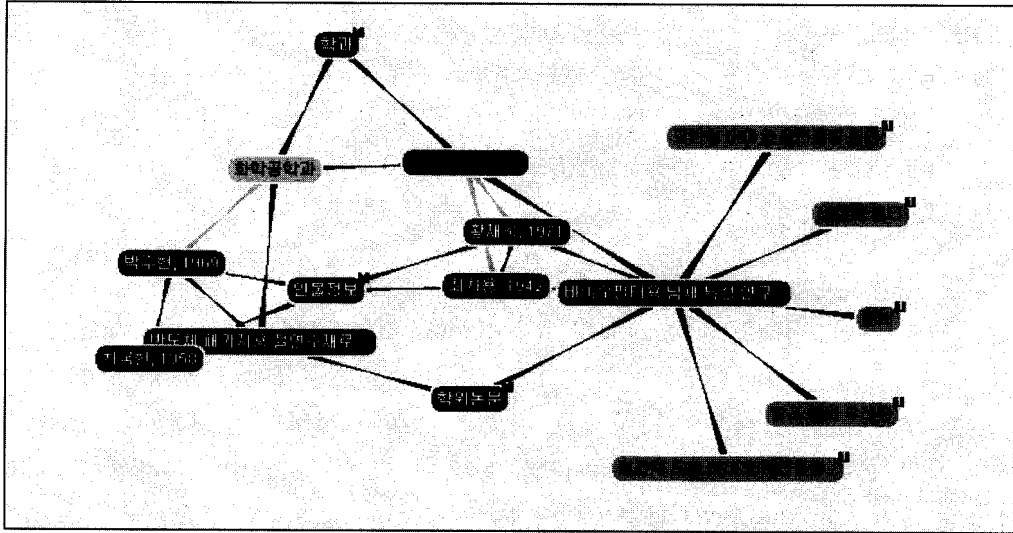
7) 옴니게이트(Omnigator) : ontopia의 OKS에서 제공하는 토픽맵 브라우저의 기능과 도식화된 표현인 vizigator를 제공하는 tool이다.



[그림 28] 옴니게이트의 학위논문 토픽맵 요소 열람

토픽맵의 주요 요소들에 맞추어 각각의 Type들을 정의해 보았고, 이제 이를 기반으로 인스턴스를 생성한다. 학위논문 토픽에 대한 인스턴스 입력 화면은 [그림 29]와 같이 Occurrence 입력부분과 [그림 30]의 Association 입력부분으로 구성된다.

[그림 29] Topic의 Occurrence 입력 화면



[그림 32] 토픽맵 vizigater

위의 그림은 움니게이트에 의하여 서비스되는 vizigater로 생성된 토픽들 간의 관계위주로 비주얼한 서비스 도식화를 보여주고 있다. 여기서 학과간의 연계와 학위논문에 대한 핵심어, 저자, 지도교수간의 관계를 보다 쉽게 확인할 수 있는 비주얼한 서비스를 제공하고 있다.

4.4 토픽맵 기반 학위논문 서비스 의의 및 활용 방안

1) 의의

- 석·박사 학위논문 서비스부분에서 다양한 사용자 맞춤형 지식 기반 서비스가 될 수 있도록 온톨로지 활용 표준 서비스 모델을 제공해 봄으로써 관련 주제간의 의미적 상호연계성이 확보되어 종합적 정보접근이 용이하며 사용자가 원하는 정보를 한 눈에 찾아볼 수 있는 시맨틱(semantic) 지식 서비스를 제공한다.
- 기존 통합 검색 서비스 및 서울대학교 전자도서관과 연계한 온톨로지 지식 기반 서비스로 확장됨으로써 검색의 효율과 정확성이 향상되어 이용자의 연구와 학습 역량이 강화된다.
- 서울대학교가 보유하고 있는 주요한 학술정보에 대한 풍부한 지식 정보를

기반으로 온톨로지를 구축함으로써, 지식을 구조화하고 의미적 상호운용성을 확대하여 지식의 표현, 공유와 재사용의 효과가 매우 크다고 할 수 있다.

2) 활용방안

- 지식자원의 서비스 제공 메커니즘을 확장하여 온톨로지 기반의 지능형 지식 서비스를 위한 비즈니스 아키텍처 설계에 활용하고 지식 서비스 내용의 구성 및 재구성을 위한, 그리고 기존의 애플리케이션 및 지식정보 데이터소스와의 상호연동을 위한 서비스 구성 시스템 설계에 활용한다.
- 지식포털 서비스를 구성, 재구성, 또는 평가할 수 있는 온톨로지 기반의 플랫폼 개발에 적용하여 지식포털 서비스의 의미와 절차를 모델링하고, 지식포털 서비스를 쉽고 일관되게 재구성할 수 있게 하며, 맞춤형의 지식 서비스 제공이 가능한 온톨로지 활용 모델을 제공하고 토픽맵 기반으로 그 효율성을 극대화 한다.

3) 기대효과

- 지식 사용의 필요를 충족하는 온톨로지 기반의 지식 서비스 시스템을 구축하고, 지식사용자 자신의 지식 수요에 맞는 메타 정보를 구축하여, 경쟁력 있는 양질의 정보와 맞춤형 지식 서비스를 제공 할 수 있는 기반을 조성한다.
- 지식 사용자로부터 요청을 받은 서비스 에이전트는 주제 기반 처리엔진, 온톨로지, 웹 엔진, 메타데이터 변환의 유기적인 정보 교환을 통하여 사용자의 의도를 정확하게 파악하여 정보의 내용을 이해함으로써 정보를 보다 정밀하게 선별하고 지능적인 분류 및 검색 실행 기술을 습득한다.
- 지식정보자원 사용자는 시맨틱 서비스 사용 과정에서 서비스에 대한 메타정보를 추가하여 더욱 양질의 서비스를 만드는 데 기여할 수 있으며, 이러한 주제와 연관된 관계 정보는 결국 지식 발전을 도모할 수 있는 양질의 정보로 축적된다.

제 5장 온톨로지 기반 디지털 도서관 고도화 전략

4장까지 온톨로지만 무엇인지, 온톨로지 표준의 양대 산맥인 ISO의 토픽맵과 W3C의 RDF/OWL에 대하여 살펴보았다. 이중 토픽맵을 기반으로 서울대학교의 학위논문 콘텐츠를 중심으로 온톨로지 서비스를 구현하고자 토픽맵 모델링 방법론을 살펴보고, 이를 적용하여 학위논문 온톨로지 구현방안을 살펴보았다. 구체적인 실현을 위하여 1단계 기초DB구축과 2단계 토픽맵 서비스 구현을 위한 기존 시스템과의 인터페이스 구현방안 등을 연구하였다.

정보화사회에서 미래의 최첨단 정보기술 사회로의 도약에 맞추어 서울대학교 도서관도 미래지향적 디지털 도서관의 밑그림과 전략을 필요로 한다. 1차로 학위논문이라는 서울대학교의 지식산출물이면서 이용자의 접근성이 높은 콘텐츠를 대상으로 온톨로지를 적용한 서비스를 시작하였다면, 이제 장기간의 고도화 전략을 세워 발전을 거듭해 나아가야 할 것이다.

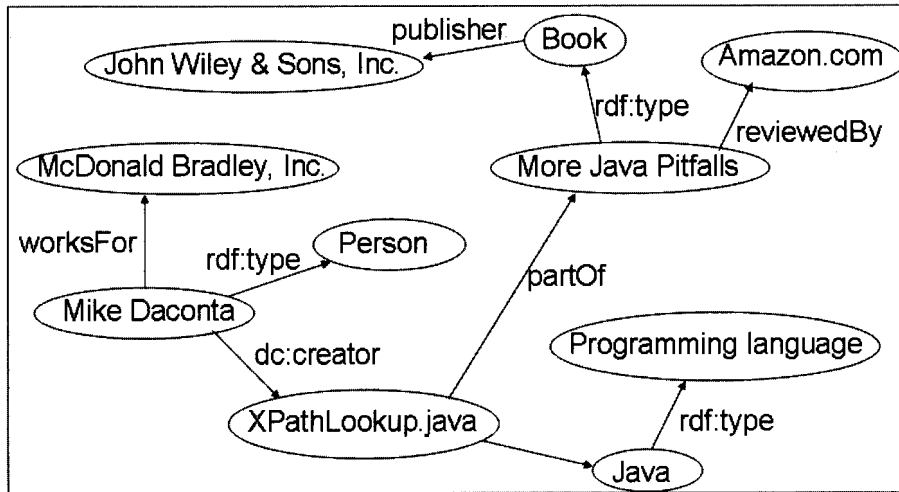
이에 웹2.0의 환경, 시맨틱 웹의 개념과 지식정보사회에서 요구하는 바를 살펴보면, 서울대학교의 온톨로지 기반 시맨틱 디지털 도서관을 지향하는 고도화 전략을 구상해 보고자 한다.

시맨틱 디지털 도서관은 이용자 참여가 활성화되는 브라우징과 인터페이스를 요구하며, 지식의 통합, 공유를 통하여 서비스의 상호운용성을 보장해야 한다. 이용자 간의 커뮤니티를 활성화하면서 사회적 협업이 이루어지는 서비스 방향을 모색하는 도서관을 지향하고 있다.

5.1 정보관리와 시맨틱 웹

시맨틱 웹은 현재의 인터넷과 같은 분산환경에서 정보자원(웹 문서, 각종 파일, 서비스 등)사이의 관계-의미 정보(Semantics)를 기계(컴퓨터)가 처리할 수 있는 형태로 표현하고, 이를 자동화된 기계(컴퓨터)가 처리하도록 하는 프레임워크이자 기술을 말한다. 즉, 시맨틱 웹은 지금과 같이 사람만이 웹에 산재한 정보의 의미를 파악하는 것이 아니라 기계도 해석할 수 있는 일종의 표준 의미정보 교환의 수단이다. 시맨틱 웹을 가능하게 하는 것은 RDF다. RDF는 정보와 자원을 기술하기 위한 데이터 포맷으로 시맨틱 웹을 구축하기 위한 기본적인 데이터 모델이다. RDF는 기계가독형으로 사물의 속성(크기, 무게, 나이, 가격)을 기술함으로써 기계

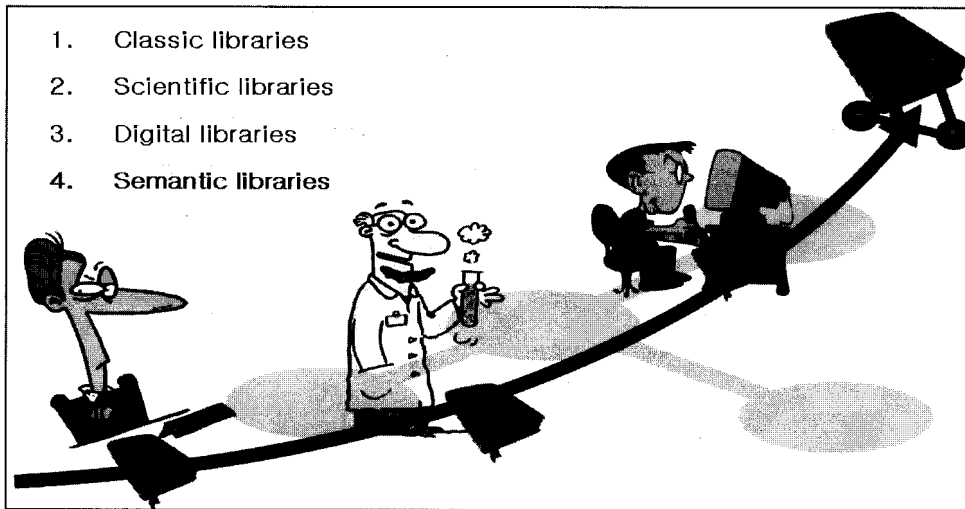
처리를 가능케하기 때문에 정보를 RDF 파일로 표현하면 RDF crawlers가 웹에 있는 정보를 탐색, 발견, 발취, 수집, 분석, 처리할 수 있다.



[그림 33] RDF의 예

5.2 시맨틱 디지털 도서관의 개념

시맨틱 디지털 도서관은 전통적인 도서관이나 디지털도서관과 다른 완전히 새로운 개념의 도서관이 아니라 기존의 디지털 도서관을 확장시킨 개념이다. 즉, 기존의 디지털도서관이 자관이 보유하고 있는 정보자원을 디지털화하여 서비스하는 개념이라면, 시맨틱 디지털도서관은 정보자원을 컴퓨터도 이해가능하도록 기술함으로써 정보자원이 서로 통합되고, 상호운용이 가능하며, 보다 이용자에게 친숙한 인터페이스를 제공하는 개념이다. 시맨틱 디지털도서관에서는 정보자원은 디지털 콘텐츠 및 그 집서(collection), 이용자 및 이용자 커뮤니티를 포함하는 개념이다. 통제어휘집, 시소러스, 분류표 등의 메타데이터도 RDF 형식으로 기술하여 정보자원의 의미를 도출시킴으로써 서로 다른 도서관 시스템간의 커뮤니케이션을 위해 중재서비스 (mediation services)도 가능하게 된다.



[그림 34] 도서관의 진화

5.3 시맨틱 디지털 도서관의 기능

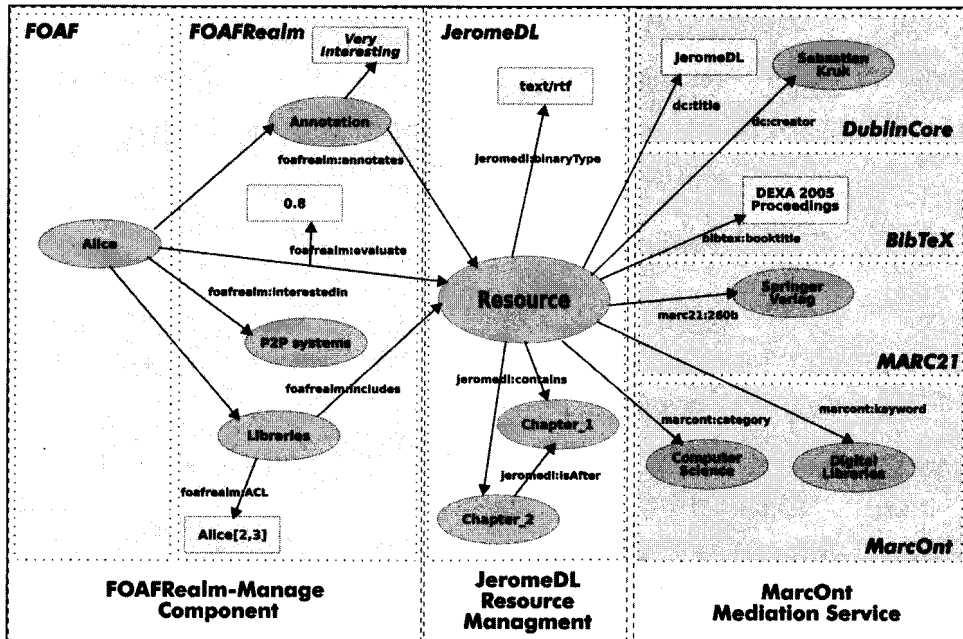
시맨틱 디지털도서관의 핵심 기능은 자원의 통합, 자원간의 상호운용성, 이용자 중심의 탐색 및 브라우징으로 요약된다. 시맨틱 디지털도서관은 통합을 통해 자원, 이용자 프로필, 북마크, 분류표 등 서로 다른 메타데이터로 기술된 정보를 통합할 수 있다. 둘째, 메타데이터 수준 또는 커뮤니케이션 수준에서 다른 도서관 시스템과의 상호운용성을 보장할 수 있다. 셋째, 의미성(semantics)의 강화로 보다 견고하고, 이용자 편의적이며 적응성있는(adaptable) 탐색과 브라우징 인터페이스의 제공이 가능하다.

시맨틱 디지털도서관에서는 정보공간의 새로운 탐색 패러다임을 제공한다. 시맨틱 디지털도서관의 이용자는 온톨로지 기반의 탐색/패킷 탐색이 가능하며, 커뮤니티 기능이 부가된 브라우징이 가능하다. 또한, RDF 등의 형식언어를 이용하여 데이터 수준의 상호운용이 가능하여 이종의 다양한 정보원에서 생성된 메타데이터를 통합할 수 있기 때문에 서로 다른 디지털 도서관 시스템에 소장된 서지레코드를 서로 연결할 수 있다. 디지털도서관의 메타데이터를 이종의 네트워크환경에서 정보검색과정을 통합하기 위한 시도로서 MARC21, DublinCore, OAI-PMH, Z39.50, SKOS, Dienst (<http://www.cs.cornell.edu/NCSTRL/protocol.html>), ODP 등의 프로토콜 등이 개발되어왔다.

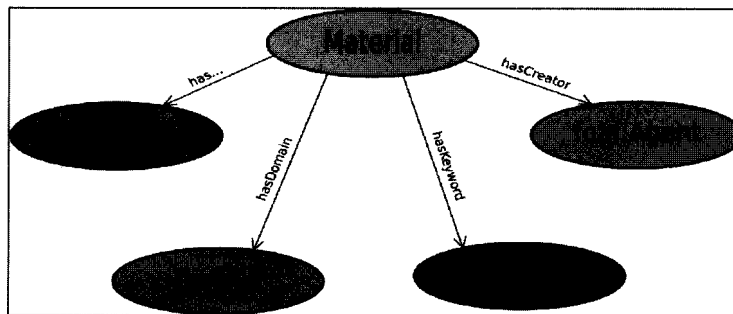
5.4 시맨틱 디지털도서관의 사례

시맨틱 디지털 도서관 시스템의 사례로는 SIMILE, BRICKS, JeromeDL 등이 있다. SIMILE(Semantic Interoperability of Metadata and Information in unLike Environments)은 기존의 Dspace 프로젝트를 확장하여 디지털 자산, 스키마, 메타데이터 서비스간의 상호운용성의 증진을 목적으로 RDF 형태의 이종의 메타데이터를 브라우징, 탐색, 매핑하기 위한 도구를 개발하였다. 이 프로젝트의 부산물로 메타데이터 관리사용 도구로서 XML 데이터를 RDF로 변환하는 도구(RDFizers)와 RDF 그래프를 조회/편집하는 도구(Welkin) 등이 개발되었다. 이용자를 위한 도구로는 웹기반의 RDF 패킷 메타데이터 브라우저(Longwell)와 개인, 집단, 커뮤니티가 데이터를 출판, 공유할 수 있는 웹기반 서버(Semantic Bank) 등이 개발되었다. BRICKS는 유럽의 전통문화 분야의 지식과 자원을 공유하기 위해 디지털 도서관 네트워크를 위한 프로젝트로 서비스 중심, 분산형, 오픈소스를 그 모토로 하고 있다.

JeromeDL은 시맨틱 웹기술과 Social Networking 기술을 활용하여 디지털 도서관의 상호운용성과 유용성을 증진시키는 것을 목적으로 하고 있다. JeromeDL에서는 MARC21, BibTeX, 더블린 코어가 다루는 개념을 포괄하는 온톨로지(MarcOnt)를 이용하여 다른 도서관(repository)과 커뮤니케이션하기 위한 중재 서비스를 제공한다.



[그림 35] JeromeDL의 온톨로지



[그림 36] MarcOnt Ontology의 주요개념

특히, JeromeDL은 인적네트워크 메타데이터인 FOAF(Friend-of-a-Friend)를 이용하여 이용자 커뮤니티 내에서 Social Semantic Collaborative Filtering(SSCF)을 지원한다. JeromeDL은 FOAF의 foaf:know 정보를 이용하여 정보자원의 소유자와 요청자간의 거리, 소유자와 요청자간의 친한 정도(level of friendship)를 social

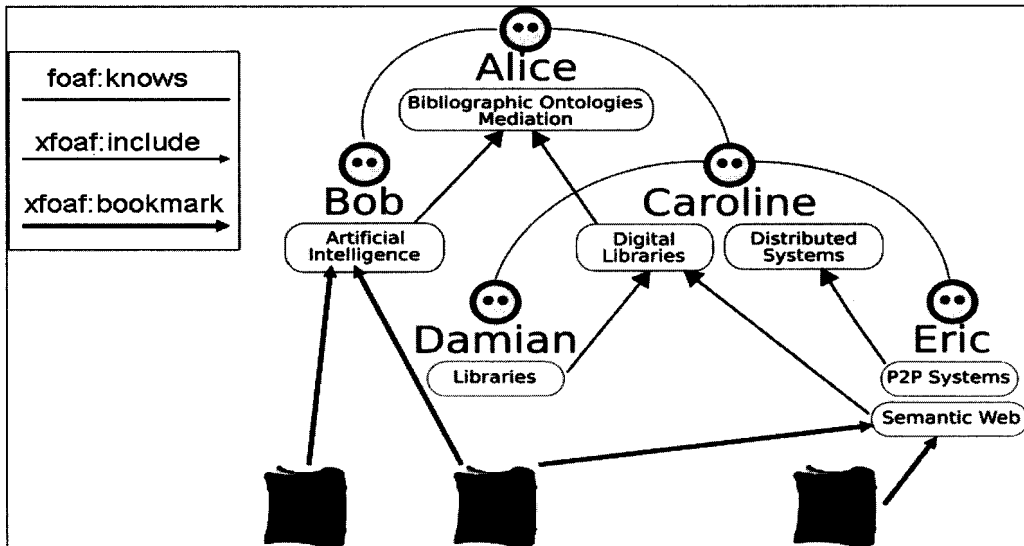
network의 그래프로 계산하여 특정 이용자의 정보자원이나 목록에 접근할 수 있는 권한을 관리할 수 있다.

```

> FOAF RDF vocabulary
<foaf:Person>
  <foaf:name>Mike Daconta
</foaf:name>
  <foaf:knows>
    <foaf:Person>
      <foaf:name> Kevin Smith
    </foaf:name>
  </foaf:Person>
</foaf:knows>
</foaf:Person>

```

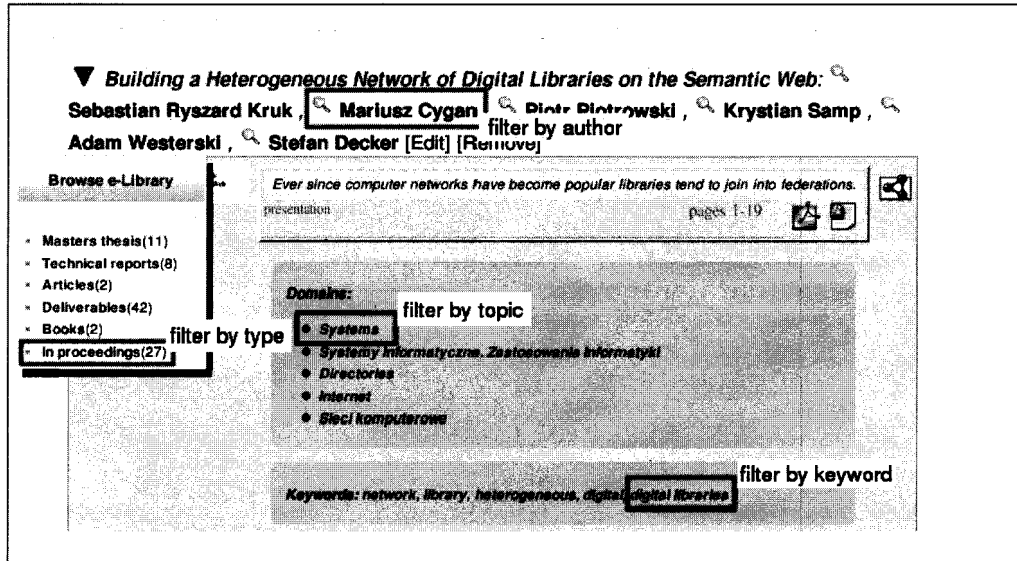
[그림 37] FOAF의 예



[그림 38] FOAF를 이용한 Social Semantic Collaborative Filtering 개념

도서관 이용자가 콘텐츠를 업로드 할 때에는 기존의 도서관 분류표(DDC)를 이용하여 도메인을 설정하고 DMoz 또는 WordNet에서 가져온 의미정보를 이용하여 특정 콘텐츠에 대해 키워드를 부여하여 이용자의 목록을 유지하고 이를 바탕

으로 앞서 기술한 social semantic collaborative filtering이 가능하게 될 뿐만 아니라 의미적 탐색과 질의 확장이 가능하게 된다.



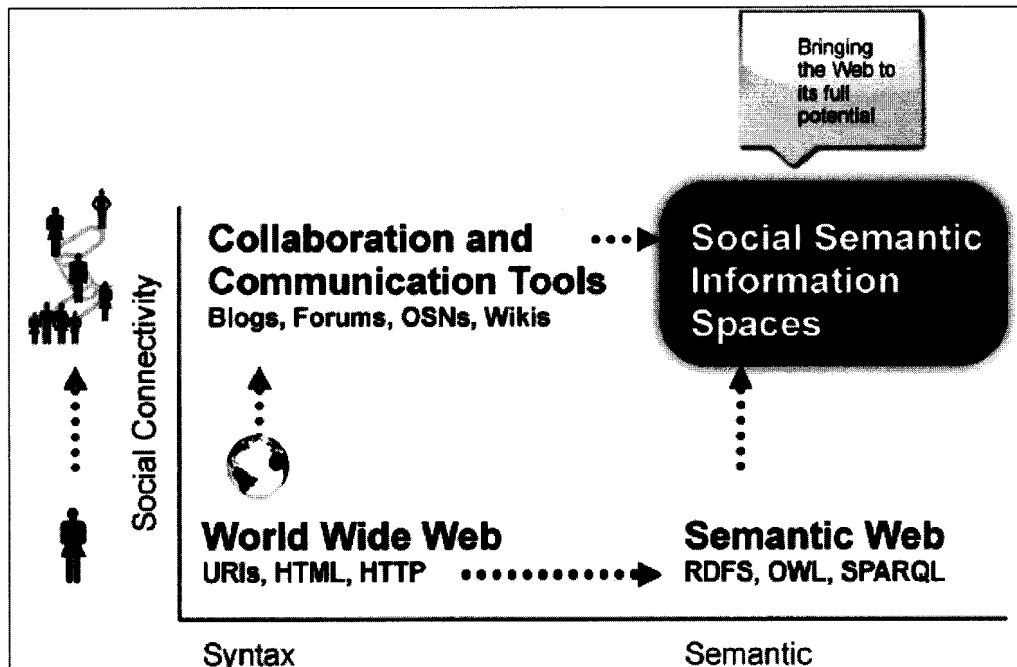
[그림 39] JeromeDL의 이용자 인터페이스

5.5 시맨틱 디지털도서관 구축을 위한 제언

웹 2.0을 규정짓는 3가지 특징은 참여, 개방, 분산이다. 이용자의 참여를 통해 정보를 공동으로 조직, 공유, 필터링, 탐색하고, 하나 이상의 소스로부터 생산되는 콘텐츠를 통합된 경험(환경)으로 연결하며, 특정한 소스로부터 생산된 정보만을 서비스하기보다는 분산된 환경에서 자발적으로 생성된 정보를 통합하는 일이 중요하다.

| Web 1.0 | Web 2.0 | Semantic Web 2.0 |
|-----------------------------|-------------------------------|--|
| Personal Websites | Blogs | Semantic Blogs |
| Content Management Systems | Wikis | Semantic Wikis |
| Altavista, Google | Google Personalised, DumbFind | Semantic Search |
| CiteSeer, Project Gutenberg | Google Scholar, Book Search | Social Semantic Digital Libraries |
| Message Boards | Community Portals | Semantic Forums and Community Portals |
| Buddy Lists, Address Books | Online Social Networks | Semantic Social Networks |
| - | - | Semantic Social Information Spaces |

[그림 40] 웹, Web 2.0, Semantic Web2.0의 특징 비교



[그림 41] Social 시맨틱 디지털 도서관의 방향

이러한 웹 2.0의 개념을 디지털도서관에 적용하기 위해서는 첫째, 이용자가 콘텐츠 어노테이션(tagging) 과정에 참여하도록 해야한다. 시맨틱 디지털 도서관에서 이용자는 이용자이면서 동시에 공헌자로서 유용한 웹자원에 대해 태깅할 수 있어야 한다(예: CiteUlike). 둘째, 이용자/독자가 그들의 지식을 커뮤니티 내에서 공유하도록 유도함으로써 소셜 네트워크를 기반으로 정보를 공동으로 필터링을 할 수 있는 환경을 제공해야한다. 셋째, 커뮤니티 내부 또는 커뮤니티간 보다 좋은 커뮤니케이션 수단을 제공하여 하나의 도서관 서비스를 뛰어넘어 다양한 서비스로부터 생성된 콘텐츠를 경험할 수 있도록 해야 한다.

제6장 결론

디지털 도서관은 전통적 도서관 서비스인 수서, 목록, 대출, 연속간행물 등의 업무 모듈에 대한 자동화 시스템 구축에서 시작되었다. 이후 자관에서 보유한 콘텐츠에 대한 관리의 강화, 이용자 서비스의 만족도와 효율성을 높이기 위한 원문, 초록, 목차 등의 디지털 콘텐츠 확보 및 구축으로 디지털 도서관이 발전하였다. 지식 정보사회가 도래하면서 지식정보의 생산과 수요가 급격히 증가하고 지식정보의 유통 채널이 다양화 되었다. 이에 따라 전자도서관도 다양한 웹 정보원과 참조 정보원을 활용한 서비스, 대학도서관 간 또는 기관과 대학도서관 간의 상호대차 및 이용서비스 등으로 확대하게 되었고, 이용자 친화적 인터페이스를 유지하면서 다양한 기능과 콘텐츠를 서비스하게 되었다.

여기에 정보기술의 새로운 패러다임인 시맨틱 개념이 도입되면서 디지털 도서관의 서비스에도 변화를 가져오고 있다. 의미적 차원에서 정보에 대한 표현이 가능하고 이를 정보검색 등을 위해 기계가 자동처리할 수 있는 서비스를 구현하고자 하는 것이 시맨틱 웹의 비전이다. 최근 정보 소비자의 참여와 오픈 플랫폼을 통한 지식 서비스가 웹의 새로운 비즈니스 모델인 Web 2.0의 중심 개념이 되면서 전자도서관의 역할과 서비스 모델 또한 변화하고 있다.

온톨로지를 표현하는 방법으로 국제표준을 지향하고 있는 것이 토픽맵과 RDF/OWL 이라고 할 수 있다. 이 중에서 토픽맵은 주제로 표현된 지식들 간의 의미적 연관성과 관련정보를 연계시켜 다양한 서비스를 모색할 수 있으며, 사용이 용이하고 시스템간의 통합을 이룰 수 있는 장점이 있다.

토픽맵 혹은 RDF/OWL로 표현된 온톨로지 기반의 전자도서관 서비스를 구축함으로써 정보자원이 서로 통합되고, 의미적 상호운용이 가능하며, 보다 이용자에게 친숙한 인터페이스를 제공하는 시맨틱 디지털도서관이 현실화 될 수 있을 것이다. 시맨틱 디지털도서관은 정보자원 간 의미적 차원의 연결성이 가능할 뿐만 아니라 이용자 간 사회적 연결성과 지식소비자의 지식 생산에의 적극적 참여가 확대되는 것을 목표로 한다.

시맨틱 디지털 도서관을 준비하는 초기 단계로서 본 연구가 시작되었고, 학위논문의 의미적 정보검색을 위한 기반이 되는 기초DB 구축의 기대효과는 다음과 같다.

- 1) 학술적 가치가 뛰어난 서울대 학위논문과 연계된 지식정보를 창출하게 된다.
- 2) 교수연구업적물 콘텐츠와 연계한 고품질 인물정보 구축으로 학술정보를 제공한다.
- 3) 학술연구의 정확도와 관련도가 높은 지식정보에 대한 연계 서비스를 실시한다.
- 4) 주제별 지식정보서비스를 위한 지식관리를 지원한다.
- 5) 해당 논문의 지도교수가 지도한 같은 전공의 학위논문이 연계되어 검색 활용성을 향상시킨다.
- 6) 학과 시소러스 구축을 통한 검색의 재현을 향상과 핵심어 시소러스 구축을 통한 검색의 정확률과 재현율을 높여준다.

학위논문 서비스는 대학도서관 이용자의 지식정보의 산물이며, 동시에 대학을 구성하는 교수와 학생간의 정보교류와 커뮤니티로 참여를 유도하는 사회적 협업을 이룰 수 있는 인물정보를 구성요소로 포함하고 있는 콘텐츠이다. 따라서 시맨틱 디지털도서관의 일차적 서비스 대상으로 학위논문 서비스를 이용하여 필요한 콘텐츠를 위한 기초 DB를 구축하고, 이를 온톨로지 기반의 토픽맵 서비스로 활용하면서 서울대학교 디지털도서관의 서비스는 진일보 할 수 있을 것이다. 다양한 정보자원의 온톨로지 구축과 함께 이용자가 참여한 지식 콘텐츠와 커뮤니티의 생성과 관리를 위한 온톨로지를 구축하여 서비스에 통합함으로써 시맨틱 디지털 도서관으로 나아가는 새로운 방향을 모색할 수 있을 것이다.

[참고문헌]

- 정보문화진흥원. 2006. "국가지식정보 온톨로지 표준개발" 연구보고 06-26
- Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, 36(2): 3-10.
- Bystrom, K. & Jarvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing and Management*, 31(2): 191-213.
- Cutter, C.A. 1875. *Rules for a printed dictionary catalogue*. Washington, DC: Government Printing Office.
- Flores, R.A. (1997). *Java Concept Maps for Learning Web*. Available at: <http://pages.cpsc.ucalgary.ca/~robertof/publications/edmedia97/>
- Garshol, L. M. (2003). Living with topic maps and rdf. Available at: <http://www.ontopia.net/topicmaps/materials/tmrdf.html>
- Garshol, L.M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Ontopia. Available at: <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>
- Gruber, T.R (1994). What is an ontology? Available: <http://www.ksl.stanford.edu/ksl/what-is-an-ontology.html>
- Guo, Y, Pan, Z., & Heflin, J. (2004). An Evaluation of Knowledge Base Systems for Large OWL Datasets. *Third International Semantic Web Conference*, Hiroshima, Japan, LNCS 3298, Springer, 2004, 274-288.
- Janes, J. (2002). Comparing the means of two groups the t-test, *Library Hi Tech*, 20(4), 469-471.
- Kang, I-H. and Kim G. (2003). Query Type Classification for Web Document Retrieval. *SIGIR Forum*, 26(2). 64-71.
- Kim, H.H. (2005). Ontoweb: Implementing an ontology-based web retrieval system. *Journal of the American Society for Information Science and Technology*, 56(11), 1167-1176.
- Kim, K-S, & Allen, B. (2002). Cognitive and task influences on web searching behavior. *Journal of the American Society for Information Science and Technology*, 53(2), 109-119.
- Kim, S. and Soergel, D. (In press). Selecting and Measuring Task Characteristics as In

- dependent Variables. *Proceedings of the 68th Annual ASIST meeting*. October 28-November 2. 2005. Charlotte, North Carolina. USA.
- Kremer, R. and Gaines, R.K. (1996). Embedded Interactive Concepts Maps in Web Documents. *Proceedings of WebNet'96: World Conference of The Web Society*. October 16-19. San Francisco, California. USA. Available at: http://pages.cpsc.ucalgary.ca/~kremer/webnet96/webnet_kremer.html
- Liao, C., & Palvia, P.C. (2000). The impact of data models and task complexity on end-user performance: an experimental investigation. *International Journal of Human-Computer Studies*, 52(5), 831-45.
- Nielsen, M L (2004). Task-based evaluation of associative thesaurus in real-life environment. *Proceedings of the ASIST 2004 Annual Meeting: Managing and Enhancing Information: Cultures and Conflicts*. Providence, Rhode Island, November 13 - 18, 2004.
- Oh et al. (In press).. Ontology-Driven Knowledge Organization Enhancing UDDI Web services in Korea using Topic Maps, *Proceedings of the 68th Annual ASIST meeting*. October 28-November 2. 2005. Charlotte, North Carolina. USA.
- Pepper, S. 2002. The tao of topic maps. Oslo: Ontopia. Available at: <http://www.ontopia.net/topicmaps/materials/tao.html>
- Pepper, S. and L, Garshol. (2002). The XML Papers: Lessons Learned on Applying Topic Maps. XML 2002 Conference and Exposition. December 8-13. 2002. Baltimore, Maryland. USA.
- Rowley, J. (1994). The Controlled Versus Natural Indexing Languages Debate Revisited: a Perspective on Information Retrieval Practice and Research. *Journal of Information Science*. 20 (2): 108-119.
- Sure, Y., & Iosif, V. (2002). First results of a semantic web technologies evaluation. In R. Meersman et al. (Eds.), *Proceedings of the Common Industry Program Held in Conjunction with Confederated International Conferences: On the Move to Meaningful Internet Systems (CoopIS, DOA, and ODBASE 2002)*, Irvine, CA: University of California, 69-78.
- Vakkari, P. (1999). Task complexity, problem structure and information actions integr

ating studies on information seeking and retrieval. *Information Processing and Management*. 35. 819-837.

ISO/IEC 13250: 2002.

Available at : http://y12web2.y12.doe.gov/sgml/sc34/document/0322_files/iso13250-2nd-ed-v2.pdf

Farquhar, A., Fikes, R., Rice, J.1997. "The Ontolingua Server: A tool for collaborative ontology construction", *International Journal of Human-Computer Studies*, 46 (6): pp.707727

Fikes, R., Farquhar, A., 1999. "Distributed repositories of highly expressive reusable ontologies", *IEEE Intelligent Systems & Their Applications*, 14(2) Mar-Apr, pp. 73-79.

Gruber, T.R. 1993. "A translation approach to portable ontologies", *Knowledge Acquisition*, 5(2), pp. 199-220.

Guarino, N. 1995. "Formal Ontology, Conceptual Analysis and Knowledge Representation", *International Journal of Human and Computer Studies*, 43(5/6), pp. 625-640.

Sowa, J. 2000. "Ontology, Metadata, and Semiotics", *International Conference on Conceptual Structures, ICCS'2000*, pp. 4-18 August 2000, Darmstadt, Germany

Uschold, M., Grüninger M., 1996. "ONTOLOGIES: Principles, Methods and Applications" *Knowledge Engineering review* 11(2), pp. 93-155.

Van Heijst, G, Falasconi, S., Abu-Hanna, A, Schrabar, G., Stefanelli, M. 1995. "A Case Study in Ontology Library Construction". *Artificial Intelligence in Medicine* 7, pp. 227-255.

Sowa, J. F. 2000. "*Knowledge representation Logical, Philosophical, and Computational Foundations*". Brooks/Cole, Pacific Groove, CA, USA

Staab, S., Schnurr, H. P., Studer, R., Sure Y. 2001. "Knowledge Processes and Ontologies." *IEEE Intelligent Systems*. 16(1) pp. 26-34.

Guarino, N. 1998. "Formal Ontology and Information systems". *Proceedings of FOIS'98*, Amsterdam, IOS Press.

- Vasconcelos, J. Gouveia, F. Kimble C. 2002. "An Organizational Memory Information System using Ontologies". *Proceedings of the 3rd Conference of the Associação Portuguesa de Sistemas de Informação*. University of Coimbra, Portugal, November
- Fonseca, F.T., Egenhofer, M.J., Davis, C., Camara, G. 2002. "Semantic granularity on ontology-driven geographic information systems", *Annals of Mathematics and Artificial Intelligence*. 36(1):121-134
- Jennings, N.R, Wooldridge M. 1995 "Applying Agent Technology, Applied Artificial Intelligence", *An International Journal*, Taylor & Francis London, 9 (4) pp.51-361
- Wooldridge, M., Jennings, N. R. 1995. "Intelligent Agents: Theory and Practice." *Knowledge Engineering Review* 10(2) pp.115-161.
- Zambonelli, F., Jennings, N. R., Omicini, A., Wooldridge, M. 2000. "Coordination of Internet Agents: Models", *Technologies and Applications*. Springer.
- Cabri G., Leonardi, L. Zambonelli F. 2000. "Mobile-agent Coordination Models for Internet Applications", *IEEE Computer Magazine*, Vol. 33, No. 2, Feb.
- Barbuceanu, M. and Fox, M.S., 1994 "The Information Agent: An infrastructure agent supporting collaborative enterprise architectures", In *Third Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*,
- Jennings, R. J. 2000. On agent-based software engineering. *Artificial Intelligence* 117: 277-296
- Ding, Y. 2001. "A review of ontologies with the Semantic Web in view", *Journal of Information Science*, 27(6): p. 377-388.
- Fensel, D., Harmelen, F., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F. 2001. "OIL: An Ontology Infrastructure for Semantic Web", *IEEE Intelligent Systems*, March/April. pp. 38-45.
- Guarino, N., Masolo, C., Vetere, G. 1999. "OntoSeek: Content-Based Access to the Web." *IEEE Intelligent Systems & Their Applications* 14(3): 70-80.
- Harmelen, F. 2002. "How the Semantic Web will change KR: challenges and opportunities for a new research agenda". *The Knowledge Engineering Review*, 2002. 17(1).
- Staab, S., Maedche, A. 2001. "Knowledge portals --- ontologies at work", *AI Magazine*,

2001. 21(2).
- Ding, Y., Fensel, D., Klein, M., Omelayenko, b. 2002. "The Semantic Web: Yet Another Hip?" - to appear in *Data and Knowledge Engineering*, 2002, 18.12.01
- Staab, S., Maedche, A., Handschuh, S. 2001. An annotation framework for the semantic web. In *Proceedings of the First Workshop on Multimedia Annotation*, Tokyo, Japan.
- Stojanovic, N., Maedche, A., Staab, S., Studer, R., Sure, Y. 2001. SEAL: a framework for developing SEMantic PortALs. In *Proceedings of K-CAP 2001*.
- Fox, M. S., Gruninger, M. 1999. "Ontologies for Enterprise Integration". Department of Industrial Engineering,, University of Ontario.
- Malone, T. W., Crowston, K., Lee, J., Pentland, B., Dellarocas, C., Wyner, G., Quimby, J., Osborn, C. S., Bernstein, A., Hermen, G., Klein, M., O'Donnel, E. 1999. "Tools for Inventing organizations: Towards a Handbook of Organizational Processes", *Management Science*, 45(3), pp. 425-443
- Ushold, M., King, M., Moralee, S., Zorgios, Y. 1998. "The Enterprise Ontology." *The Knowledge Engineering Review*, 13 (Special Issue on Putting Ontologies to Use).
- Gruninger, M., Fox, M. S. 1996. The Logic of Enterprise Modelling. In *Modelling and Methodologies for Enterprise Integration*, pages 140-157.
- Gruninger, M., Atefi, K., Fox, M. S., 2000. "Ontologies to Support Process Integration in Enterprise Engineering", *Computational & Mathematical Organization Theory* 6, 381-394, 2000.
- Kim, H. M., Fox, M. S., Grüninger, M. 2001. "Construction of Ontology Based Enterprise Models for E-Commerce", *Work Report*, Schulich School of Business, York University, 4700 Keele St., Toronto, Ontario Canada, M3J 1P3.
- Fox, M., Barbuceanu, M., Teigen, R. 2000. "Agent-Oriented Supply-Chain Management", *The International Journal of Flexible Manufacturing Systems*, 12, pp. 165-188.
- Swaminathan, J. M., Smith, S. F., Sadeh, N. M. 1998. "Modeling Supply Chain Dynamics: A Multiagent Approach" *Decision Sciences*, 29 (3), Summer

[부록] 국내외 학위논문 서비스 및 KERIS 온톨로지 기반 서비스 사례

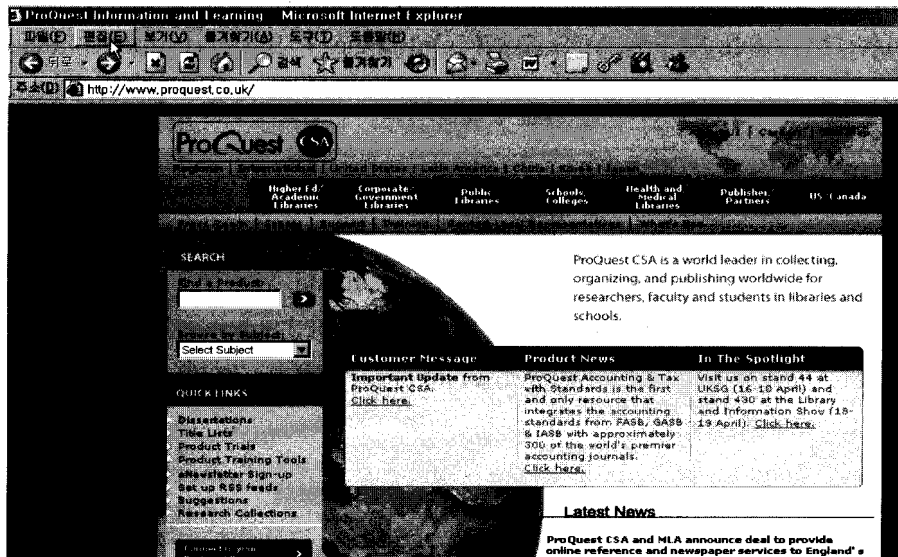
1. 국외 학위논문 서비스 사례

1.1 ProQuest Company & UMI

1) 개요

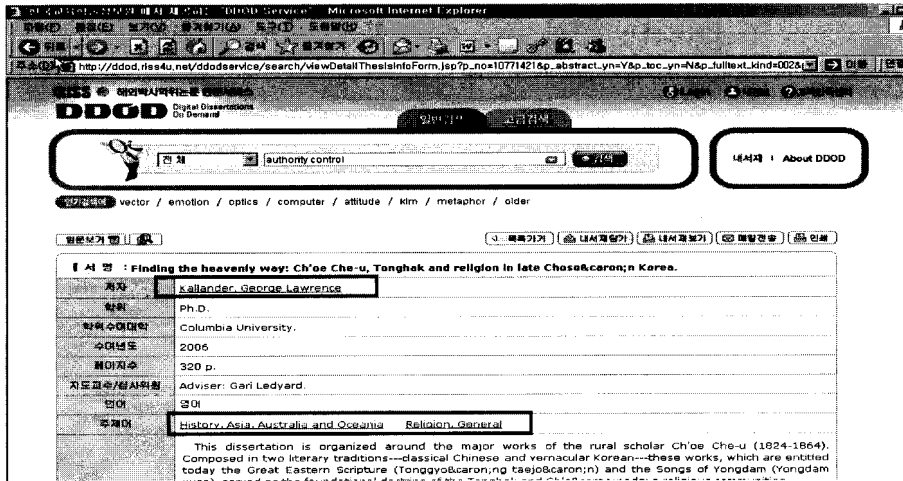
- ProQuest Information and Learning은 1985년 학위논문 데이터를 보유하고 있는 UMI를 인수, ProQuest CSA(<http://www.proquest.com/>)를 통해서 학위논문 서비스 실시
- ProQuest Dissertations & Theses database (PQDT)는 전세계 최대 규모의 석박사 논문데이터베이스
- ProQuest가 보유하고 있는 석박사 학위논문 2백3십만건 중 2백만건 이상은 인쇄물, 마이크로폼, 전자형태로 이용 가능
- MIT를 제외한 모든 북미대학 및 유럽지역의 논문을 포함하고 있고, 매년 8만건 이상 업데이트

2) ProQuest CSA 홈페이지

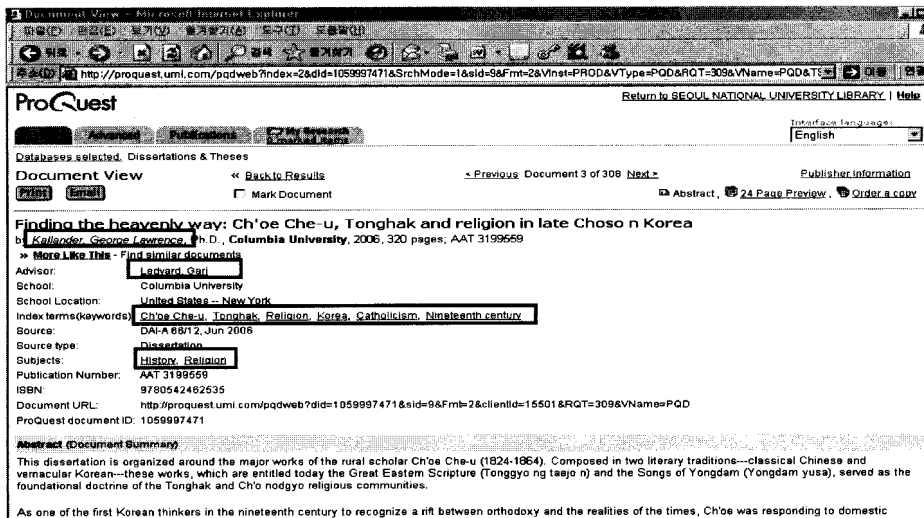


3) KERIS(한국교육학술정보원)에서 제공하는 DDOD 검색 화면

- DDOD(Digital Dissertations on Demand) 서비스는 KERIS에서 제공하는 해외박사학위논문서비스에 수록된 분야별 상위권 대학의 우수 박사학위 논문 원문(PDF)을 손쉽게 검색 및 다운로드할 수 있도록 UI를 한글에 맞게 구성하여 서비스
- 재검색 : 저자, 주제어만 가능



4) ProQuest에서 직접 학위논문 검색



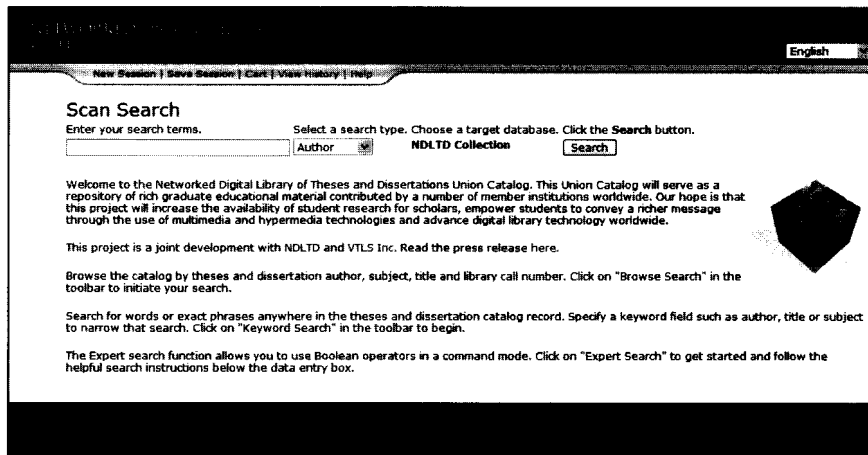
- 재검색 : 저자, 지도교수, 핵심어(keyword), 주제 가능
- KERIS 제공 UI는 재검색 기능이 축소되어 있음

1.2 NDLTD

(Networked Digital Library of Theses and Dissertations)

1) 개요

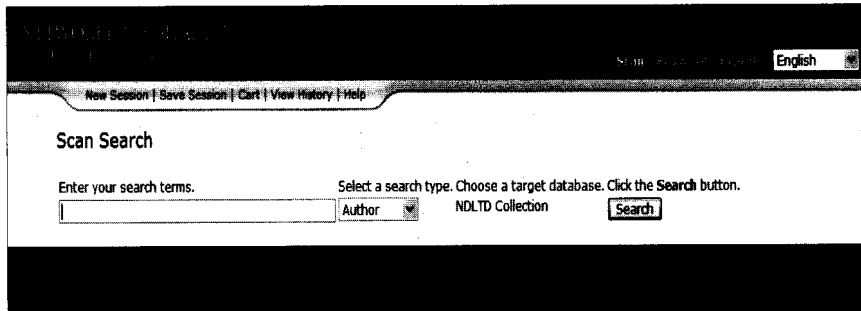
- 전자학위논문(ETD) 생성 및 공동 활용 체제 구축을 위한 협의회인 NDLTD는 전 세계 160여 개 대학과 24개 기관 참여
- ETD 시스템 개발 및 보급, ETD 유통시스템 개발 및 공유 환경 구축, OAI 표준 프로토콜을 활용한 종합목록DB 구축 등을 수행
- 학위논문 원문구축 현황 : 232,953건 ('06.10 현재)



2) 검색방법

○ Scan Search

- 검색조건 : 저자, 주제, 제목, 청구기호

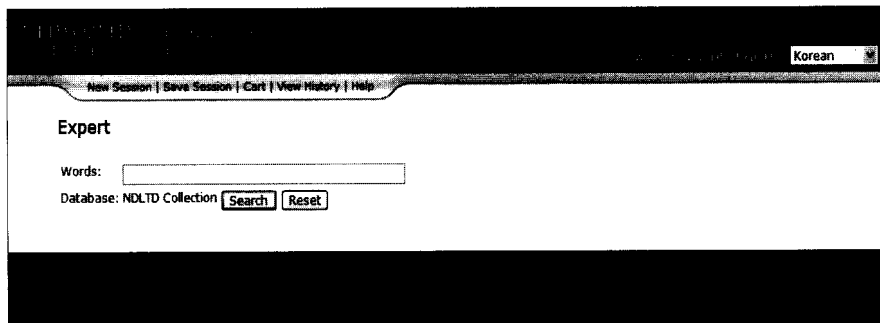


○ Keyword Search

- 검색조건 : 전체(Anywhere), 저자(Author), 주제(Subject), 제목(Title), Committee Member, 발행자/기관(Publisher/Institution)

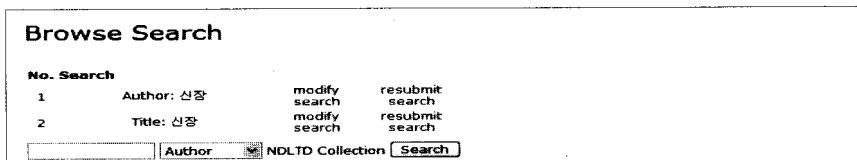
○ Expert Search

- 논리연산 명령문을 사용
- 검색조건 : 전체(Anywhere), 저자(Author), 주제(Subject), 제목(Title), Committee Member, 발행자/기관(Publisher/Institution)



3) 검색특징

- 검색시 11개의 언어 지원
- 검색히스토리 보기 메뉴가 있어 이전 검색항목을 조회할 수 있음



4) 검색결과

○ Scan Search 검색결과

- 학위논문의 제목만 Scan
- 해당 목록의 알파벳 앞뒤 순서의 제목 Scan

First Page | Previous Page | Next Page | Last Page

You searched NDLTD Collection - Title: patient

| Hit Count | Scan Term |
|-----------|---|
| 1 | Patience and urgency in biblical conversion |
| 1 | Patient admission characteristics influencing discharge destinations from a Geriatric Medicine In-Patient Unit |
| 1 | Patient age, number and type of clinical encounters, and provider advice to quit smoking |
| 1 | Patient age, number and type of clinical encounters, and provider advice to quit smoking, BRFSS 2000 |
| 1 | Patient and provider determinants of the use of ambulatory physician services for chronic illness episodes in an adult population in Quebec |
| 1 | Patient-based outcomes, older adults' perceptions of hospital and recovery experiences |
| 1 | Patient care system: proposal of a software - prototype. |
| 1 | Patient-centered advance care planning in open-heart surgical patients / |
| 1 | Patient evaluation of using marketing elements at Outpatient Polyclinic |
| 1 | Patient-Image Registration using A-mode ultrasound localization of features |

First Page | Previous Page | Next Page | Last Page

Scan Search

Enter your search terms. Select a search type. Choose a target database. Click the Search button.

Select below to return to the last:

○ Keyword Search 검색결과

- 저자(Author), Title(제목), Publisher/Institution(발행기관) 제공

NEWSPIDER ONLINE

English

New Session | Save Session | Cart | View History | Help

next set | last set

Sort by: None

Ascending Descending

Records 1 through 10 of 49

| | | |
|------------------------------|-----------------------|--|
| 1 | Author | KAPATZAS, ΓΑΒΡΙΗΛ Μ. |
| full | Title | ΜΕΘΟΔΟΣ ΣΥΝΤΗΡΗΣΕΩΣ ΝΕΦΡΙΚΩΝ ΜΟΣΧΕΥΜΑΤΩΝ ΕΠΙ 24 ΩΡΑΣ (ΠΕΙΡΑΜΑΤΙΚΗ ΕΡΓΑΣΙΑ) |
| marc | Publisher/Institution | ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ ΤΜΗΜΑ ΙΑΤΡΙΚΗΣ. |
| Select to add record to cart | | |
| <input type="checkbox"/> | | |
| 2 | Author | Tang, Chi-wai, Sydney. |
| full | Title | Human kidney as an organ of complement synthesis : its regulation by tubular protein / Tang Chi Wai, Sydney. |
| marc | Publisher/Institution | University of Hong Kong |
| Select to add record to cart | | |
| <input type="checkbox"/> | | |

○ 상세 검색 결과

- 저자(Author), Title(제목), Subject(주제), Summary Note(요약),
Dissertation note, Contributor, Publication(발행), Institution, 원문
URL(Multimedia)
- 꼬리물기 검색 지원
저자(Author), Title(제목), Subject(주제), Publication(기관)

| | |
|--|--|
| You searched NDLTD Collection - Title: Forms of Trypanosoma cruzi in mouse kidney during chronic infection and analyses of infected kidney antigen | |
| Author | Huang, Ya-Chi |
| Title | Forms of Trypanosoma cruzi in mouse kidney during chronic infection and analyses of infected kidney antigen |
| Publication | NSYSU 2002-08-27 |
| Summary Note | Chagas' disease induced by Trypanosoma cruzi infection causes damages in heart and digestive tract of the host, but trypanosomes were rarely found in blood and tissue during chronic infection. Although trypanosomes could not be found in the peripheral blood of Trypanosoma muscui infected mice during chronic stage, trypanosomes detected in vasa recta of kidney became a unique kidney form (KF). In this investigation, BALB/c mouse was used as animal model to examine if kidney form of T. cruzi exists in chronic stage. Kidney sections were made on the 20, 40, 60, and 80th day after infection of each mouse with 1,000 bloodstream trypanosomes. In order to compare with chronic stage, kidney sections obtained from mice 20 and 40 days after infection were regarded as acute phase. Amastigotes of T. cruzi as well as significant tissue damages were found in both acute and chronic stages. The results indicated that T. cruzi may exist in amastigote form in mouse kidney during chronic infection. ELISA was employed to measure mouse serum specific antibody titers which recognize mouse kidney antigen infected with trypanosomes 20 to 300 days respectively. Peak titer was observed on the 40th day of infection. Results showed that the ability of acute serum recognizing infected kidney antigen was better than chronic serum. On the other hand, antibody titers that recognize amastigote antigens increased gradually during the infection. It seems correlated with trypanosomes existing and releasing antigen during chronic infection. Antisera from rat immunized with epimastigote or amastigote could recognize infected and normal mouse kidney antigen. Anti-epimastigote sera detected 27 and 49 kDa proteins and anti-amastigote sera detected 28 and 48 kDa proteins of infected mouse kidney antigen. Functions of these molecules are still not clear and require further characterization in the future. |
| Dissertation note | Master (none) -- NSYSU. |
| Contributor | Chao, David |
| Contributor | Shin, Jyh-Wei |
| Contributor | Chang, Gan-Nan |
| Institution | NSYSU Biological Sciences |
| Subject | T. muscui |
| Subject | amastigote |
| Subject | T. cruzi |
| Subject | KF |
| Multimedia | Click to link to URL application/pdf http://etd.lib.nsysu.edu.tw/ETD-db/ETD-search/view_etd?URN=etd-0827102-163807 |

- 검색결과물을 [Save to Cart]를 통하여 세션에 저장할 수 있음
- MARC 제공

5) 기타

- Session 방식의 개인 환경 설정

You may load or save a former searching session or current searching session from here. Please be advised that the duration of time that your session remains on the server is at the discretion of the sys admin.

Current Session ID 2007041921513817075

Enter the session ID in the box below and click the Load Session button.

Load Session

End & Save Session

OR

Return with New Session ID

- 개인 환경 설정을 로그인 방식이 아니라 Session방식으로 개인 Cart와 검색
- History를 저장 하여, 후에 다시 홈페이지를 접속하였을 때 개인 Cart와 검색 History 보기 가능

1.3 MIT

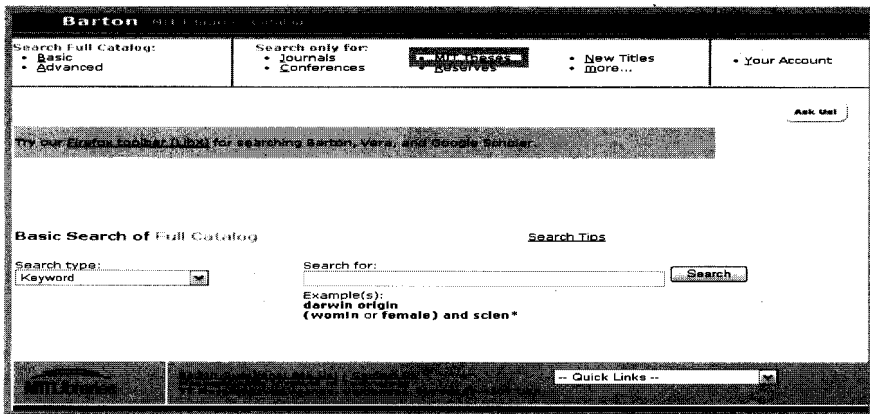
1) 검색방법

- MIT 학위논문, 비 MIT 학위논문을 구분하여 MIT 학위논문만을 검색

| | |
|--------------|----------------------|
| MIT 학위논문 | Catalog(Barton) |
| MIT 학위논문(전자) | Dspace |
| 비 MIT 학위논문 | 외부기관을 통한 검색/구매 방법 안내 |

- MIT 학위논문

- MIT 도서관 Catalog인 Barton을 통하여 검색
- 전체검색 상세검색에서 학위논문(MIT Theses) 유형 제한을 제공
- 학위논문 검색 페이지가 별도로 존재, 간략검색과 상세검색을 별도로 제공하지 않고 한 화면에서 Basic Search와 Search by Multiple Terms 제공



Search for MIT Theses only

See also: [MIT Theses in DSpace](#)

Basic Search

Search type: Search for:

Search by Multiple Terms

Search type: Search for:

Limit search to:

Thesis Department:
Thesis Degree:
Year from: Year to: Location:

*NOTE: The Thesis Department Keyword and Thesis Degree Keyword search uses abbreviations for the different departments and degrees. Consult [Department Name and Degree Abbreviations](#) for help, or use the menus in the "Limit search to" section.

Brief Results Display from MIT Theses only

Results for W-all keywords= library; sorted by : Year
Records 1 - 10 of 5992

[Add to Bookshelf](#) [Save/Email](#) [View Selected](#)
[Limit by Date](#) [Limit to Available Items](#)

[Select All](#) [Deselect](#) [Search within results](#)

Sort options : Year Year(ascending) Author Title
Change Display : Table View List View

- 1 [\[Display full record \]](#)
Author [Alper, Hal \(Hal Samuel\)](#)
Title [Development of systematic and combinatorial approaches for the metabolic engineering of microorganisms / by Hal Alper.](#)
Published c2006.
Format Book
Online Ed. URL <http://dspace.mit.edu/handle/1721.1/95134>
Availability Click [All items](#) to check current status
Location [Institute Archives - Noncirculating Collection 9 | THESIS Thesis Chem Eng 2006 Ph.D.](#)
Location [Hayden Library - Stacks | Science THESIS Thesis Chem Eng 2006 Ph.D.](#)
- 2 [\[Display full record \]](#)
Author [Antipova, Alena A.](#)
Title [Gene expression-based screening of inhibitors of signal transduction / by Alena A. Antipova.](#)
Published c2006.
Format Book
Online Ed. URL <http://dspace.mit.edu/handle/1721.1/94500>
Availability Click [All items](#) to check current status
Location [Institute Archives - Noncirculating Collection 9 | THESIS Thesis Chem 2006 Ph.D.](#)
Location [Hayden Library - Stacks | Science THESIS Thesis Chem 2006 Ph.D.](#)

Full Record

[Results List](#) | [Add to Bookshelf](#) | [Save/Email](#)

Choose format: [Standard](#) | [Citation](#) | [MARC tags](#)

Record 2 out of 440

Author Alper, Hal (Hal Samuel)

Title Development of systematic and combinatorial approaches for the metabolic engineering of microorganisms / by Hal Alper.

Online Ed. URL <http://dspace.mit.edu/handle/1721.1/35134>

Availability [All items](#)

Location [Institute Archives - Noncirculating Collection 3 | THESIS Thesis Chem Eng 2006 Ph.D.](#)

Location [Hayden Library - Stacks | Science THESIS Thesis Chem Eng 2006 Ph.D.](#)

Published c2006.

Description 261 p. : ill. (some col.) ; 29 cm.

Format Book

Thesis Note Thesis (Ph. D.)--Massachusetts Institute of Technology, Dept. of Chemical Engineering, 2006.

Thesis Supervisor Supervised by Gregory Stephanopoulos.

Bibliography Includes bibliographical references (p. 243-261).

Abstract Explorations and optimizations through the genomic space are a daunting undertaking given the complexity and size of the possible search space. To approach this problem, systematic and combinatorial approaches were employed for the engineering of cellular

- 꼬리물기 검색 : 저자, 서명

○ MIT 학위논문(전자자료)

- Dspace를 통하여 검색 (<http://dspace.mit.edu/handle/1721.1/7582>)
- MIT의 일부 석·박사 학위논문만을 포함

2) 검색항목

○ MIT 학위논문

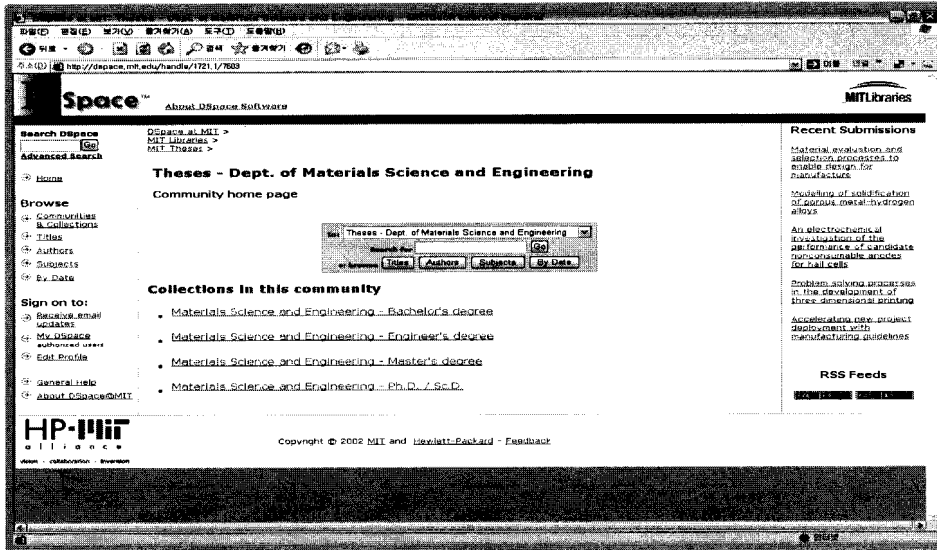
- 기본검색 : 키워드, 지도교수 키워드, 서명(우측절단), 서명(키워드), 저자, -1986년 이전에 한하여 - 주제(우측절단), 주제(키워드)
- 용어 검색 : 전체, 저자, 서명, 지도교수 키워드, 학과 키워드, 학위 키워드, 주기사항
- 제한항목 : 전공학과별, 학위명별, 수여연도별, 소장위치별
- 재검색항목 : 저자, 서명, 주제, 다른 저자(other author)

○ MIT 학위논문(전자자료)

- 소속 커뮤니티별 검색 가능 : 커뮤니티는 분야별로 각각 생성됨
- 서명, 저자, 주제, 날짜별로 브라우징 가능
 - 서명, 저자, 주제 브라우징시 ABC 리스트 제공
 - 날짜 검색시 년, 월 지정 가능

3) 검색특징

- 전공학과와 학위명은 약어를 사용(별도리스트 제공)
- Dspace의 커뮤니티를 이용하여 학위별 검색 제공



4) 원문, 부가정보 연계

- 원문 : 자관학위논문의 경우 Dspace와 연계하여 Online Ed. 제공
 - 학외 공개 : 원문에 대한 Preview, 열람가능, 저장가능, 출력불가
 - 학내 공개 : 원문에 대한 열람가능, 저장가능, 출력가능
 - 접속장소에 따라 권한 달라짐

URI: <http://hdl.handle.net/1721.1/33405>
 Appears in Collections: [Materials Science and Engineering - Ph.D. / Sc.D.](#)
[Materials Science and Engineering - Ph.D. / Sc.D.](#)

| Files in This Item: | | | |
|---------------------|--------------------------------------|--------|-------------------------------------|
| File | Description | Size | Format |
| 62712948.pdf | Preview, non-printable (open to all) | 9709Kb | Adobe PDF View/Open |
| 62712948-MIT.pdf | Full printable version (MIT only) | 9714Kb | Adobe PDF View/Open |

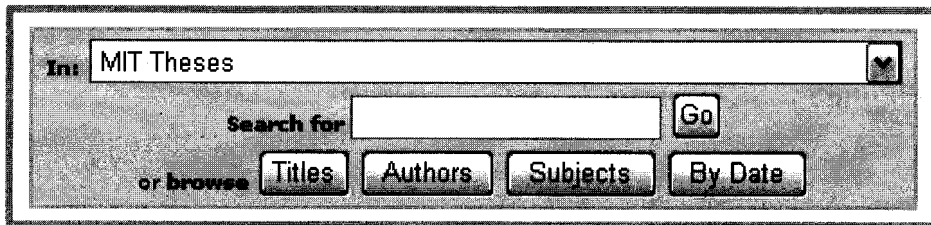
- 연계정보 : 초록

5) 타 서비스 및 시스템과의 연계

- 자관학위논문의 경우 Dspace와 연계하여 Online Ed. 제공

6) DSpace

- 개요 : MIT와 휴렛팩커드사에 의해 개발된 디지털 연구를 위한 보관소
- 14,000건 이상의 MIT 학위논문 원문을 MIT Dspace를 통해 서비스
- 커뮤니티 단위로 구성
 - 학위논문 커뮤니티
 - 전공분야별 커뮤니티
 - 학위구분별 커뮤니티(Bachelor, Engineer, Master, Ph.D. / Sc.D.)
- 각 커뮤니티별 직접검색, 브라우징 제공



- 상세서지사항
 - 서명, 저자, 지도교수, 소속(Department), 공저자(other contributors), 키워드, 출판년, 출판사, 초록, Description, URI, Appears in Collections, 원문파일(PDF)
 - URI : 해당 서지에 접근할 수 있는 고유 URL
 - Appears in Collections : 해당 서지가 소속되어 있는 커뮤니티 (동일한 커뮤니티명이라 할지라도 탐색 방법에 따라 가지는 고유 정보가 상이함)
- 해당 컬렉션에 대한 신착자료 정보를 매일 e-mail로 제공하는 서비스
- 신분에 따라 제공 원문이 다름
 - MIT 소속 : 온라인에서 프린트 가능한 원문 다운로드 가능
 - 비 MIT 소속 : 온라인에서 열람만 가능한 원문 다운로드 가능
프린트 가능한 원문PDF나 복사본은 구매를 통해 이용

1.4 UC Berkeley

1) 검색방법

○ Pathfinder

- UC버클리대학 내 20여개 도서관의 자료 검색
- Quick Search, Advanced Search, Browse 제공
- 학위논문에 대한 별도 유형은 구분하지 않고 Manuscript로 구분함

※ 학위논문만 검색하고자 할 때

- Advanced Search

- 검색 항목 : 'Note Table do Contents Keyword'로 두고 아래 항목을 키워드로 입력하여 검색
- 학위 (bachelor, MA, MS, PDH)
- 출판 유형 (Thesis, Professional report 등)
- 주제분야 (architecture, landscape architecture, city planning 등)
- 학위수여연도
- 입력 예) achelor architecture berkeley 1963
City Planning berkeley 1974
Thesis architecture berkeley 1994
Thesis PHD architecture berkeley 1985
professional report landscape berkeley 1999
professional landscape berkeley 1982
urban design berkeley 1995
ms arch* 1990-2004

- Browse Search

검색 항목 : 'subject begin with로 두고 기관명으로 검색

- 입력 예) architecture, city and regional planning, landscape

UCB LIBRARY Pathfinder LIBRARY CATALOGUE ONLINE

Quick Search **Advanced Search** Browse Previous Searches Show Saved Contact Us Help Exit

Advanced Search Search for in Notes/Table of Contents keyword(s) **SEARCH**

You are searching: **UCB Library Catalog** and in Author (last name first)

Search for: Limit by Location and ALL locations

Recent Acquisitions: Limit by Year: and

Choose Database: Limit by Language: and ALL languages More Choices

Limit by Publication Format: and ALL publication formats

Limit by Date(s) Added: yymmdd-yyymmdd and

Automatic Sort for search result count less than 3000 records:

Date Ascending Descending

Author Ascending Descending

Title Ascending Descending

Tips: Advanced Searches

Combining terms
To combine two searches at one time, enter a different search statement in each of the two boxes above (for example, title keywords and an author's name). You may also combine terms within a box:

- 간략서지

- 기본출력 : 서명, 저자, 소장위치, 전자자원위치, 통일서명, 포맷, 출판사항
- CHANGE DISPLAY : 간략 검색 결과 디스플레이 항목 이용자가 개별적으로 편집 가능
- AVAILABILITY : 소장정보 제공
- 원문 정보 포함 안함
- 저장방식 : 목록 저장 및 출력

- 상세서지

- 저자, 주제 꼬리물기 검색 지원
- 원문 아이콘 제공하지 않으며, UCB

UCB Library Catalog (1 of 1)

Title Predictors of peer rejection, acceptance, and victimization among girls with and without ADHD / by Dara Renee Blachman.
Author Blachman, Dara Renee.
Date 2004.
Description v, 91 leaves : ill. ; 28 cm.
Notes Conferred date of degree is December 2003.
Thesis (Ph.D. in Psychology)--University of California, Berkeley, Spring 2004.
Includes bibliographical references (leaves 58-70).
Subject University of California, Berkeley. -- Dept. of Psychology -- Dissertations.
Headings

Location(s): Main Stack 308t 2004 165
Loan period: non-circulating; Not checked out.

UCB Library Catalog (1 of 1)

No Records Currently Saved

○ Melvyl

- 10개 UC 캠퍼스내 도서관 장서 및 캘리포니아 주 소속 전문장서 등 포함
- Quick Search, Advanced Search, Command, Browse 제공
- 검색 포맷에 Dissertation을 두어 유형으로 구분하고 있음
- 학위논문은 주제표목을 갖지 않아 주제론 검색 불가능

※ 학위논문만 검색하고자 할 때

- 검색항목 : Keyword로 두고 저자, 서명, 주기 검색
- 검색방법 : Title에서 3개의 키워드를 조합 검색

Melvyl

Advanced Search: Entire Collection [Change Collection](#)

Search:

and

and

Words as Phrase

No Yes

No Yes

No Yes

View Search Results:

Optional Limits:

Library:

Format:

Language:

Year:

From (yyyy) To (yyy)

Tip: Want more or fewer results? Edit your search directly on this page.
[More Tips and Search Examples](#)

- 간략서지

- 저자에 대한 꼬리물기 검색 제공
- 디스플레이 항목 : 저자, 서명, 출판사, 포맷, 소장도서관
- 정렬 : 년도, 저자, 서명, 통일서명 (최초 1000건의 데이터를 정렬함)

Search results: 1 item(s) [Modify Search](#)

Searched: **Keywords= Predictors of peer rejection and Library= UCB or NRLF or GTU and Format= DS** **Display: Short Long Review**
 Sort: [Year](#) [Author](#) [Title](#) [Uniform Title](#)
 Sorted by:
 Collection: Entire Collection **Sorting and display limited to first 1,000 records**

[Print/Email](#) [Save](#) [Request](#) [Previous](#) [Next](#) Item # [Go](#)
[select all](#) [deselect all](#)

1 [Details/locations](#) **Author** [Blachman, Dara Renee.](#)
Title Predictors of peer rejection, acceptance, and victimization among girls with and without ADHD / by Dara Renee Blachman.
Publisher 2004
Format Dissertation
 Archive/Manuscript
Library UCB

[Print/Email](#) [Save](#) [Request](#) [Previous](#) [Next](#)
[select all](#) [deselect all](#)

- 상세서지

- 저자, 주제 꼬리물기 검색 가능
- 재검색 기능 지원
- MARC 보기 지원 및 상호대차, 원문배달 서비스 연계

Search results: 1 item(s) [Modify Search](#) **Display: Full MARC**

[Print/Email](#) [Save](#) [Request](#) [Previous](#) [Next](#)

Item 1 of 1 Total
[Return to Search Results List](#)

Author [Blachman, Dara Renee.](#)
Title [Predictors of peer rejection, acceptance, and victimization among girls with and without ADHD / by Dara Renee Blachman.](#)
Publisher 2004.
Description v, 91 leaves : ill. ; 28 cm.
Note Conferred date of degree is December 2003
Note Thesis (Ph.D. in Psychology)--University of California, Berkeley, Spring 2004.
Note Includes bibliographical references (leaves 98-70).
Language English
Subject [University of California, Berkeley, Dept. of Psychology -- Dissertations.](#)
[Dissertations, Academic -- UCB -- Psychology -- 2001-2010.](#)
Format Dissertation
 Archive/Manuscript
Library [UC Berkeley](#) [All](#)

| Library | Call Number | Availability | Notes |
|---------------------|---------------|--------------|-------|
| UC Berkeley Main | 306t 2004 165 | Circ. status | |

[Previous](#) [Next](#)

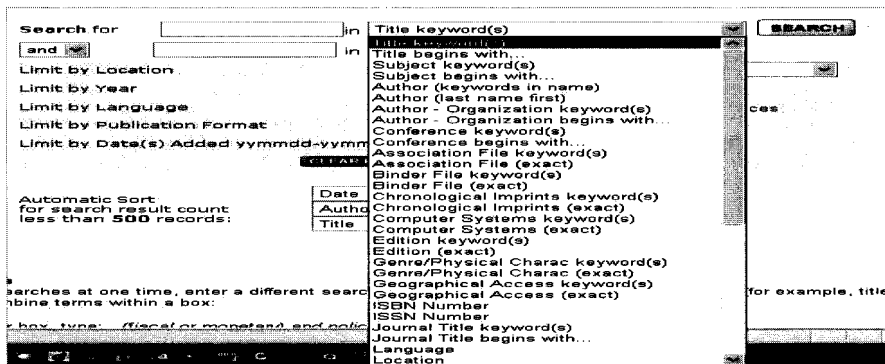
2) 검색항목

○ Pathfinder

- 검색항목 : 서명(키워드), 서명(전방일치), 주제(키워드), 주제(전방일치), 저자(키워드), 저자(성이름순), 저자(단체명 키워드), 저자(단체명 전방일치), 회의명(키워드), 회의명(전방일치),

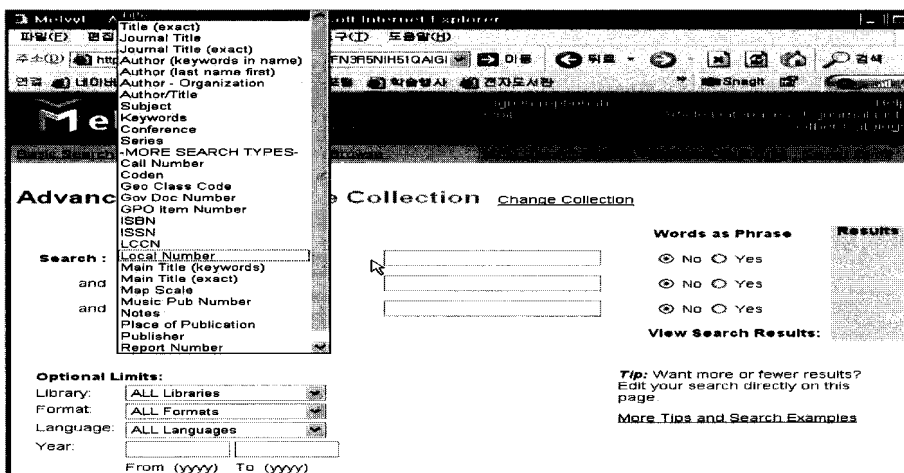
ISSN, ISBN, 언어, 소장위치 등

- 제한항목 : 소장위치, 년도, 언어, 출판형식, 날짜
- 정렬옵션 : 날짜, 저자, 서명, 소장위치, 청구기호, 통일서명에 의한 순차식, 역순차식 검색시 설정하면 500개까지 자동 정렬



○ Melvyl

- 검색항목 : 서명(키워드), 서명(완전일치), 저널명, 저널명(완전일치), 저자(키워드), 저자(성이름순), 저자(단체명), 주제, 키워드, 회의명, 시리즈, 청구기호, ISSN, ISBN, LCCN 등
- 제한항목 : 소장도서관, 년도, 언어, 출판형식



3) 검색특징

- 학위논문 검색을 위한 별도의 시스템 구현 미비.
- 검색 방법 어려움

4) 원문, 타 서비스 및 시스템과의 연계

- 원문을 제공하는 자료에 대해선 Electronic Location 항목에서 URL을 제공하나 학위논문에서 원문을 제공하는 자료 찾을 수 없었음
- 원문 제공에 대해 ProQuest 데이터베이스 연계 제공함

Current Research@

UC users can search and access full-text UC Ph.D. dissertations

- 1996~ 현재 버클리대의 논문 원문 제공. 다운로드 가능
- 학내에서 free접근 가능. (프락시 서버 제공)
- 검색항목: 키워드에서 주제 또는 토픽으로 검색
- PDF파일, 24page 프리뷰 제공

Digital Dissertations

Search for *non-UCB* theses or dissertations on your topic.

- 북미와 유럽 대학의 논문 제공
- 1861~ 현재
- 검색항목: 키워드로 주제 또는 토픽 검색
- PDF파일, 24page 프리뷰 제공

The screenshot shows the ProQuest search interface. At the top, there are navigation tabs for 'Advanced', 'Topics', 'Publications', and 'Current Research@'. Below this, it says 'Databases selected: Multiple databases'. The 'Basic Search' section has a search input field and buttons for 'Search' and 'Clear'. Below the search field, there are several filters: 'Database' (Multiple databases), 'Date range' (All dates), and 'Limit results to' (Full text documents only, Scholarly journals, including peer-reviewed). There is also a 'More Search Options' section with expandable fields for 'Publication title', 'Author', 'Look for terms in', 'Document type', 'Publication type', and 'Exclude from results' (Book Reviews, Dissertations, Newspapers). On the right side of the 'More Search Options' section, there are links for 'Browse publications' and 'About'.

<Current Research@화면>


| | | | | | |
|---|----------------------|---|--------------------------------|---------------------------------------|--|
| SEARCH | | DIGITAL DISSERTATIONS | | LIMIT | |
| BROWSE | | RESULTS LIST MARKED LIST SHOPPING CART DOWNLOAD CONTACT | | HELP PRICING | |
| Search - basic | | | BASIC ADVANCED | | |
| You are searching all dissertations in the Dissertation Abstracts database with a degree date of 2006 or 2007. | | | | | |
| Enter up to three search terms in the boxes below. View your search history . | | | | | |
| and | <input type="text"/> | in | Keyword (KEY) | <input type="button" value="Search"/> | |
| and | <input type="text"/> | in | Choose a Field | | |
| and | <input type="text"/> | in | Choose a Field | | |
| Degree Date from <input type="text"/> to <input type="text"/> (e. g. 1985) | | | | | |
| Important Migration Message | | | | | |
| On July 21, 2007, ProQuest Digital Dissertations will be going away! | | | | | |
| If your library is a Digital Dissertations subscriber, use ProQuest Dissertations & Theses (PQDT) on the ProQuest platform instead. Ask your library staff how to access the ProQuest platform. | | | | | |
| If your library doesn't subscribe, use Dissertation Express instead. You can order unbound print copies of all available dissertations. | | | | | |

<Digital Dissertations 화면>

1.5 Harvard Univ.

1) 검색방법

- Search : 검색 유형만 지원
(키워드, 서명 우측절단, 서명 키워드, 저자, 저자키워드 등)
- Expanded Search : 키워드의 종류, 소장위치, 언어, 출판년도, 자료유형으로 제한 검색 지원
- Command Search : 키워드의 종류, 소장위치, 언어, 출판년도, 자료유형
- 학위논문 검색 방법 : 자료유형 중 학위논문 제공하지 않음

| | | | |
|---|--|--|--|
| Harvard University | | HOLLIS CATALOG | |
|  | | <ul style="list-style-type: none"> Search Expanded Search Command Search | <ul style="list-style-type: none"> Results List Previous Searches Display Options |
| My Account/Signin Comments/Requests Start Over Help | | | |
| Full Catalog Find Articles Journal Titles Digital Resources Reserves | | | |
| Expanded Search Select keywords from: <input type="text"/> Search for: <input type="text"/> | | Search on: <input type="radio"/> words <input type="radio"/> two-word phrase <input type="radio"/> words <input type="radio"/> two-word phrase <input type="radio"/> words | |
| <input type="text"/> Anywhere <input type="text"/> <input type="text"/> and <input type="text"/> Anywhere <input type="text"/> <input type="text"/> and <input type="text"/> Anywhere <input type="text"/> <input type="text"/> | | <input type="button" value="Search"/> or <input type="button" value="Clear"/> | |
| Limit Search for: Language: <input type="text"/> | | Locations: <input type="text"/> | |
| English Abkhaz Achinese | | Networked Resource Afro-American Studies Andover-Hav. Theol | |
| Format: <input type="text"/> all formats | | Year Range: <input type="text"/> From: <input type="text"/> Through: <input type="text"/> | |
| New! Download LibX search tool | | About the HOLLIS Catalog | |
| <small>Copyright © 2006 ProQuest and Fellow of Harvard College</small> | | | |

- 검색결과 (간략 및 상세서지)

- 자료유형 중 학위논문이 별도로 존재하지 않음
- 검색결과에 학위논문인 경우 해당 서명이 아닌 학위수여연도의 학위논문이라는 것만 검색되어짐
(MARC 입력 내용을 보면 서명(245태그)에 Ph.D Theses 2005 Novemver. 라는 형식으로 입력되며, 해당 연도에 수여된 학위논문은 모두 저자사항(700태그)의 개별저자(\$t)로 처리되고 있음)
- 특정 학위논문에 대한 개별적인 서지사항은 입력되고 있지 않음

[Home](#) [About HOLLIS](#) [Help](#) [Feedback](#)

[Search](#) [Expanded Search](#) [Command Search](#) [Results List](#) [Previous Searches](#) [Display Options](#) [My Account/Block Comments/Requests](#) [Start Over](#) [Help](#)

[Full Catalog](#) [Find Articles](#) [Journal Titles](#) [Digital Resources](#) [Reserves](#)

Results List
 Records 1 - 5 of 5

[Select All](#) | [Deselect All](#) | [Print/Save/Send](#) | [Create Subset](#) | [Add to List](#) | [View List](#) [Last Browse](#) | [Modify Search](#)

Results for: W-all keywords= Governing global markets
 Sorted by: Year (Descending), then Author
 Sort options: - Author(D)/Year - Author(A)/Year - Title(D)/Year - Title(A)/Year - Year(D)/Author - Year(A)/Author - Author/Title

[Jump to #](#)

| | | | |
|--------------------------|---|---|------|
| <input type="checkbox"/> | 1 | Offshoring : understanding the emerging global labor market / | 2005 |
| <input type="checkbox"/> | 2 | Governing financial globalization : international political economy and multi-level governance? / | 2005 |
| <input type="checkbox"/> | 3 | Governing the corporation : regulation and corporate governance in an age of scandal and global | 2005 |
| <input type="checkbox"/> | 4 | Ph.D Theses. 2005 November. | 2005 |
| <input type="checkbox"/> | 5 | Reinicke, Wolfgang M. Global public policy : governing without government? / | 1998 |

[Select All](#) | [Deselect All](#) | [Print/Save/Send](#) | [Create Subset](#) | [Add to List](#) | [View List](#) [Last Browse](#) | [Modify Search](#)

SEARCH: FULL CATALOG

Search for:

Title beginning with
 Title keywords
 Author (last name first)
 Author keywords
 ----- Scroll down for more choices -----

[New! Download LibX search tool](#) [About the HOLLIS Catalog](#)

Full View of Record : FULL CATALOG

Print/Save/Send | Add to List | View List | Last Browse | Back to Results List
Choose format: Full view | Short view | MARC

Record 1 out of 1

Title : Ph.D. Theses, 2005 November.

Locations/Orders : Availability

Location : Harvard Archives [UNPROCESSED ACCESSION 17262: availability may be limited] Holdings Availability

Notes : Thesis (Ph.D.)--Harvard University, 2005.

Authors : Aguirre, Andrew James. Genetic Determinants of Pancreatic Ductal Adenocarcinoma.
Albion, Alexis K. The Spy In All of Us: The Public Image of Intelligence in the 1960s.
Ali, Siraj Mahamed. Tale of Two mTOR Complexes.
Alvarez, James Valeriano. STAT3 Signal Transduction in Human Malignancy.
Angelova, Dillana Nikolova. Gender and Imperial Authority in Rome and Early Byzantium, First to Sixth Centuries.
Arnold, Mark A. Categorization of the Hitpa7el of Classical Hebrew.
Auerbach, Paul Asher. Mutational Specificity of Abasic Sites in Saccharomyces cerevisiae.
Aznar, Carolina Ana. Exchange Networks in the Southern Levant During the Iron Age II: A Study of Pottery Origin and Distribution.
Bathke, Lara Lynn. Mechanisms Underlying Genetic Diversity in Malaria Parasites.
Biaruk, Michael Jordan. Local Gate Control in Carbon Nanotube Quantum Devices.
Blackburn, David James Hart. Essays on the Economics of Copying and the Recorded Music Industry.
Bonanos, Alceste Zoe. Determining Accurate Distances to Nearby Galaxies.
Caldarone, Christopher Thomas. Iterated Branching Reaction Pathways: A Strategy to Access Diverse Small-Molecule Libraries Enabled by DNA-Templated Synthesis.
Carothers, James Matthew. Experimental Investigations of Informational Complexity and RNA Functional Activity.
Castreno, Adam B. Control of Lipid Synthesis and Distribution During Membrane Biogenesis.
Chen, Irene Ann. Building a Protocol: Physical Aspects and Emergent Behaviors.
Chiu, Anna. Studies Toward the Total Synthesis of Spiro-Procentrimine.
Clarkwest, George Andrew. African American Mental Distraction: What's History Got to Do With It?
Copolovitch, Mark Steven. **Governing Global Markets: Private Debt and the Politics of International Monetary Fund Lending.**
Custodio, Angel O. Analysis of Genetic Modifiers of Murine Iron Homeostasis.
Dalkic, Isin. Insights into the Role of ELNc in Muscle.
Diluzio, Willow Ruth. Behavior of Bacteria Near Surfaces and in Microchannels.
Dinea, Eleni. Lower Bounds for the Capacity of Channels with i.i.d. Deletions and Insertions.
Dunn, Travis Bailey. Synthesis of the C21-C40 Fragment of Azaspiracid-1.
Eden, Uri Tzvi. Point Process Filters in the Analysis of Neural Spiking Models.
Endries, Carrie Anne. Exiled in the Tropics: Nazi Protesters and the Get-It Vargas Regime in Brazil, 1933-1945.
Erwin, Phillip Allen. Dynamic S-Nitrosylation of Endothelial Nitric Oxide Synthase in Vascular Endothelial Cells.
Farazi, Paraskevi A. Modeling Liver Cancer in the Mouse.
Fogarty, Hugh Stephen. Critical Edition of the Middle Irish Saga Aided Gull meic Carbada agus Aided Gairb Glinni R. gi.
Fraligh, Matthew Patrick. Kanshibun in Modernizing Japan: The Case of Nanshima Ryuhoku (1837-1884).
Gafni, Chanan. Emergence of Critical Scholarship on Rabbinic Literature in the Nineteenth Century: Social and Ideological Contexts.
Gavle, Mahalia Christina. Imperiled Inheritance: The Decline of Politeness in 20th Century French Literature.
Gattelman, Debra Lynn. Reverie, Reading, and the Victorian Novel.
Glenn, Jason Edwin. Medicalizing Addictions, Criminalizing Addicts: Race, Politics and Profit in Narratives of Addiction.
Gonzales, Cynthia Inez. Text-Music Relationships in the Early Songs of Arnold Schoenberg.
Gross, Joshua Benjamin. Assessing Contributions of Cranial Neural Crest to the Skull in Xenopus laevis.
Harzylak, Jaroslaw. Using Functional Regression Models to Analyze Longitudinal Data.
Harris, Alison Meghan. Face-Selectivity of the M170 Response in Developmental Prosopagnosia and Double-Pulse Adaptation.

2) 검색 항목

[일반검색]

- 검색 항목 : 키워드, 서명 우측절단, 서명 키워드, 저자, 저자키워드, 주제 우측절단, 의학주제 우측절단, 총서명, 서명 우측절단, 중국 : 저자, 중국 : 서명 우측절단, 일본 : 저자, 일본 : 서명 우측절단, 한국 : 저자, 한국 : 서명 우측절단, ISBN, ISSN, Music Publisher number, Government number, 청구기호 (LC), 청구기호(NLM), Widener 청구기호, Tozzer 청구기호, Harvard-Yenching 청구기호, Dewey 청구기호, Other 청구기호, Su DOCS, Hierarchical place, HOLLIS number

[확장검색]

- 검색 항목 : 저자, 서명, 주제, 총서명, 출판사, 출판지, 서명 일부, 형식/장르, 모든 번호, 청구기호
- 제한검색 항목 : 언어, 소장위치, 자료유형, 연도제한
- 자료유형 종류 : Audio, Audio Music, Audio Spoken, Books, Digital, Journals/Serials, Manuscripts, Objects, Score, Visual
- 검색방법 : 단어, 구문 검색 지원

[Command 검색]

- 검색방법 : 검색대상 항목을 명령어리스트로 제공하여 여러개의 질의어를 동시에 검색할 수 있도록 지원. 전문가 검색 기능

wsn=education and wtn=journal
 retrieves records with the subject word "education" and the title word "journal"
 wsn="strategic planning" and wtn=management
 retrieves records with the subject phrase "strategic planning" and the title word "management"

- Commands list 예시

| 구분 | 명령어(약어) | 내용 |
|------------------|---------|--------------------------------|
| Numbers | SYS | HOLLIS No. |
| | LCN | LCCN(010) |
| | ISN | ISBN(020) |
| | KON | OCLC all(022) |
| | GPN | GPO Item No.(074) |
| Browse Commands | AUN | Authors, All |
| | AUNK | Authros, All Korean |
| | CDD | Call No. Dewey(PST) |
| | COT | Call No. Other(PST) |
| | TIN | Titles, All |
| | TIO | Title on piece(2xx_740) |
| | SUL | Subjects, LCSH(6xx_0) |
| | SFN | Subjects, Form/Genre, all(655) |
| | URE | URLs, 856 source(856) |
| Keyword Commands | NNN | Numbers all |
| | WRD | W-all keywords |
| | WAN | W-author all |
| | WFN | W-Notes |

3) 검색특징

- 학위논문을 위한 별도의 인터페이스 제공하지 않음
- 학위논문 서지사항도 간략하게 구성되어 있음
 - 서명, 저자만으로 구성되어 있으며 700태그 사용
 - 245태그(기본서명)에는 학위수여 연도로 입력되어 있음
- 학위수여연도별 서지레코드 생성으로 개별적인 학위논문 접근이 매우 불편함
- 간략결과 화면에서 학위논문 서명, 저자를 알 수 없음
- Archive Accession No. 로 학위논문 접근 가능

4) 원문, 부가정보 연계

- Harvard Archive에서 학위논문 원문 디지털화하여 소장 및 관리
- Archive Accession No.를 부여, 권한이 있는 이용자만 접근 가능 (제한접근)
- 초록, 목차, 참고문헌 등의 부가정보는 기본 결과화면에서 제공하지 않음

5) 타 서비스 및 시스템과의 연계

- Harvard Archive에서 원문을 디지털화하여 관리 및 온라인 검색시스템 운영
 - 1636년부터의 자료를 수집하여 원문 구축
 - 소장자료 내용
대학출판물, 종신교수의 연구물, 일반적인 대학 기록물, 대학관련 기록물, 학위논문(Dissertation, Theses), 수상작 연구물, 대학내의 시각자료(건물, 인물, 초기 사진자료 등)
 - Online Guide 및 Catalog 지원
- HOLLIS 검색 지원 : 소장자료의 70% 검색
- * OASIS (Online Archival Search Information System) : 별도 페이지 운영
<http://oasis.harvard.edu>

Quick search

[OASIS Home](#) | [Search](#) | [Process Finding Aids](#) | [Search History](#) | [Portfolios](#) | [Help](#)

Search OASIS

OASIS (Online Archival Search Information System) provides centralized access to a growing percentage of finding aids for archival and manuscript collections at Harvard. These finding aids are detailed descriptions of collections that contain a wide variety of materials, including letters, diaries, photographs, drawings, printed material, and objects. For each collection described in OASIS there is a summary description in [HOLLIS](#).

Tip: On searching

Search for: in

in

in

Limit by repository:

Number of items per search result page:

Hints:
 Use * as a wildcard. Examples: wom*n, photo* and *itics*

Thu Apr 19 04:15:04 EDT 2007 / Copyright 2004 by the President and Fellows of Harvard College | [Copyright Statement](#) | [Questions or Comments](#)

[OASIS Home](#) | [Search](#) | [Process Finding Aids](#) | [Search History](#) | [Portfolios](#) | [Help](#)

* VIA (Visual Information Access) : 별도 페이지 운영

<http://via.harvard.edu>

Visual Information Access

Search VIA

The Visual Information Access (VIA) system is a union catalog of visual resources at Harvard, focusing on access and search.

VIA includes catalog records for objects or images created, held or licensed by Harvard. Access to the catalog is open to the general public. All subject records and thumbnail images are available to everyone. Access to higher resolution images is usually available to the Harvard community. It always is determined by an individual repository and is often dependent on copyright.

Access to original object or image is determined by the individual repository. Restrictions on access may be noted in the VIA record.

For more detailed information, see [about VIA](#).

Tip: On searching | [On display preferences](#)

Search for: in

in

in

Limit search to records that have digital images

Limit search to records that have originals at Harvard

Limit search to records with dates from to

Limit repository to:

Hints:
 Use * as a wildcard. Examples: cat*, *operable and *itics*

Display/Sort Preferences:

Image grid size:

Sort by:

Result sets greater than 2000 will not be sorted

- 디지털 자원 서비스
- 직원에 의한 복사서비스 실시
- 마이크로필름 프린터 등

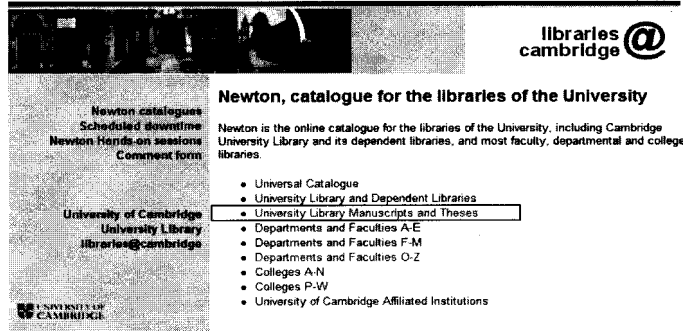
1.6 Cambridge Univ.

1) 검색방법

- 캠브리지 대학에서는 대표적으로 NEWTON, JANUS, COPAC Cataloge 통하여 여러 자료 유형 검색 지원
- 학위논문은 NEWTON과 COPAC을 통해 검색 가능

○ NEWTON

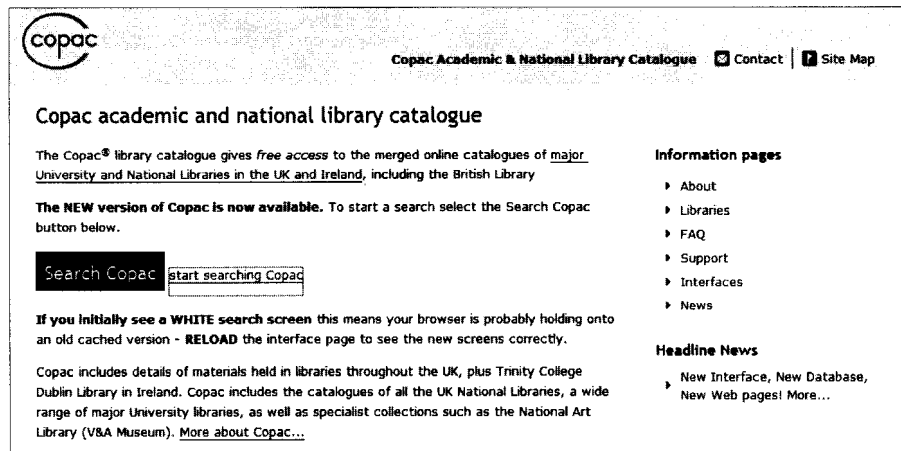
- 캠브리지대학의 도서관과 학부 및 관련 도서관의 온라인 목록제공
- NEWTON에서는 별도의 학위논문 인터페이스 구축
- 검색방법 : Basic Search, Guided search로 학위논문 검색 가능



○ JANUS : 캠브리지 Repository의 온라인 목록 제공

○ COPAC

- 영국도서관, 북아일랜드에 있는 주요대학 및 국립도서관의 온라인목록 제공
- E-Resource, 학술저널, 비도서자료, 학위논문등의 유형의 자료를 일괄 검색 또는 원하는 자료 유형을 선택하여 검색 가능
- 검색방법 : Quick Search, Main Search
- 상세결과 : 꼬리물기 검색 지원 (Title(제목), Subject(주제), Other names)



Copac Main Search

Quick Search Main Search **Advanced Search**

Author (etc): eg. T.S. Eliot. eg. BBC More...

Title words: eg. Birds of prey More...

Publisher: eg. Wild Swan More...

ISBN/ISSN (etc): eg. 1665920015 eg. 0262-4079 More...

Subject: eg. Railroads Norfolk More...

Keyword: eg. Tolkien Hobbit More...

Limit search by:

Date published: eg. 1997 eg. 2000- eg. 1750-1790 More...

Place published: eg. Leipzig. eg. London More...

Material type: All materials More...
 Electronic resources, Computer programs etc.
 Journals & other periodicals
 English & All Languages

2) 검색항목

○ NEWTON

- 학위논문 검색항목 : All, Title(논문명), Author(저자), Subject(주제), ISBN/ISSN, Publisher:Date, Name, ClassMark,

○ COPAC

- 학위논문 검색항목 : All, Title(논문명), Author(저자), Publisher, Subject(주제), ISBN/ISSN, keyword
- 제한 항목 :
 - Date published(발행날짜)
 - Place published(발행기관)
 - Material type(자료유형) : All, Electronic resources, Computer programs, Journals & other periodicals, Music scores & sound recordings, Theses, Visual materials
 - Language(수록언어) : ALL, KOR, ENG, JPN, CHI, FRE, GER 등
 - Library(도서관) : 영국도서관, 북아일랜드에 있는 주요대학 및 국립도서관

3) 검색특징

○ NEWTON

- Relevance(근접성)을 통하여, 검색어와의 근접도 순서대로 정렬가능
- 간략검색에서 검색어에 Boolean search 지원(AND, OR, NOT 지원)
- 해당 상세 정보를 인쇄 또는 개인 메일로 발송
- 원문보기 기능 없음
- COPAC
 - 논문 제목이 포함된 논문 검색
 - 해당 논문 소장 위치 제공
 - XML Record 제공

4) 원문, 부가정보 연계

- DSpace에서 Material Science and Metallurgy 관련 학위논문만 원문 제공되어 있으나 연계되어 있지 않음
- COPAC에서 "find a copy"를 통해 복본을 찾을 수 있는 정보 제공

5) 타 서비스 및 시스템과의 연계

- COPAC 목록 시스템을 통하여, 영국도서관, 북아일랜드에 있는 주요대학 및 국립도서관의 온라인 목록 제공
- 운용하고 있는 DSpace에서 Material Science and Metallurgy Collection에서만 학위논문 원문 제공
- 간략검색 결과 : Issue Date, Title, Author(s)

DSpace

Home | DSpace at Cambridge > Materials Science and Metallurgy > Device Materials Group > Theses >

Browse by Title

Jump to: 0-9 A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
or enter first few letters:

Showing items 1-21 of 26.

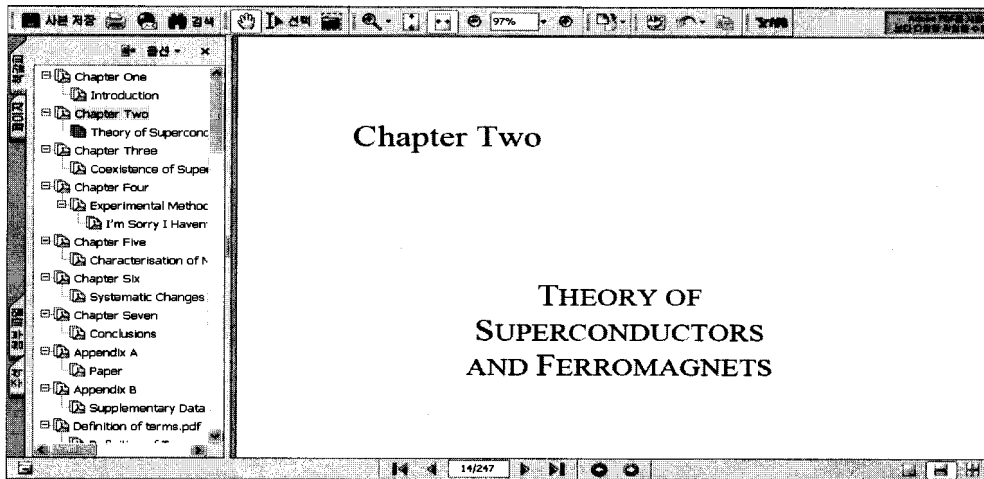
| Issue Date | Title | Author(s) |
|------------|--|-------------------------|
| Apr-2004 | ACTIVE CONTROL OF SUPERCONDUCTIVITY BY MEANS OF THE FERROMAGNETIC EXCHANGE INTERACTION | Kinsey, Robert |
| Mar-2005 | Co2MnSi Heusler alloy thin films | Singh, Laura Jane |
| Jun-2004 | Control of Superconductivity in Cuprate/Manganite Hetrostructures | Pang, Brian Siew-Jan |

○ 상세검색 결과

- Title, Author, Keywords, Issue Date, Citation, Abstract, 원문다운로드

Title: ACTIVE CONTROL OF SUPERCONDUCTIVITY BY MEANS OF THE FERROMAGNETIC EXCHANGE INTERACTION
Authors: Kinsey, Robert
Keywords: Ferromagnetic
Exchange Interaction
Superconductivity
Devices
Issue Date: Apr-2001
Citation: Thesis, University of Cambridge
Abstract: Recent theoretical studies have suggested that the observed suppression of superconductivity in superconductor/ferromagnet (S/F) heterostructures could be modulated by controlling the ferromagnetic exchange interaction in the superconductor. The exchange interaction in the superconductor is the sum of the exchange interaction from the ferromagnetic regions, which has

○ 원문 서비스 : PDF 파일(Acrobat reader 이용)



- 서비스 제한 없음 : 보기, 저장, 출력 가능
- DRM, 워터마크 미적용

1.7 Stanford Univ.

1) 검색방법

- 꼬리물기 검색 : 논문 저자, 논문 수여기관의 학과명, 지도교수
- 자료유형별 검색에서 학위논문만을 위한 검색 지원하지 않고 있음

Conceptual foundations of cultural evolution
Reisman, Kenneth.

Find

↕ Jump to call number/location/availability ↕

Author: Reisman, Kenneth.

Title: Conceptual foundations of cultural evolution / Kenneth Reisman.

Imprint: 2005.

Physical Description: xv, 254 p.

Note: Submitted to the Department of Philosophy.

Note: Copyright by the author.

Note: Thesis (Ph. D.)--Stanford University, 2005.

Organization: Stanford University, Dept. of Philosophy.

Local author: Godfrey-Smith, Peter advisor

Catkey: 6212884

Call Number, Location, and Availability

Stanford Auxiliary Library - SAL1&2 [Nearby items on shelf]

| Copy | Holdings | Status |
|-------------|----------|--|
| 3781 2005 R | 1 | Request this item [with SUNetID] or [without SUNetID] <input type="checkbox"/> |

Special Collections & University Archives [Nearby items on shelf]

| Copy | Holdings | Location |
|-------------|----------|---|
| 3781 2005 R | 1 | * University Archives -- Request [with SUNetID] or [without SUNetID] <input type="checkbox"/> |

Philosophy Library [Tanner] [Nearby items on shelf]

| Copy | Holdings | Location |
|-------------|----------|---|
| 3781 2005 R | 1 | Due: 4/10/2007 -- Request [with SUNetID] <input type="checkbox"/> or [without SUNetID] <input type="checkbox"/> |

2) 원문, 부가정보 연계

- 원문은 서비스 하고 있지 않으며, UMI(ProQuest)를 통해서 이용 가능
- 초록, 목차, 참고문헌 등의 정보를 따로 서비스 하고 있지 않음
- DSpace 운용하지 않음

1.8 Univ. of Tokyo

1) 학위논문 검색방법

- 도쿄대학에서 수여된, 신 제도 과정 박사논문, 박사논문의 서지사항, 요지 수록
- 수록범위 : 서지 1957.4~2005.3 / 요지(초록) 1994.4~2005.3
- 수록건수 : 36,775건(요지가 있는 데이터는 14,295건)
- 2002년 2월 도쿄대학학위논문논제데이터베이스(서지DB)와 도쿄대학학위논문요지데이터베이스(초록DB)를 통합하여 시행
- 원문은 제공하지 않음

2) 검색항목

- 검색항목 : 전체 항목, 논문제목, 저자명, 요지, 논문심사위원, 보고번호, 학위기(學位記)번호, 연구과명/전공명
- 학위수여년월일, 학위 종류(과정박사, 논문박사), 학위 종류별 제한 가능

学位授与年月日: 年 月 日 - 年 月 日

学位の種類別: 課程博士 論文博士

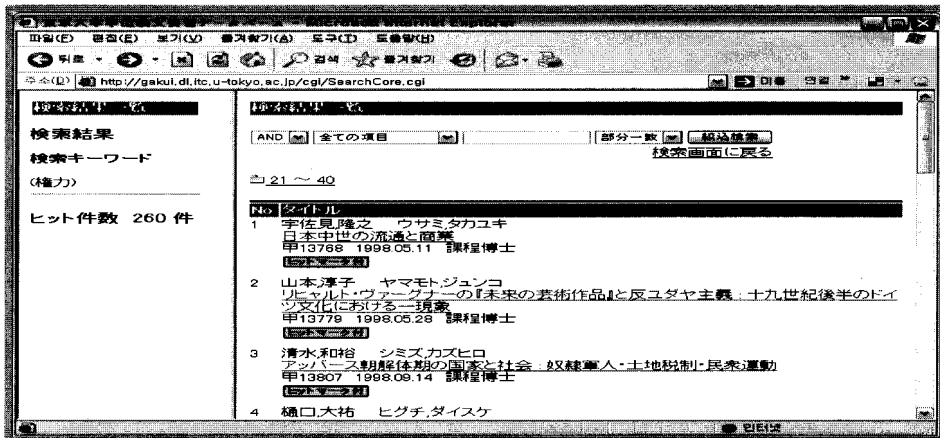
学位の種類:

| |
|-------|
| 文学 |
| 心理学 |
| 社会学 |
| 社会心理学 |

1ページに表示する件数

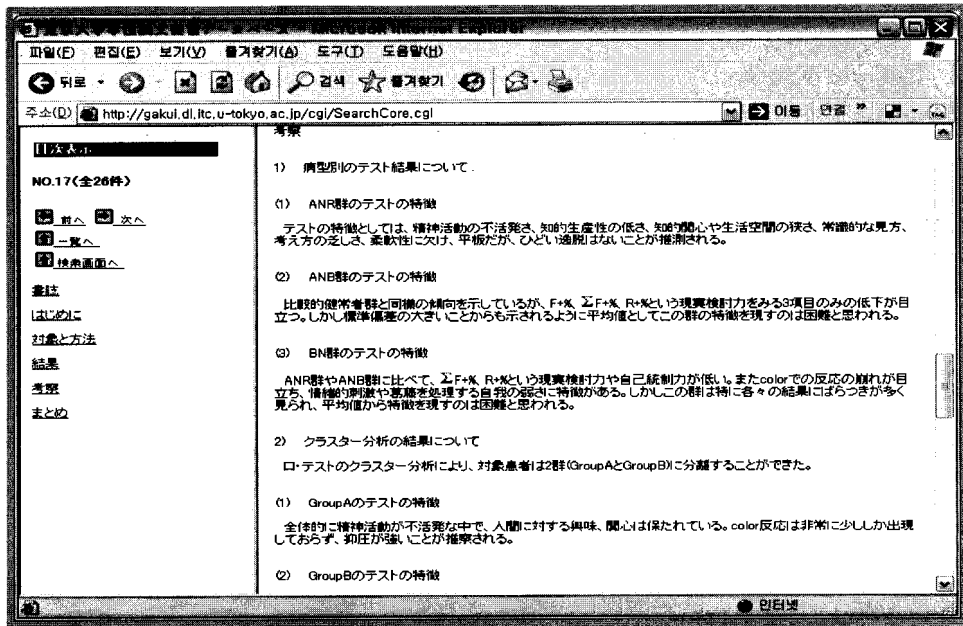
- 부분일치, 전체일치 검색 가능
- 재검색 항목은 없음

3) 검색특징



4) 부가정보 연계

- 제공정보 : 서지, 내용요지, 검사요지
- 내용요지의 분량이 많을 경우 화면 왼쪽에 북마크 제공



- 2000년 이후의 요지는 Acrobat으로 확인 가능

5) 타 서비스 및 시스템과의 연계

○ FELIX(Front End of Library Information eXpansion)

- 일본 국내·외의 주요한 학술잡지 논문·기사 검색 가능
- 이용자 제한 (교육용계산기시스템(ECCS)의 계정이 있을시 이용 가능)

○ CiNii : NII 논문 정보 내비게이터

- 일본 내의 학술논문, 학술협회지, 대학의 연구기요(紀要)를 중심으로 한 논문정보 데이터베이스
- 이용자 제한

2. 국내 학위논문 서비스 사례

2.1 KAIST

1) 검색방법

- 전체 검색에서는 소장 일반도서, 학위논문, 보고서, 학술저널, 비도서 자료, 저널수록논문 등과 같은 여러 유형의 자료를 한꺼번에 검색 또는

원하는 자료 유형을 선택하여 검색 가능하며, 검색방법으로 키워드검색 (By Keyword)과 전방일치검색(By Front Match)이 지원

- 별도 학위논문 검색페이지 존재 - 키워드검색(By Keyword) 지원

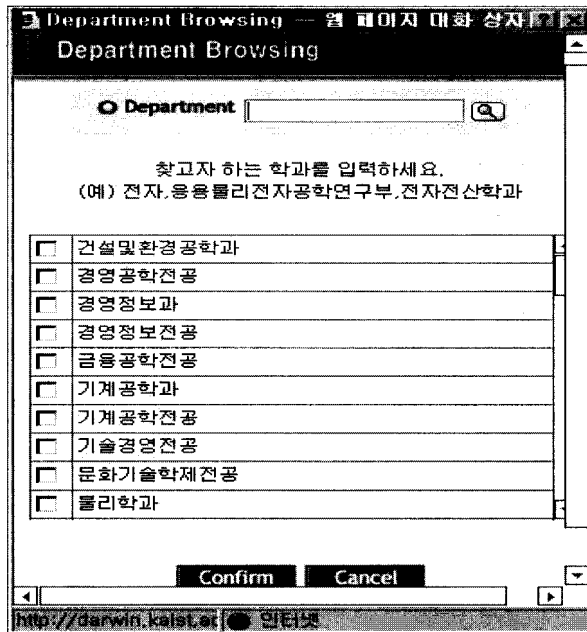
The screenshot shows a search interface for Theses at KAIST. The search method is 'Search by Keyword'. There are three search input fields: 'All', 'Title', and 'Department', each with an 'AND' operator. Below the search fields are filters for 'Format' (All, Paper, Online), 'Degree' (All, Ph.D, Master), 'Language' (All), and 'Year'. At the bottom are 'Search' and 'Clear' buttons.

2) 검색항목 및 제한

- 학위논문 검색항목
 - All, Title(논문명), Author(저자), Advisor(지도교수), Abstract(초록), Department(학과)
- 제한 항목
 - Format(자료형태) : All, Paper, Online
 - Degree(학위) : All, Ph.D(박사), Master(석사)
 - Language(수록언어) : ALL, KOR, ENG, JPN, CHI, FRE, GER
 - Year(학위년도)

3) 검색특징

- 지도교수명, 초록, 학과로 검색 가능
- 학과별 리스트 제공하며 복수 선택이 가능



4) 검색 결과

○ 간략검색 결과

Search > Holding Materials > Theses > Brief Information

Korea Advanced Institute of Science and Technology KAIST

Guide

Send to My Basket Total : 4 / 4 건

| <input type="checkbox"/> | No | Title / Author / Department (Call No.) | Year | Type | Abstract | Fulltext |
|--------------------------|----|---|------|--------|----------|----------|
| <input type="checkbox"/> | 1 | Cloning and characterization of the gene encoding bdellin from <i>Hirudo nipponia</i> = Hirudo nipponia로부터 분리한 bdellin 유전자의 클로닝과 구조분석 / Choi, Jung-Kyoon / 생물과학과 (MBS 00028) | 2000 | Thesis | | |
| <input type="checkbox"/> | 2 | 수확연구정보센터의 의의와 활용에 관한 고찰 및 유한요소법의 안정화 = A study on the center of mathematical research information and the stabilization of finite element method / 심문보 / 수학과 (MMA 99010) | 1998 | Thesis | | |
| <input type="checkbox"/> | 3 | Atomic force microscope를 이용한 $\text{Ge}_{2}\text{Sb}_{2}\text{Te}_{5}$ 박막의 상변화 거동 및 전기적 특성평가 = The estimation of the phase change behavior and electrical property of $\text{Ge}_{2}\text{Sb}_{2}\text{Te}_{5}$ thin films using atomic f / 노기민 / 신소재공학과 (MAME 05037) | 2005 | Thesis | | |
| <input type="checkbox"/> | 4 | 제품 시스템 디자인의 아이디어 전개 도구를 위한 개념모델에 관한 연구 = A concept model for idea deployment tools of product system design / 이원 / 산업디자인학과 (MID 05004) | 2005 | Thesis | | |

FIRST 1 LAST

○ 상세검색 결과 : 서지정보, 초록, 소장위치, 원문 URL

Search > These > Detail Information KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY KAIST

Record Position : 1 / 4

| | |
|----------------------|---|
| Call No | MBS 00028 |
| Title | Cloning and characterization of the gene encoding bdellin from Hirudo nipponia = Hirudo nipponia 도부러 분리한 bdellin 유전자의 클로닝과 구조분석 / |
| Personal Author | Choi, Jung-Kyoon 최정균 |
| Imprint | [대전 : 한국과학기술원, 2000]. |
| Physical Description | iv, 48 p. : 삽도 : 26 cm |
| Language | 영어 |
| Advisor | Kim, Sun-Chang 김선창 |
| Dissertation Note | 학위논문(석사) - 한국과학기술원 : 생물과학과, |
| Bibliography Note | 참고문헌 : p. 43-45 |
| Subject | Leech Bdellin Gene cloning Genomic library 거머리 델린 유전자 클로닝 유전자 도서관 |

KOREAN

Bdellin - KL은 Hirudo nipponia도부러 분리된 bdellin으로서 trypsin, plasmin, acrosin inhibition activity를 가지고 있다. Bdellin의 생리적 기능 총 알기 위한 연구의 하나로서 bdellin-KL의 유전자를 library의 screenings을 통해 클로닝하고 그 염기서열을 분석하였다. 전체 open reading frame과 5' 3'인접 서열을 가지고 있는 clone를 찾아낸 후 그 구조를 분석한 결과 이 유전자는 3개의 exon과 2개의 intron으로 이루어졌으며 splicing position에서 GT-AG rule를 따르고 있음이 밝혀졌다. Primer extension analysis를 통해 transcription 시작점이 시작 codon으로부터 약 60 염기쌍 정도 떨어진 곳임을 알 수 있었고 그 곳에서 28염기쌍 위 쪽으로 TATA box가 존재하고 있음을 밝혔다. 각각 두 곳의 GC, CAAT box서열이 TATA box 위 쪽으로 위치해 있었고 AP-1, AP-3 결합 부위가 있었다. 이러한 factor들의 전체 혹은 일부가 bdellin의 조직 비특이적 발현을 조절하고 있을 것으로 생각되었다. 조직 특이성을 가지고 발현되는 몇 가지의 transcription factor들이 binding할 수 있는 서열들을 찾아 내었는데 대표적으로 GATA-1, NF- κ B, E2A 결합 부위, E-box 등을 포함하고 있었다. 이 factor들은 bdellin의 거머리 내에서의 생리적 기능과 밀접한 관련이 있을 것으로 생각되며, 이들 factor들의 발현을 조절하는 유전자들을 찾아내어 유전자 발현을 조절할 수 있는 방법을 연구할 수 있을 것이다.

| No | Accession No. | Location | Volume No. | Copy No. | Book Status | Due Date | Loan Status | Check Location |
|----|---------------|----------------|------------|----------|-------------|----------|-------------|----------------|
| 1 | 8010711 | 대전 중앙관2층 학위논문실 | | | 배가중(도서관에) | | 대출가능확인 | 실물확인요청 |

※2002년도 이전의 학위논문 원문 이용시에는 이미지뷰어를 다운로드 받아서 설치하신후에 이용하시기 바랍니다.

http://darwin.kaist.ac.kr:20080/csl-bin/Maestro?func=image_link&bib_no=158716&gubun=hw

Copyright © 2003 Korea Advanced Institute of Science and Technology All rights reserved.

5) 원문 연계 서비스

- 2002년도 기준으로 원문 뷰어 다른 형태로 제공
- 서비스 제한
 - 서비스 제한 없음 : 보기, 저장, 출력 가능
 - DRM, 워터마크 미적용

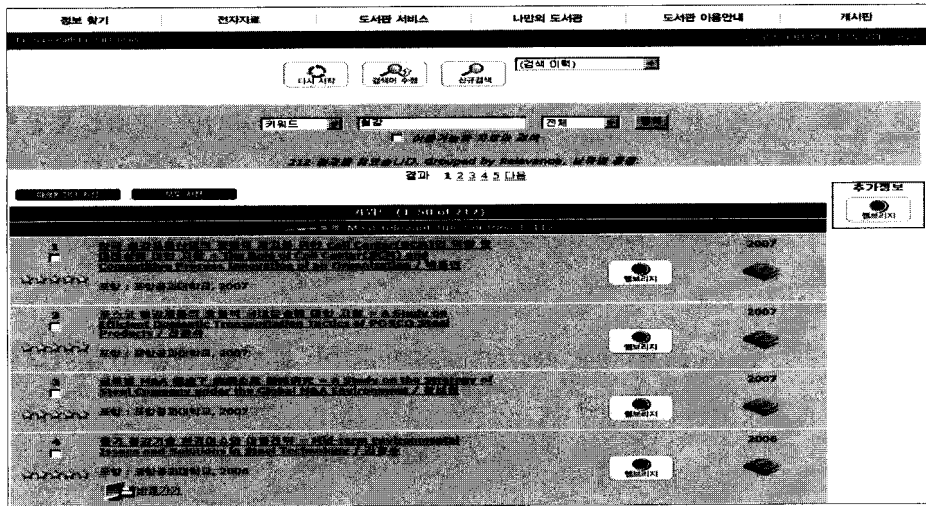
6) 학위논문 PDF 구축

- 자체 학위논문 제출 시스템 구축
 - "Library"의 메인화면에서 "Dissertation Submission"에 로그인(ID:학번, 암호:주민등록번호 뒷부분 7자리)
 - 입력사항 입력 및 온라인 논문파일 제출
 - 온라인 제출 완료 후 전자학위논문 공개동의서 출력
 - 인쇄된 학위논문 7부와 동의서를 기한 내에 교무팀에 제출

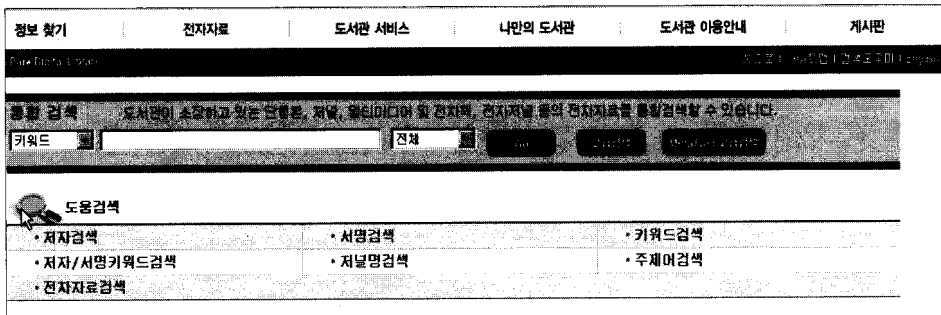
2.2 포항공대

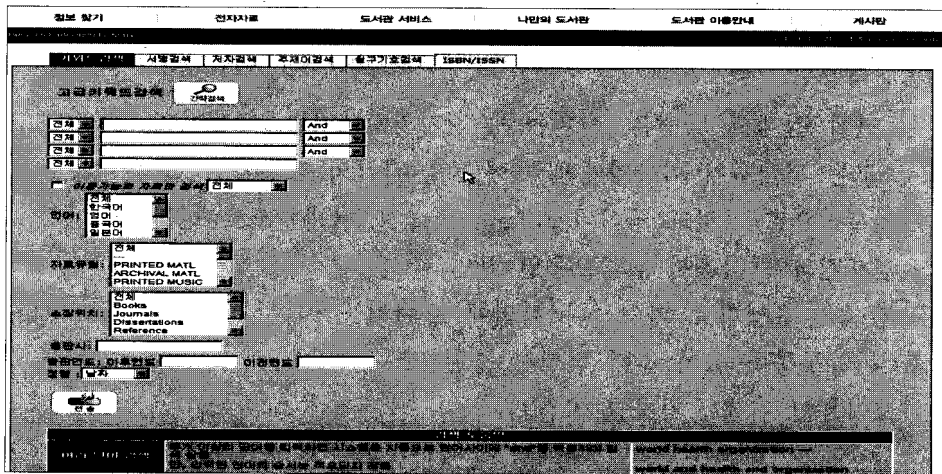
1) 검색방법

- 전체 검색시 단행본, 학위논문, 전자자료, 저널, 멀티미디어, 참고자료를 모두 검색하여 이용자에게 보여주며, 각각의 유형별로 검색 가능
- 기본적으로 키워드 방식이 default로 설정, 날짜순으로 정렬
- 검색결과가 없을시 검색도움말 화면 출력
- 검색은 유형별로 가능하나 화면에 보여지는 아이콘이 유형별로 구별되지 않은 부분이 있어 전체 검색시 불편
- 꼬리물기 검색 : 저자만 지원



2) 검색항목

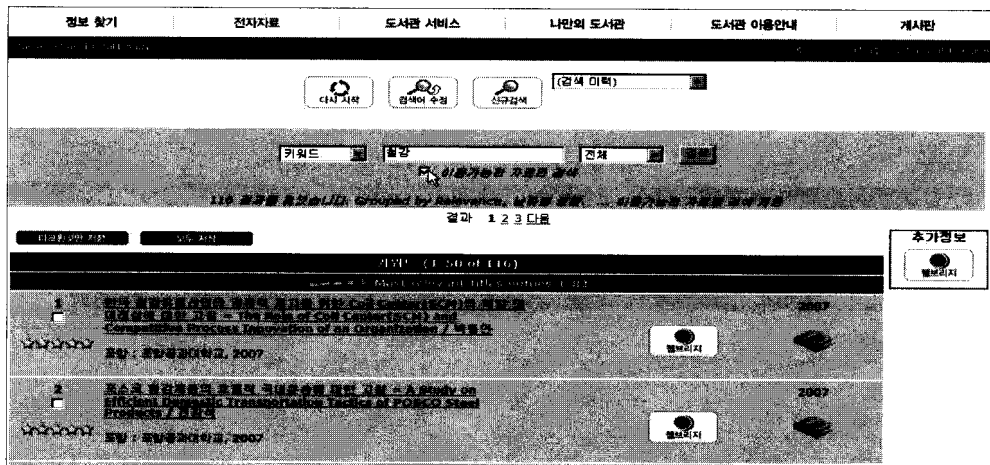




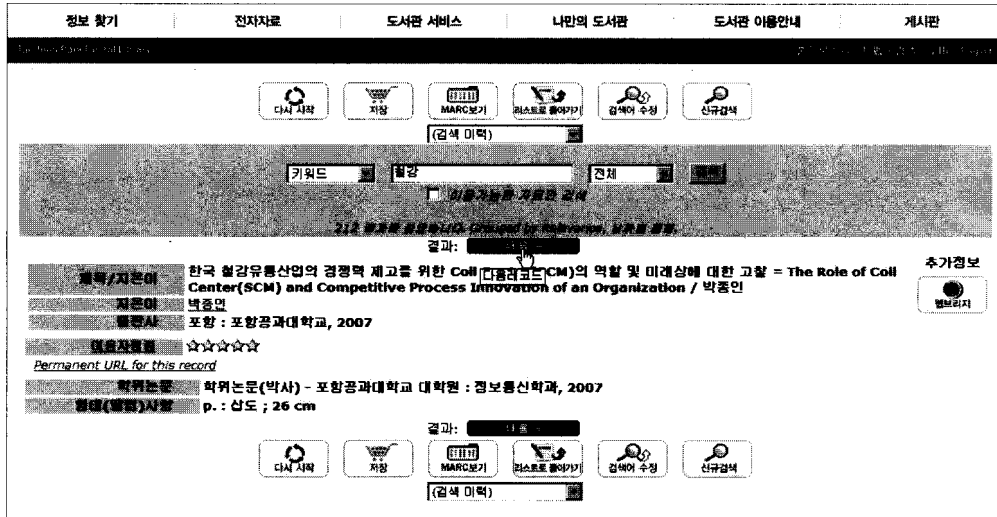
- 여러 유형의 자료에 대해 통합검색을 통해 보여주지만 학위논문만의 구별된 항목들인 자료형태(Paper, Online), 학위, 수록언어, 학위년도 등의 검색은 지원하지 않음

3) 검색특징

- '이용가능한 자료만 검색'이라는 항목이 있어 실제로 권한이 부여되어 이용가능한 자료만 보여짐
- 학내로 제한된 자료는 로그인을 통해 이용가능



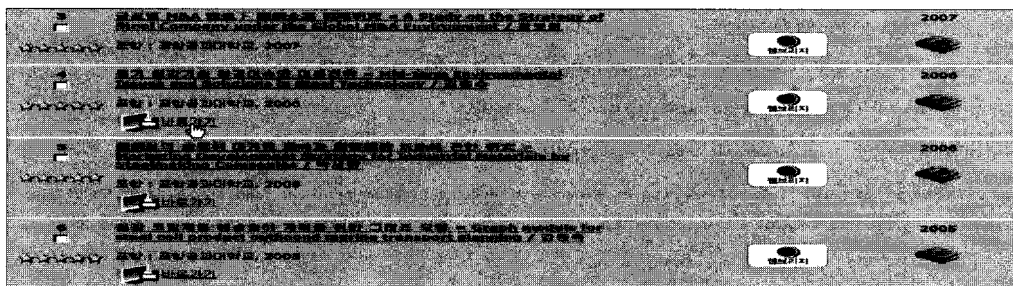
- 상세화면에서 다음 결과물의 상세화면으로 쉽게 이동이 가능, 다시 간략 검색화면으로 가지 않아도 바로 상세화면으로 접근 가능



- 상세화면에 '이용자평점'이라는 항목을 두어 이용자들이 평가한 결과를 별의 숫자로 구분하게끔 되어 있으나 실제 평점이 부여된 자료는 매우 적음

4) 원문, 부가정보 연계

- 원문 여부는 아이콘을 통해 구별가능, 상세서지에서 자세한 열람범위를 보여주고 있으며, 자료의 이용을 돕기 위하여 소장위치를 풍선도움말을 통해 지도로 보여줌



library.postech.ac.kr/htdocs/use

키워드 검색

학위논문

학위논문(석사) - 포항공과대학교 정보통신대학원 : 정보통신학과(기술경영 전공), 2006

iii, 31 p. : 삽도 ; 26 cm

결과: 1 2 3 4 5 6 7 8

5) 타 서비스 및 시스템과의 연계

- Z39.50 프로토콜을 이용해 타기관 및 타도서관의 자료 열람 가능

검색조건 - 시명 education

in MULTIPLE SERVERS

| 번호 | 500 찾는 항목, 항목 1-8 | Server |
|----|--|--------|
| 1 | Accountability in education / edited by P. R. Chippendale and Paula V. Wilkes. Chippendale, Peter Richard, 1931- | KERIS |
| 2 | Adult and worker's education periodicals. Unesco | KERIS |
| 3 | Adult education and leisure-time activities in Czechoslovakia. Unesco | KERIS |
| 4 | Adult education in community development. Unesco | KERIS |
| 5 | Aerobic pas de deux / Choreography [and] notator: Lori Katterhenry 1982 ; Music: Dance of the Swans Tchaikovsky. Katterhenry, Lori | KERIS |
| 6 | The American schools in the past and present VI / by Ostap E. Oryshkewych. Oryshkewych, Ostap E. | KERIS |
| 7 | And after the journey / Music: Heitor Villa-Lobos ; choreography and notation: Georgette Welsz Amowitz. Villa-Lobos, Heitor, 1887-1959 | KERIS |
| 8 | Belief, morals, and education : collected essays on the philosophy of education / Wolfgang Brezinka ; translated by James Stuart Brice. Brezinka, Wolfgang, 1928- | KERIS |

2.3 부산대

1) 검색방법

- 학위논문 원문구축 현황 : 10,661건('06.03.01 현재)

2) 검색항목 및 제한

- 학위논문 검색항목
 - All, Title(논문명), Author(저자), Advisor(지도교수), Abstract(초록), Department(학과)
- 제한 항목
 - 학위구분 : 전체, 박사, 석사
 - 수록정보 : 전체, 원문, 초록, 목차
 - 언어 : 전체, 한국어, 영어, 일어, 불어, 독어, 중국어, 러시아어, 기타
 - 발행년도

3) 검색특징

- 키워드검색(By Keyword), 전방일치검색(By Front Match)
- 주제별리스트
- 검색조건에 완전일치를 선택할 수 있음
- 지도교수명, 초록, 학과로 검색 가능
- 검색히스토리 보기 메뉴가 있어 이전 검색항목을 조회할 수 있음.

부산대학교 도서관
BUSAN NATIONAL UNIVERSITY LIBRARY

로그인 | P.제출 | 내노출명세서 | 발행도서명 | 외역도서명 | STEMAP | ENG

MY LIBRARY | 소장자료검색 | 신착자료검색 | 도서관서비스 | 도서관안내

정책자료 | 단행본 | 연속간행물 | 학위논문 | 멀티미디어 | 고분자 | 기사 | 학술행사자료 | 교수자료 | 신성자료 | 지정도서

학위논문

키워드 | 전방일치 | 주제별리스트 | 도움말

서명 | AND |
저자 | AND |
발행처 | 다국어입력

학위구분 전체 석사 박사
수록정보 전체 원문 초록 목차
언어 전체
발행년도
최대 검색건수 500

검색 | 다시쓰기 | **검색이 스토리**

4) 검색 결과

○ 간략검색 결과

부신대학교 도서관
1998 P.888 나노생물과학도서관 법학도서관 국제도서관 SITEMAP ENG

MY LIBRARY 소장자료검색 전자자료검색 논문서비스 투데이뉴스

전체자료 | 단행본 | 연속간행물 | 학위논문 | 멀티미디어 | 교양헌 | 기사 | 학술행사자료 | 교수저작물 | 신착자료 | 지정도서

학위논문

검색어: [전방일지/ 서명:경명학] 결과내검색 [서명] 국어 검색

이전페이지 | 검색결과

4건중 4건 출력 [1/1]

| 선택 번호 | 논문명 | 저자 | 발행처 | 소장처 | 발행년 | 자료 유형 | 원문 | 초록 | 목차 |
|----------------------------|--|-----|---------|-----|------|-------|----|----|----|
| <input type="checkbox"/> 1 | OWL이 등록 접근과 통찰스트레스에 미치는 影響에 관 한 實驗的 研究 | 윤희경 | 朝鮮大學校 | 제2 | 1998 | 학위논문 | | | |
| <input type="checkbox"/> 2 | 經濟學 分野의 學術情報 流通 및 利用에 관한 研究 | 이진환 | 淑明女子大學校 | 제2 | 1989 | 학위논문 | | | |
| <input type="checkbox"/> 3 | 經濟學 分野 디소오라스作成的 自動化 | 김중현 | 경북대학교 | 제2 | 1984 | 학위논문 | | | |
| <input type="checkbox"/> 4 | 經濟學 科 大學院生들의 情報利用 및 圖書利用 行態 | 김귀영 | 경북대학교 | 제2 | 1985 | 학위논문 | | | |

상세정보 | 비귀납키 | 비귀납키 | 서명 | 원문 | 초록

○ 상세검색 결과

- 제공정보 : 상세정보, 동일자료, Online Access, 대출정보, 초록, 목차

MY LIBRARY 소장자료검색 신착자료검색 도서관서비스 도서관안내

전체자료 | 단행본 | 연속간행물 | 학위논문 | 멀티미디어 | 고문헌 | 기사 | 학술영사자료 | 교수저작물 | 신학자료 | 지정도서

상세정보

도움말

| | | | | | | | |
|---------|--|------|-----------|------|----|-----|----|
| 이전화면 | 검색정보 | 검색화면 | [9/499 건] | MARC | 원문 | URL | 바탕 |
| 자료유형 | 학위논문 | | | | | | |
| 개인저자 | 김을집 | | | | | | |
| 출판사 | 부산대학교 | | | | | | |
| 서명/저자사항 | 사회과 교과서 재구성이 중학부 청각장애학생의 사회과 학업성취도 및 어휘습득에 미치는 영향= The influences of reforming of social study textbook in the academic achievement and vocabulary acquirement of the students of the middle school for deaf/ 김을집. | | | | | | |
| 발행사항 | 부산: 부산대학교, 2003. | | | | | | |
| 형태사항 | iv, 74 장 : 26 cm. | | | | | | |
| 학위논문주기 | 학위논문(석사)- 부산대학교 교육대학원: 특수교육전공, 2003년 2학기 | | | | | | |
| 서지주기 | 참고문헌 | | | | | | |

동일자료

사회과 교과서 재구성이 중학부 청각장애학생의 사회과 학업성취도 및 어휘습득에 미... [원문](#)

Online Access

| | |
|----|---|
| 번호 | 소장사항 |
| 11 | http://dcollection.pusan.ac.kr/jsp/common/DcLoOrgPer.jsp?ItemId=00000020515 |

대출정보

| 번호 | 등록번호 | 소장처/자료실 | 청구기호 | 도서상태 | 반납예정일 | 예약 | 캠퍼스간 | 자료대출 | 신청 |
|----|--------|-------------------------|------------------------|------|-------|----|------|------|----|
| 1 | 634983 | 제2도서관 1층 인문사회과학자료관 학위논문 | HDT 371.912 김68 a. c.1 | 열람가능 | | | | | |

V. 결론 및 제언 51

1. 결론 51
2. 제언 53

참고 문헌 54

영문 초록 59

부록 62

바구니담기 바구니보기 저장 메일 출! 복사/인쇄/다운로드

- 상세결과 특이사항
· 메일 보내기

메일

도움말

이전화면

| 선택 번호 | 서명 | 저자 | 발행처 | 소장처 | 발행년 |
|-------------------------------------|---|-----|-------|---------|------|
| <input checked="" type="checkbox"/> | 사회과 교과서 재구성이 중학부 청각장애학생의 사회과 학습성취도 및 어휘습득에 미치는 영향 | 김홍집 | 부산대학교 | 371.912 | 2003 |

저장항목 간략정보 상세정보 소장정보 초록 목차

메일주소

메일 My Collection 저장 다시보기

- 찾지 못한 자료신고
- 이용자 ID로 로그인 후 신고

5) 원문 서비스

- KERIS의 dCollection 사업 참여 : 이미지 뷰어 Acrobat 7.0

3. KERIS(한국교육학술정보원) 온톨로지 기반 서비스 사례

1) 개요

- 교육학 세부 10개 분야 정보를 온톨로지로 구축
 - 6,000여개 교육학 개념 및 용어에 대한 의미기반 검색서비스 제공
- 주요 서비스 내용
 - 의미기반검색, 추천 교수, 관계논문, 연관 주제저
 - 관계 논문 대상 : 교육학 분야 논문 11만건
 - 연관 주제어 서비스 : 10만개 단어의 연관도를 비교하여 추천
 - 검색어 자동완성 기능 : 검색어 입력시 시스템에서 제공하는 리스트 제공

2) 검색특징

- 특정 검색항목이 없이 접근 가능

- 키워드, 자연어 입력으로 검색 지원
- 해당 분야의 기초(기반)지식이 필요

○ 질의어 검색 시 결과내용

- 개념, 추천교수, 논문, 저자, 발행처, RISS 검색, 연관주제어 등이 보여짐
- 결과항목에 대한 세부건수 명시되어 제공
- 자연어의 문장 검색시 의미기반의 검색을 완전하게 수행하고 있지 못함
예) 질의어 : 평생교육에 있어서 안드라고지는?

[검색결과]

- 질의어가 총 4개로 구분되어 검색이 수행됨
(평생교육법, 평생교육사, 평생교육학, 안드라고지)
- 개념 : 평생교육법과 관련된 법 등이 제시됨
- 논문 : 교육법에 관련된 논문으로 검색 실행
- 저자 : 평생교육과 관련된 교수의 저작물이 검색 실행
- 발행처 : 평생교육과 관련된 기관으로 검색 실행
- 국내학술지논문 : '제목+저자+출판사 => 교육법'으로 검색 실행

