

학술정보자원에 대한 Topic Map 자동구축 방안

장화수, 고일주
송실 대학교, 미디어학과
e-mail: {jhsu01, andy}@ssu.ac.kr

Topic Map automatic construction Study for research information resource

HwaSu Jang, IlJu Ko
Dept. of Media, Soongsil University

요 약

Topic Map을 구축하는데 있어서 봉착하는 문제는 특정분야 전문가들이 Topic Map의 구성과 체계에 익숙하지 않다는 점이다. 이를 해결하기 위해서 Topic Map의 모든 요소들을 새로이 작성하는 것 보다는 작성하려는 분야에 대해 기 구축된 정보자원이 존재할 경우 이를 최대한 재활용하여, 모든 요소들을 추출한 다음 Topic Map 온톨로지로 변환하고 이용하는 것이 시간과 비용을 절약할 수 있는 효율적인 방법일 것이다.

본 연구에서는 기 구축된 학술DB로부터 Topic Map에서 재활용할 수 있는 요소들을 추출하기 위한 정보 소스로서 데이터베이스 스키마와 MARC에서 언급하는 메타데이터를 이용하는 것은, 기초 학문자료의 복잡한 관계의 개념구조, 자료유형 및 자료간의 의미적 상관관계 표현에 있어 효율적인 개발방법임을 제안한다.

키워드 : 도서관 시스템, marc, komarc, marcxml, 시맨틱웹, 토픽맵, TAO, xtm, marcxm, web2.0

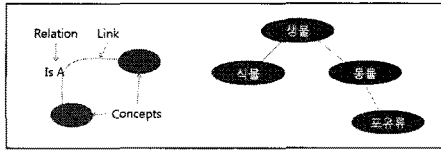
I. 서론

현재 웹 자원들에 관한 메타데이터의 의미를 기계가독형으로 표현할 수 있는 시맨틱웹을 구축하기 위해서 많은 노력을 기울이고 있다. 하지만 학술정보자원과 같은 특정분야의 Topic Map을 구축하기 위해서는 각 특정분야의 전문가들의 도움이 절대적이다. 특정분야 온톨로지를 구축하는데 있어서 봉착하는 문제는 이 특정분야 전문가들이 온톨로지의 구성과 체계에 익숙하지 않다는 점이다. 이를 해결하기 위해서는 특정분야 전문가와 Topic Map 구축 전문가가 협력하여 공동으로 온톨로지를 구축하여야 하나 모든 특정분야에서 항상 같이 작업하기가 현실적으로 어려우며, 기존 시스템에서 새로운 Topic Map을 구축하는 것은 많은 비용과 시간이 소요되는 등 어려운 일이다. 서지정보를 포함하고 있는 도서관 정보시스템의 DB로부터 메타데이터의 의미를 기계가독형으로 표현할 수 있는 Topic Map을 이용하여 재 설계하고, 변환모듈을 통하여 온톨로지를 구축할 때, 보다 효과적인 개발방법으로서 데이터베이스 응용업무에 대해 작성되어 있는 ERD(Entity-Relationship Diagram)를 이용하여 Topic Map 개발에 활용하는 방안을 제안하고자 한다.

II. 시맨틱 메타데이터와 Topic Map

시맨틱 웹을 간단히 정의하면 컴퓨터가 네트워크 상의 정보에 대해 그 의미를 이해하고 사람이 관여하지 않아도 이를 가공하고 처리할 수 있는 웹이라고 할수 있다. 기존의 월드와이드웹을 비롯한 환경에서는 검색에 필요한 추가적인 정보가 주어지지 않으면 사용자 개개인이 모든 정보의 유용성이 합당성을 일일이 판단해야 하는 부담을 안게 된다. 반면 시맨틱 웹에서는 어떤 정보에 대해 정보제공자, 사용자가 제공한 경험, 관련정보에 대한 체계 등을 참조해 메타데이터에 미리 정의해 놓아 컴퓨터가 이를 활용, 정보를 가공해 사용자의 판단에 도움을 준다. 이런 시맨틱 웹의 핵심이 되는, 즉 의미와 개념을 정의하는 데이터 모델이 바로 시맨틱 메타데이터로, 현재 RDF(Resource Description Framework), OWL, Topic Map 등의 표준이 존재한다. 이 중 Topic Map은 ISO에서 제정한 표준이며, 이미 널리 쓰이는 표준인 RDF에 비해 고수준의 추상화를 제공한다라는 특징이 있다. 고수준의 추상화가 중요한 이유는, 종래의 시맨틱 메타데이터는 실제계의 거의 모든 정보를 표현할 수 있는 강력한 표현력을 지녔지만 그 강력함이 도리어 실제 사용자가 그런 데이터를 정의하고 관리하는 데는 커다란 장벽이 되고 만다. 실제로 RDF 등으로 구성된 온톨로지를 구축하는 것은 관련분야의 전문가

에게도 쉬운 일은 아니다. 반면 Topic Map이 제공하는 추상화 요소들은 일반적으로 정보를 구성하는 개념들을 추려낸 것이기 때문에 일반 사용자도 비교적 손쉽게 데이터 모델을 생성할 수 있다.



〈그림 1〉토픽을 정의한 사례

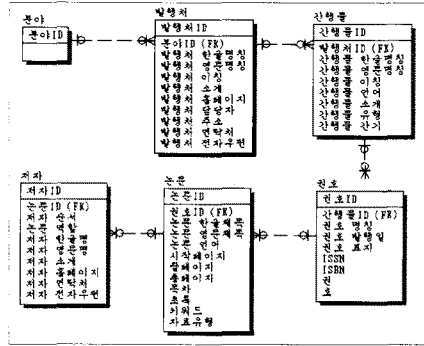
Topic Map이 정보를 표현하는 방법을 살펴 볼때, Topic Map에서 정보는 〈그림 1〉과 같은 구조로 표현된다. 어떤 특정 개념은 다른 개념과의 관계를 통해 서로 링크가 되고 이러한 기본적인 정의가 모여 의미있는 정보를 이룬다. 예를 들어 〈그림 1〉과 같이 한 번 토픽의 관계를 정의하면 '식물이 동물은 아니지만 포유류와 같은 생물'이라는 판단처럼, 정의된 의미에 기초해, 컴퓨터 또는 사용자 자신이 여러 가지 부가적인 정보를 획득할 수 있다.

Topic Map에서는 기본적으로 이러한 의미와 관계들을 Topic, Association, Occurrence라는 소위 TAO라고 부르는 세 가지 요소로 표현하며, Public Subject Identity, Scope, ResourceRef, InstanceOf 등 여러 가지 속성으로 이용해 데이터 모델을 손쉽게 작성할 수 있도록 지원한다. Topic은 일반적으로 정보의 주제를 나타내며 어떤 것이든 그 대상이 될 수 있다. 즉 사람, 사물, 개념, 의미 등 실제 존재하는 것 또는 특정 속성이나 어떤 의미 등을 표현할 수 있다. 또한 토픽에 이름과 URI, 타입 등의 속성들을 이용해 해당 주제를 명확하게 정의할 수 있으며, Association는 둘 이상의 토픽들 사이에 상하관계가 아닌 의미적인 연관성을 정의하는 관계이다. 또한 Association Type이나 Association Role이라는 속성을 통해 더 풍부한 관계를 표현할 수 있다. 그리고 각 토픽은 자신이 참조하는 하나 이상의 지식항목과 연결될 수 있다. 이러한 연결정보를 Occurrence라고 하며 문서파일, 이미지 파일, 데이터베이스내 특정 레코드 등 다양한 형태의 정보가 모두 Occurrence의 대상이 될 수 있다. 이러한 Occurrence에 의해 참조되는 자원들은 일반적으로 Topic Map과는 별도로 저장된다. 이렇게 Topic Map과 실제 지식 자원 사이를 분리할 수 있다는 것은 실제 지식 항목의 이동 없이 이분화된 구조로 지식을 분류하고 색인할 수 있는 또 다른 장점을 낳는다.

III. ERD분석을 통한 Topic Map 구축

일반적으로 어떠한 시스템을 완전히 새로 구축하는 것 보다는 기존에 존재하는 것을 최대한 재활용하여 이용하는 것이 효율적인 방법이다. Topic Map을 작성할 때도 모든 요소들을 새로이 작성하는 것 보다는 작성하려는 분야에 대해 기 구

축된 정보자원이 존재할 경우 이를 최대한 재활용하여, 모든 요소들을 추출한 다음 Topic Map으로 변환하고 이용하는 것이 시간과 비용을 절약할 수 있는 효율적인 방법일 것이다.



〈그림 2〉일반적인 학술정보자원의 논리 ERD

본 논문에서는 자료흐름도 및 ERD를 이용하여 분석된 데이터베이스를 Topic Map으로 변환하고자 할 경우, 기 작성된 자료흐름도 및 ERD를 최대한 활용하여 일관성 있게, 객체의 자료구조 및 오퍼레이션이 보다 충실하게 표현된 Topic Map 모델을 단시간에 효율적으로 생성할 수 있어야 한다는 것 등을 Topic Map 모델로의 변환방법 생성시의 주요 고려사항으로 간주하였다. 이에 따라 본 논문에서는 기본적으로 ERD를 활용하여 객체구조 및 애트리뷰트를 식별하고, MARC를 활용하여 객체의 오퍼레이션을 식별하는 방안을 제안한다.

학술 메타데이터의 서지항목들을 관리하기위한 전자도서관 시스템의 기본적인거나 공통적인 스키마 설계 구성은 〈그림 2〉와 같다.

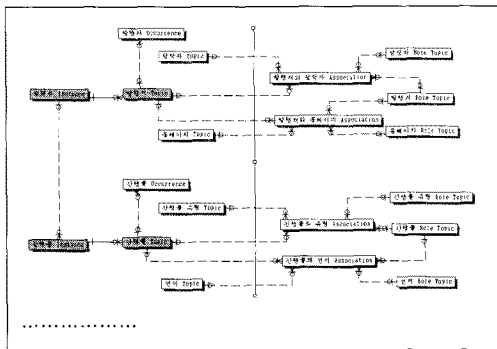
〈그림 2〉의 ERD에 대한 이해는 다음과 같이 설명될 수 있다. 학술정보자원은 발행처, 간행물, 권호, 논문, 저자 등의 개체가 있으며, 논문개체의 속성은 한글제목, 영문제목, 논문언어, 시작페이지, 끝페이지, 총페이지, 목차, 초록, 키워드, 원문자료유형 등이 있으며, 논문ID는 논문개체의 키 속성이다. 간행물 개체의 속성은 간행물명칭, 언어, 소개, 유형, 간기 등이 있다. 그리고 간행물ID는 간행물 개체의 기본 키 속성이고, 발행처ID는 간행물 개체의 외래키이다. 발행처는 간행물을 출판하는 관계를 형성한다. 출판관계에는 권, 호 속성이 존재한다. 간행물과 권호는 1:N 관계이다. 일반적으로 가장 많이 사용되고 있는 관계형 데이터베이스는 테이블, 열/행, 기본키, 외래키 등의 요소들로 구성되어 있으며, 이러한 요소들의 분석을 통해 변환 가능한 요소들을 찾아 본다면, 예를 들어 A라는 테이블이 의미하는 바가 B라는 토픽 타입과 일치하고, A 테이블 내의 필드들이 이름, 키 등의 정보를 담고 있다면, 이를 통해 B에 속하는 인스턴스, 토픽들을 자동 생성할 수 있을 것이다. 본 연구에서 진행한 실험방법에 대한 관계형 데이터베이스의 요소와 Topic Map 개체들의 매핑은 〈표 1〉과 같다.

〈표 1〉 관계형 데이터베이스와 Topic Map의 매핑

방법1	
관계형 데이터베이스	Topic Map
테이블	Instance
레코드	Topic간의 Association
필드	Topic 또는 Occurrence
기본키	Topic ID
외래키	Instance 또는 scope

방법2	
관계형 데이터베이스	Topic Map
테이블	Instance 또는 Topic Type
레코드	Topic
필드	Instance 또는 Topic 또는 Occurrence
기본키	Topic ID
외래키	Association

〈표 1〉의 방법1은 엔티티간에 관계를 부모와 자식간의 관계 또는 트리구조로 간주하였을 때 설계할 수 있는 방법이다. 하지만 콘텐츠의 내용이 가지고 있는 복잡한 관계의 개념구조, 자료유형 및 자료간의 의미적 상관관계를 고려하기 위해 방법2를 구상하였다. 방법1은 발행처, 간행물, 권호, 논문, 저자 엔티티가 가지는 속성들을 어떠한 토픽에 대한 부가정보로 축소시킴으로써 지나치게 단순화하는 결과를 초래한다. 즉 Topic Map을 표현할 때 사용하는 XTM으로 기술하는데 아무런 문제가 없을 지라도, 시맨틱하지 않다면 굳이 Topic Map으로 구축할 이유도 없을 것이다.



〈그림 3〉 학술정보자원에 대한 Topic Map식 논리 ERD

〈그림 3〉은 〈그림 2〉에 대한 Topic Map식 논리 ERD로서 〈표 1〉의 방법2를 구체화하였다. 〈표 1〉의 방법2와 같이 발행처, 간행물, 권호, 논문, 저자 엔티티는 인스턴스로 기술하고 외래키를 〈instanceOf〉로 기술하지 않고 〈association〉로 기술하여 〈topicRef〉으로 참조하도록 하여야 한다. 그리고 각 엔티티의 속성들을 토픽과 어커런스로 구분하여 토픽

들의 의미적 상관관계를 기술할 수 있도록 하여야 한다. 이때 어커런스 타입은 테이블의 필드와 매핑하여 자동생성하게 되는데, 엔티티와 관련된 속성 중에는 특정한 값만을 갖는 경우가 있다. 즉 하나의 토픽 내의 속성 값으로만 의미가 있고, 이 속성 값이 다른 토픽과의 관계를 의미하지 않는 경우는 어커런스로 정의할 수 있다. 의미적 상관관계는 〈association〉으로 기술하게 되는데 하나의 레코드를 구성하는 필드들 중 토픽으로 선언되는 것에 한하여 관계를 기술하는데, 이는 자료흐름도를 통하여 구체화할 수 있다. 이후 작성된 XTM을 그래픽 네비게이션으로 시각화하였을 때 〈instanceOf〉으로 기술된 것은 하나의 토픽을 선택하였을 때 노드가 단계별로 확장되어 가는 기준이 되고 〈association〉으로 기술된 것은 각 단계에서 연관된 토픽들이 보여지는 기준이 된다.

〈표 2〉 Topic Map 명세서

엔티티	속성	Key	Null	Topic Map
발행처	발행처 ID	PK		발행처 Topic ID
	분야 ID	FK		Instance간 계층으로 표현
	발행처 한글 명칭			발행처 Topic
	발행처 영문 명칭	Null		발행처의 Occurrence
	발행처 이칭	Null		발행처의 Occurrence
	발행처 소개	Null		발행처의 Occurrence
	발행처 담당자	Null		담당자 Topic
간행물	간행물 ID	PK		간행물 Topic ID
	발행처 ID	FK		Instance간 계층으로 표현
	간행물 한글 명칭			간행물 Topic
	간행물 영문 명칭	Null		간행물의 Occurrence
	간행물 언어			언어 Topic
	간행물 소개	Null		간행물의 Occurrence
	간행물 유형			간행물의 Topic
권호	권호 ID	PK		권호의 Topic ID
	간행물 ID	FK		Instance간 계층으로 표현
	권호 명칭			권호 Topic
	권호 발행일			발행일 Topic
	ISSN	Null		권호의 Occurrence
	ISBN	Null		권호의 Occurrence
논문	논문 ID	PK		논문의 Topic ID
	권호 ID	FK		Instance간 계층으로 표현
	논문 한글 제목			논문 Topic
	논문 영문 제목	Null		논문의 Occurrence
	논문 작성 언어			언어 Topic
	페이지	Null		논문의 Occurrence
	목차	Null		논문의 Occurrence
	초록	Null		논문의 Occurrence
	키워드			키워드 Topic
	저자	저자 ID	PK	
논문 ID		FK		Instance간 계층으로 표현
논문역할				논문역할 Topic
저자명				논문 Topic
저자 소개		Null		논문의 Occurrence

〈표 2〉는 〈그림 3〉에 대한 Topic Map 명세서이며, 먼저 topic 타입과 occurrence 타입의 요소를 나열해 보았다. 주로 속성이 Null값이 허용되면 Occurrence 타입으로, Not Null이면 Topic으로 매핑되는 결과를 보였으며, association 타입은 데이터베이스 테이블의 한 레코드에 대한 필드들 중에서 Topic으로 선언한 것들의 관계로 매핑 하였다. topic 타입과 topic 타입과의 관계 설정 시에는 각 topic의 역할을 role 타입으로 추출하여야 한다. role 타입은 데이터베이스 테이블 상에 구분자로 쓰이는 속성이거나, 공통코드로 관리하는 속성일 수도 있지만, topic 간의 의미적 해석에 따라 계속해서 추가될 수도 있다.

IV. 학술정보자원의 내용기반 지식 표현

〈topic〉은 일반적으로 어떠한 것이든 될 수 있지만 자동구축을 위해 학술정보자원에 대한 토픽은 〈표 2〉에서 언급한 토픽 요소를 기반으로 분야, 발행처, 학술지, 권호, 논문, 저자, 키워드를 토픽 타입으로 분류한다. 그리고 이들 토픽들은 다른 토픽들의 실제 인스턴스가 된다. 또한 오직 〈occurrence〉가 아닌 〈topic〉으로 구현된 정보자원을 만이 이름, 어커런스, 다른 토픽과의 관계 등의 속성을 가질 수 있다는 사실을 항상 인식해야 한다. 즉 〈표 2〉의 논문 엔티티를 예로 XTM을 작성한다면 〈그림4〉와 같다.

```

<topic id="DS_article">
  <instanceOf><topicRef xlink:href="#Field245_a"/></instanceOf>
  <baseName><baseNameString>논문</baseNameString></baseName>
</topic>
<topic id="art_name">
  <instanceOf><topicRef xlink:href="#DS_article"/></instanceOf>
  <baseName><baseNameString>논문제목</baseNameString></baseName>
</topic>
<topic id="ara_cont">
  <instanceOf><topicRef xlink:href="#DS_article"/></instanceOf>
  <baseName><baseNameString>초록</baseNameString></baseName>
</topic>
.....
<topic id="DS_issue">
  <instanceOf><topicRef xlink:href="#Field773▼d"/></instanceOf>
  <baseName><baseNameString>권호</baseNameString></baseName>
</topic>
<topic id="is_name">
  <instanceOf><topicRef xlink:href="DS_issue"/></instanceOf>
  <baseName><baseNameString>권호명</baseNameString></baseName>
</topic>
  
```

〈그림 4〉 논문 엔티티에 대한 XTM작성 예

〈그림4〉의 Topic ID인 "DS_article"은 테이블명이며, Topic ID "art_name"은 논문제목에 대한 영문필드명이다. 그리고 테이블명은 해당 필드의 〈instance〉가 된다.

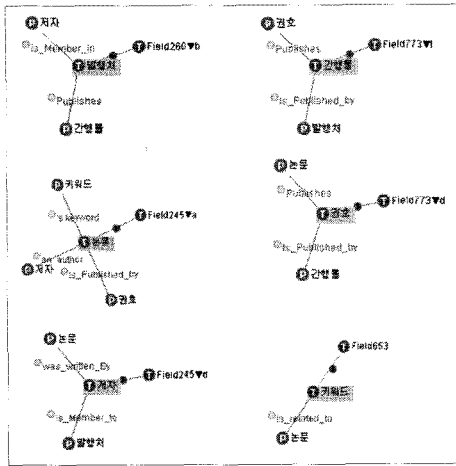
〈association〉는 토픽과 토픽을 연결시켜 주는 관계를 표현한 것으로 둘 이상의 토픽들 사이에 상하관계가 아닌 의미적인 관계성을 정의한다. 토픽들이 토픽 타입이라는 것으로 분류가 되는 것처럼 토픽들 간의 연관관계도 연관관계 타입

이라는 것으로 분류된다. 아래 〈표 3〉의 "member of", "Has language", "Is related by", "written by" 등이 연관관계 타입이다. 즉 예를 들어 "논문A" 라는 논문이 "Thomas"에 의해 쓰여 졌다면, 다른 논문 도 다른 사람에게 의해 쓰여 졌다는 관계를 설정할 수 있는 것이다. 그리고 트리구조에서 벗어난 Map이라는 특징을 살리기 위해 본 연구에서는 토픽들을 조회할 때 "발행처-간행물-권호-논문-저자"의 순서로 끝나지 않고 계속 순환할 수 있도록 저자와 발행처를 연결함으로써 〈표3〉과 같이 "발행처-간행물-권호-논문-저자-발행처"와 같은 토픽간의 어소시에이션을 구성하였다.

〈표 3〉 토픽간의 어소시에이션

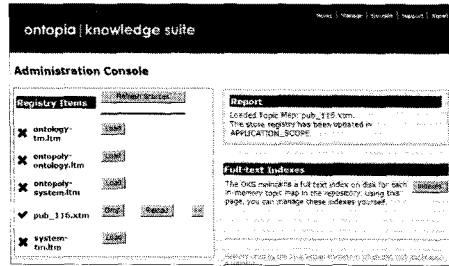
Type	내용	Topic 1	Role 1	Topic2	Role 2
Is related to	논문A는 Web2.0과 관련이 있다.	논문A	논문	Web2.0	키워드
was written by	논문A는 Thomas에 의해 저술되었다.	논문A	논문	Thomas	저자
is Member in	Thomas는 발행처A의 회원이다.	Thomas	저자	발행처A	발행처
Is Published by	발행처A는 간행물A를 출판한다.	발행처A	발행처	간행물A	간행물
belong to	간행물A의 3권 2호다.	간행물A	간행물	3권 2호	권호
Is Produced by	3권 2호에 논문A가 배정되었다.	3권 2호	권호	논문A	논문

위 〈표 3〉는 논문 메타데이터 사이의 의미적 연관성을 각각의 토픽에 대한 어소시에이션을 정의한 표이다. 토픽들은 이름, 어커런스, 그리고 다른 토픽과의 연관관계에서의 역할이라는 3가지 종류의 성격을 가진다. 연관관계에는 타입 외에도 "역할"이라는 중요한 정보가 있다. "Thomas는 발행처A의 회원이다." 라는 예를 보면, Thomas는 회원이라는 역할을 가지고 발행처A는 발행처라는 역할을 가진다. 이 말은 연관관계에 어떤 제약조건을 줄 수 있다는 것을 의미한다. 즉, "member of" 라는 연관관계에서는 회원 역할의 토픽과 발행처 역할의 토픽이 있어야 하며 나오는 순서도 저자가 발행처보다 먼저 와야 한다는 것이다. 이때 MARC 필드들 사이의 연관성은 그 자체가 하나의 중요한 사상이고 개념일 수 있다. 다른 예로 논문과 키워드 같은 경우는 한 단어로 설명되지 않는 그 자체가 여러 페이지에 나누어 설명되는 중요한 개념일 것이다. 이 경우 "Is related to"와 같은 일반적 연관성으로 규정하기 어려우므로 연관 관계 이 외에 두 개념 사이의 관계를 설명하는 여러 개념들을 추출한 다음 개념들 사이에 구조를 정의해야 한다. 아래의 〈그림 5〉는 위의 표에서 정의한 관계를 바탕으로 논문과 저자명, 키워드, 권호, 간행물, 발행처를 토픽으로 선정하고, 모델링 한 것을 토픽맵 형식으로 표현한 것이다.



〈그림 5〉 모델링한 토픽간의 어시메이션

먼 두 토픽을 병합하고 새로운 관계를 생성해 최종적으로 변경된 Topic Map을 생성하여야 한다. 본 연구에서는 Topic간의 중복을 제거하기 위하여 DB에 Topic 개체를 추가하여 Topic이 생성될 때 마다 누적된 Topic이 중복되는지를 체크하였다.



〈그림 6〉 Omnigator 관리 콘솔

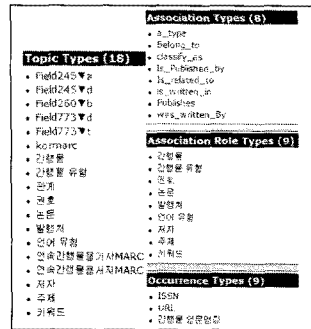
V. Topic Map 프로토타입 구현

Topic Map 기반의 학술정보자원을 ERD로부터 자동 생성하려면 일정한 알고리즘에 의해 생성되어야 하는데 생성절차에 대한 알고리즘은 〈표 4〉와 같다.

〈표 4〉 Topic Map 자동 생성 알고리즘

구분	단계	설명
단계 1	빈 Topic Map을 만든다.	등록 절차 시작
	토픽 생성	ERD상의 모든 개체를 처리할 때 까지 1-5단계를 반복한다.
단계 2	토픽에 basename 추가하고 인스턴스 부여	토픽들의 중복 제거
단계 3	토픽 간의 계층 형성	토픽간의 superclass-subclass 부여
단계 4	어커런스 생성	어커런스에 resourceData, resourceRef 추가
	어커런스에 인스턴스 추가하고 topicRef 추가	어커런스에 인스턴스 추가하고 topicRef 추가
단계 5	Association 생성	Association에 인스턴스 추가하고 topicRef 추가
	Association 생성	Association에 멤버 추가
	Association 생성	멤버에 해당하는 토픽의 role 추가

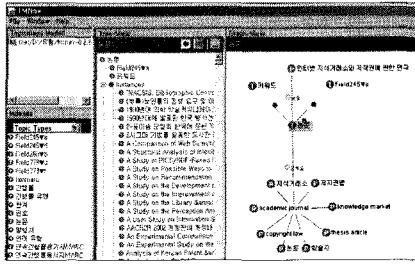
〈그림6〉은 자동생성 알고리즘을 기반으로 개발한 Topic Map 변환 모듈로서 자동변환한 XTM을 검증하기 위하여 오픈소스인 "Ontopia.net"에서 배포하는 Omnigator에 Topic Map을 등록하는 관리콘솔 화면이다. 토픽맵에서 토픽은 하위 토픽의 포함 여부에 따라 클래스 토픽(class topic)과 인스턴스 토픽(instance topic)으로 나눌 수 있는데 클래스 토픽은 토픽 계층 구조에서 'is-a' 또는 'part-of' 와 같은 관계로 연결된 하위 토픽을 가지는 토픽이고 인스턴스 토픽은 하위 토픽을 가지지 않는 말단 토픽을 말한다. 또한 참조 유형에 따라 토픽 타입(topic type), 어커런스 타입(occurrence type), 연관관계 타입(association type), 역할 타입(role type)의 4가지 유형으로 나누어진다. 〈그림7〉은 Omnigator에서 자동변환된 XTM을 등록한 결과화면이다.



〈그림 7〉 XTM이 등록된 결과

중요한 점은 2단계에서 Topic을 생성한 후 이 Topic들을 효율적으로 통합하는 중복제거 작업이다. 새롭게 생성된 토픽과 기존의 Topic이 같은 내용으로 존재할 경우 실제로 두 지식간에 밀접한 연관이 있는데도 관계를 맺지 못해서 무의미하게 동떨어지게 될 것이다. 이를 위해 토픽을 구분하는 요소 중 토픽을 식별하는데 추가 되는 <subjectIdentity>와 <baseName>을 이용하여 토픽과 해당 토픽이 참여하는 Association 등을 대조해 동일한 속성 값을 지녔다고 판단되

여기서 사용된 실험데이터는 1개 학회를 선정하여 XTM으로 자동변환하였다. 이 학회에서 실험에 사용한 논문은 1253개 였다. 자동변환시 약 10분이 걸렸으며 Omnigator에서 변환한 XTM을 등록할때도 약 10초 정도 걸렸으나 "논문 Topic Type"에서 상세화면으로 이동하기에는 약간의 부하가 생겼다. 아래 〈그림8〉은 TM4J Engine을 사용한 TMNav라는 토픽맵 네비게이션에 자동변환한 XTM을 등록한 화면이다.



(그림 8) TMNave에 등록된 결과

VI. 결론

이미 구축된 학술DB로부터 Topic Map에 재촬영할 수 있는 요소들을 추출하기 위해 상세하게 분석해 봐야 할 정보 스스로 데이터베이스 스키마로 방향을 잡았다. 학술정보자원은 발행처, 간행물, 권호, 논문, 저자로 한정하고 Topic Map을 자동구축하기 위하여 매핑표를 작성한 후 자동생성 알고리즘을 기반으로 개발한 Topic Map 변환 모듈로 자동변환하였다. 그리고 이를 검증하기 위해 구축된 Topic Map을 오픈소스인 Omnigator와 TMNave에 등록해보았다. 이때 각각의 토픽은 정보 검색의 검색 포인트가 되고, 이것은 관련된 정보만을 연결하고 있기 때문에 정보의 필터링이 용이하여, 자신이 원하는 목적에 따라서 정보를 네비게이션 할 수 있어서 검색의 결과를 얻기까지의 과정이 결과내 재검색을 해야하는 현재의 검색방법과는 달랐다. 또한 자동변환된 Topic Map이 해당분야의 전문가가 이해하는 수준의 학술정보자원의 메타데이터간의 의미적 상관관계를 표현하기는 다소 미흡하다 할지라도 구축된 DB로부터 빠르게 Topic을 구축할 수 있는 방법이다.

참고문헌

- [1] 한국 더블린코어 DCMI "더블린코어 메타데이터에서 서지인용정보의 입력지침", 2005, URL: "http://www.dublincore.go.kr/documents/uidelidc-gcitationnes/"
- [2] The Library of Congress "MARC21 XML Schema", 2007, URL: "http://www.loc.gov/standards/marcxml/"
- [3] 김규훈, 신성기, "웹 학술정보자원 DB 구축을 위한 메타데이터 요소 분석에 관한 연구", 한국정보관리학회, 한국정보관리학회 제9회 학술대회 논문집, 2002. 8, pp. 123-130
- [4] 이응봉, 류범중, "해외 학술정보 메타 데이터베이스 구축 및 서비스에 관한 연구", 한국문헌정보학회, 한국문헌정보학회 학술발표논문집 제13집, pp. 23-36, 2002. 4.
- [5] 정도현, 이상환, 신기정, "이기종 학술정보 분류체계간 상호운영에 관한 연구", 한국콘텐츠학회, 한국콘텐츠학회 종합학술대회 논문집 제16권 제2호(상), pp. 360-364, 2007. 11.
- [6] 장동춘, "디지털대학교도서관 설계를 위한 이론적 고찰", 한국도서관정보학회, 한국도서관정보학회지, vol.30, no.1 pp.119-150, 1999.

- [7] Ahmed, Kal. "Topic Map Design Patterns For Information Architecture. XML", 2003, URL: "http://www.techquila.com/tmsinia.html".
- [8] ISO/IEC. 2002. ISO/IEC 13250: Topic Map. ISO/IEC.
- [9] Steve, "The TAO Topic Maps" 2005, URL: "http://www.gca.org/papers/xml europe2000/papers/s11-01.html".
- [10] TopicMaps.org Authoring Group, "XML Topic Maps(XTM) 1.0 Specification", 2001, URL: "http://www.topicmaps.org/xtm/1.0/index.htm"
- [11] Marut Buranarach. The Foundation for Semantic Interoperability on the World Wide Web, 2001.
- [12] Amit P. Sheth, *Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics.*
- [13] Erhard Rahm and Philip A. Bernstein, "A survey of approaches to automatic schema matching," VLDB, vol. 10, issue 4, pp. 334-350, 2001.
- [14] Graham Moore. "Topic Map technology - the state of the art," XML 2000 Conference & Exposition, Washington, USA, December 2000.
- [15] T. Berners-Lee. "Semantic Web Roadmap," <http://www.w3.org/DesignIssues/Semantic.html>.
- [16] T. Berners-Lee, J. Hendler, O. Lassila. "The Semantic Web," Scientific American, May 2001.
- [17] Bernstein, A., Klein, M. : Towards High-Precision Service Retrieval. Proceedings of The International Semantic Web Conference(2002) 84-101
- [18] Steve P. : The TAO of Topic Maps : Finding the Way in the Age of Information. Proceedings of XML Europe 2000, GCA(2000).
- [19] Garshol, L.M. (2004). Metadata? Thesauri? Taxonomies? Topic maps! Ontopia. Available at: <http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>
- [20] 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.
- [21] Pepper, S. 2002. The tao of topic maps. Oslo: Ontopia. Available at: <http://www.ontopia.net/topicmaps/materials/tao.html>
- [22] Martin Lacher and Stefan Decker. "On the integration of Topic Map data and RDF data," Extreme Markup Languages 2001 Conference & Exposition, Montréal, Canada, August 2001.
- [23] Modelling Distributed Knowledge Management System with Topic Maps. Thomas Schwotzer, 2004