

X-TOP: 레거시 시스템상에서 온톨로지 구축을 위한 토픽맵 플랫폼의 설계와 구현

(X-TOP: Design and Implementation of TopicMaps Platform for Ontology Construction on Legacy Systems)

박 여 삼 [†] 장 옥 배 ^{**} 한 성 국 ^{***}
(Yeo-Sam Park) (Ok-Bae Chang) (Sung-Kook Han)

요 약 토픽맵은 기존의 온톨로지 언어와는 달리 위치정보를 이용하여 형태를 변환하지 않고도 많은 양의 이종 정보 자원을 통합할 수 있는 온톨로지 언어이다. 지금까지 토픽맵 구축을 위한 여러 편집기등이 개발되어 있으나, 이들은 XTM 문서를 독자 방식으로 처리하고 있다. 따라서 대용량 자료를 처리하는데 많은 시간이 소요되고, 현재 RDB기반에서 운영되고 있는 레거시 시스템에 적용하여 실용화하기에는 많은 문제가 있다.

본 논문에서는 XTM 1.0 규격 기반의 대용량 토픽맵을 RDB 구조로 모델링하여, 처리시간을 최소화하고 레거시 시스템 상에서 온톨로지 구축이 가능하도록 하였다. 기존에 사용하고 있는 SQL 도구와 어플리케이션 개발 도구를 토픽맵 온톨로지 구축에 활용할 수 있도록 하여, 온톨로지 구축의 효율성을 높이고 XTM 문서와 데이터베이스간의 상호 호환이 가능한 토픽맵 플랫폼 X-TOP을 설계하고 구현하였다. X-TOP은 향후 사용자 인터페이스 변경과 다양한 DBMS를 지원할 수 있도록 3계층 구조를 갖고 있다.

웹스캐어의 압 온톨로지 관리에 X-TOP을 적용하여 기존 시스템과의 성능 비교와 실무 응용의 유효성을 보였다.

키워드 : 지식표현, 온톨로지, 토픽맵, 대용량온톨로지, XTM

Abstract Different from other ontology languages, TopicMap is capable of integrating numerous amount of heterogenous information resources using the locational information without any information transformation. Although many conventional editors have been developed for topic maps, they are standalone-type only for writing XTM documents. As a result, these tools request too much time for handling large-scale data and provoke practical problems to integrate with legacy systems which are mostly based on relational database.

In this paper, we model a large-scale topic map structure based on XTM 1.0 into RDB structure to minimize the processing time and build up the ontology in legacy systems. We implement a topic map platform called X-TOP that can enhance the efficiency of ontology construction and provide interoperability between XTM documents and database. Moreover, we can use conventional SQL tools and other application development tools for topic map construction in X-TOP. The X-TOP is implemented to have 3-tier architecture to support flexible user interfaces and diverse DBMS.

This paper shows the usability of X-TOP by means of the comparison with conventional tools and the application to healthcare cancer ontology management.

Key words : Knowledge Representation, Ontology, TopicMaps, Large-scale Ontology, XTM

[†] 정 회 원 : 전북대학교 컴퓨터과학과
maronie61@gmail.com
^{**} 중신회원 : 전북대학교 컴퓨터과학과 교수
okjang@chonbuk.ac.kr
^{***} 중신회원 : 원광대학교 컴퓨터공학 교수
skhan@wku.ac.kr
논문접수 : 2008년 1월 10일
심사완료 : 2008년 3월 4일

Copyright©2008 한국정보과학회 : 개인 목적이나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지 : 컴퓨팅의 실제 및 레터 제14권 제2호(2008.4)

1. 서론

차세대 웹 기반 기술은 현재와 같이 사람을 위한 웹 환경이 아니라 컴퓨터가 자동으로 처리할 수 있는 웹 환경을 제공하기 위한 것이다. 기존 웹이 단순히 HTML 기반으로 사용자에게 브라우징 할 수 있도록 하고, 서버 측에 있는 프로그램을 실행하여 그 결과를 볼 수 있도록 하는 사용자 중심의 환경이었다. 하지만 차세대 웹 기반 기술이 도입되면 현재 사용하고 있는 인터넷의 표준 프로토콜을 이용하여 원격지에 있는 웹 객체를 XML 기반으로 접근/이용/재사용 할 수 있는 웹 환경의 분산 컴포넌트 모델을 제공한다. 또한 웹 환경의 응용 프로그램 사이에서 웹 기반 지식 처리 및 공유가 가능하게 될 것이며, 이를 위해 시맨틱 웹 기술이 필요하게 되었고 이러한 시맨틱 웹을 구현하고자 온톨로지(ontology) 개념이 등장하게 되었다.

온톨로지는 특정도메인 지식을 개념화(conceptualization)하여 형식적(formal)이고 명시적(explicit)으로 표현한 공유된 개념체계로서, 개념에 대한 일의적 해석과 의미적 상호 운용성을 제공한다. 또한 온톨로지는 의미태그(semantic tag)를 사용하는 XML의 등장으로 간편하고 효과적으로 표현할 수 있게 되었다[1]. 그러나 XML은 단순히 정보의 의미 모델링만을 제공하고 개념의 특성이나 상호관계를 표현 하는 데는 미흡하였다. 온톨로지가 지식기반시스템의 기반 요소가 됨에 따라 정확하게 개념화 구조를 표현하기 위하여 온톨로지 전용 언어가 필요하게 되었고, RDF/RDFS, DAML, OWL, TopicMaps 등의 표준 언어가 단계적으로 개발되었다[2-4]. 이러한 온톨로지 언어를 이용해서 도메인의 지식 정보를 표현할 수 있게 됨에 따라 웹은 점차 지식기반의 시맨틱웹(Semantic Web)으로 진화하고 있으며, 새로운 응용시스템 개발에 토대가 되고 있다. 이중 의료, 생명과학 분야에서 온톨로지 프로젝트가 제일 활성화 되어 있는데, 분야 특성상 개념과 개념들 사이에 대한 관계 정보 컨텐츠가 제일 많이 누적되어져 왔으며, 이러한 온톨로지 컨텐츠의 활용 효과가 상대적으로 가시적이기 때문이다.

많은 온톨로지 언어가 제안되고 있지만, XTM은 ISO 국제 표준으로 토픽간의 관계 중심 모델을 기반으로 하고 있기 때문에, 다른 온톨로지 언어에 비하여 활발하게 응용되고 있다. 본 논문의 목적은 토픽맵을 구축하는 틀을 개발함에 있어 기존의 문제점들을 해결하기 위한 방안을 제시하고자 한다. XTM 1.0 DTD의 개념 모델을 관계형 데이터베이스 스키마로 구성하고, 토픽맵 온톨로지 모델을 관계형 데이터베이스 내 각 테이블로 저장하여, XTM 문서를 효과적으로 데이터베이스화 하는 방법을 개발하였다. 데이터베이스에서 기존 데이터와 XTM

문서가 상호 호환이 가능하고, RDB의 SQL을 XTM 문서에서 그대로 활용하는 방법을 보였다.

토픽맵은 지식층과 정보층의 이중 구조를 나타내는데, 지식층은 상위 계층으로 토픽(topic)과 토픽간의 연계(association)로 구성된다. 토픽은 특정 주제를 나타내는 표현이고, 연계는 주제들 간의 관계를 나타내는 표현이다[5]. 이러한 토픽맵 온톨로지 응용을 위해서, 토픽맵 구축 및 수정이 가능하고 검색 및 네비게이션 할 수 있는 TMDesign, Pro-TM Editor 등의 토픽맵 편집기가 개발되어 있다[6,7]. 그러나 현재 사용되어지고 있는 토픽맵 구축 도구들은 XTM 문서를 직접 접근하여 처리하는 독자적(stand-alone) 방식이 대부분이다. 이러한 방식은 XTM 문서만 처리할 수 있으나, XTM 문서 처리에 장시간이 요구되며 대용량 토픽맵 자료를 처리할 수 없는 문제점이 있다[8]. 또한 일부 데이터베이스를 지원하더라도 데이터베이스에서 XTM 문서를 생성하지 못하는 한계점이 있어, 본 논문에서 이를 위한 새로운 해결 방법을 제시한다.

본 논문에서 제안하는 토픽맵 플랫폼 X-TOP은 RDB 구조에서 직접 기존의 토픽맵 편집기들이 제공하는 기본적인 편집기능과 토픽에 대한 정보를 시각적으로 보여줄 수 있도록 사용자 친화적인 시각화 기능과 멀티유저를 위한 Lock과 Unlock 기능을 제공하고 있으며, XTM 문서를 데이터베이스로 импорт(import)할 수 있는 파서 기능과 데이터베이스에서 XTM 문서로 익스포트(export)할 수 있는 생성기능도 구현하였다. 헬스케어 암 관리 온톨로지를 구축하여 구현된 시스템의 유효성을 보였다.

본 논문의 전체적인 구성은 다음과 같다. 2장에서는 토픽맵 개념을 서술하고 기존 모델들의 사례와 문제점들을 기술한다. 3장에서는 본 논문에서 구현한 X-TOP 토픽맵 플랫폼의 시스템 아키텍처 및 데이터베이스 모델링 등의 설계와 구현 그리고 특징에 대하여 서술하며, 4장에서는 구현된 시스템의 결과에 대한 적용 사례와 기존 도구들과의 기능을 비교하고, 마지막 5장에서는 결론과 향후 연구에 대하여 기술한다.

2. 관련 연구

토픽맵은 2000년에 ISO/IEC 13250 국제표준으로 채택된 온톨로지 구축 모델이다. 처음 ISO/IEC 13250에서 제안한 토픽맵 표준 규격은 SGML(Standard Generalized Markup Language)구조와 HyTM 언어였으나, 2001년에 TopicMaps.Org에서 개발한 XTM(XML TopicMaps)을 통합하면서 지금은 XTM 1.0이 표준 규격으로 자리잡고 있다. 또한 초기에는 서적의 인덱스(index), 용어(glossary), 시소러스, 목차 등의 기술을 위한 도구

로 개발되었지만, 온톨로지 개념이 대두되면서 그 응용 영역이 크게 확대되었다. XTM 1.0은 도메인의 지식정보를 온톨로지 형식으로 모델링하여 지식정보의 공유와 상호 교환을 실현해 주는 국제 표준 언어이다[9]. XTM 1.0은 RDF/S, OWL 등과 더불어 시맨틱웹을 실현하는 방법을 제공하며, 관계형 데이터베이스의 정보 표현에도 활용된다. 데이터베이스는 정보객체사이의 관계만을 표현하지만, XTM 1.0은 정보객체가 존재하는 다양한 위치를 연결할 수도 있다[7]. 즉 개념간의 관계뿐만 아니라 개념과 정보자원을 연결하여 지식베이스(knowledge base)를 구축하여 준다. 이것은 XTM 1.0이 기존 또는 앞으로 연구 개발되어질 지식 표현 방식까지 적극적으로 수용할 수 있도록 설계되어 있기 때문이며, XTM 1.0은 이처럼 광대한 범위의 지적 표현에 적용할 수 있는 특성이 있다[9,10].

토픽맵을 표현하는 기술 언어로 XTM이 있으며, XTM은 토픽맵 모델의 각 요소를 나타내는 태그 집합과 엘리먼트들 사이의 구조를 정의하고 토픽맵을 기술하는 표준형식으로 사용되고 있다[11]. 현재 토픽맵의 표준연구는 표 1과 같이 진행 중에 있으며, 이중 현재 가장 활발히 연구가 이루어지고 있는 TMCL(ISO 19756:Topic Map Constraint Language)은 토픽맵을 위한 질의 언어 표준으로 CD(committee draft) 단계에 있고, TMQL(ISO 18048: Topic Maps Query Language)은 토픽맵 질의 언어 표준으로 역시 CD 단계에 있다.

2.1 토픽맵 개념

토픽맵은 주제를 실체화하는 <topic>을 기본요소로 하여, <topic>간의 관계를 설정하는 <association>, <topic>의 연관 정보자원을 결합하는 <occurrence>등 3개의 중요요소로 구성된다. <topic>에 다양한 명칭을 부여하기 위한 명명법, <topic>의 개념 범위를 한정하는 <scope>, <topic>간의 관계정의 시 <topic>의 역할을 규정하는 <role> 등 3가지 보조 개념도 사용된다.

- 토픽맵의 기본 구성요소 : <topic>

주제를 컴퓨터가 저장, 질의, 처리할 수 있는 형태로 실체화 한 것을 <topic>이라 하며 XTM 1.0을 구성하는 기본단위이다. 구성된 토픽맵의 개념요소는 모두 <topic>이 되며, <topic>은 주제의 실체화, 토픽의 형(type) 또는 클래스 선언, association 선언, occurrence의 원천, association의 멤버(member)와 역할(role) 정의, 명칭 파라메타 정의, scope 선언 그리고 다른 토픽맵과 통합하기 위한 토픽 등 다양한 형태로 활용된다.

- 토픽맵의 연관관계 : <association>

<association>은 토픽간의 관계를 부여하는 지식정보 모델링의 필수 도구이다. XTM 1.0에서 <association>은 방향이 없는 N-ary 관계를 형성할 수 있다. 관계에는 방향이 없기 때문에, <association>은 선언적 지식(assertion)을 정의하게 된다. <association>은 관계의 속성이나 형을 지정하는 클래스의 인스턴스가 될 수 있지만, 다중클래스의 인스턴스가 되어서는 안 된다. <association>은 관계를 구성하는 토픽들의 집합(aggregation)으로 구성되며, 토픽은 집합 내에서 역할(role)에 따라 구별된다. <association>의 멤버(member)가 될 수 있는 토픽의 형태에는 멤버 토픽을 지시하는 <topicRef>, 멤버가 되는 정보자원을 지시하는 <resourceRef>와 멤버가 되는 주제 지시자를 지시하는 <subjectIndicatorRef>의 3가지 종류가 있다.

- 정보자원 결합 : <occurrence>

토픽과 연관되는 정보자원을 결합하는 것이 <occurrence>이다. 토픽에 관한 추가적인 정보를 부가하고자 할 때는 <occurrence> 형태로 기술한다. <occurrence>는 일반적으로 <occurrence>의 속성이나 형(type)을 표시하는 클래스의 인스턴스가 될 수 있지만, 다중 클래스의 인스턴스가 될 수는 없다. <occurrence>가 하나의 클래스만을 갖기 때문에, <occurrence>의 클래스-인스턴스 관계는 속성-속성값의 구조로 사용될 수 있다. 컴퓨터가 접근가능한 자원은 모두 <occurrence>가 될 수 있고, <occurrence>는 정보자원을 지시하는 <resourceRef>

표 1 분야별 연구 현황

| 구분 | 국제표준번호 | 표준 제목 |
|------|----------------------|----------------------------------------------------------------------------|
| TM | ISO/IEC 13250:2003 | Information technology - SGML applications - Topic maps |
| TM | SO/IEC WD 13250-1 | Information technology - Topic Maps - Part 1 : Overview and basic concepts |
| TMDM | ISO/IEC FDIS 13250-2 | Information technology - Topic Maps - Data model |
| XTM | ISO/IEC FCD 13250-3 | Information technology - Topic Maps - XML syntax |
| CXTM | ISO/IEC FCD 13250-4 | Information technology - Topic Maps - Canonicalization |
| TMRM | ISO/IEC CD 13250-5 | Information technology - Topic Maps - Reference model |
| CTM | ISO/IEC WD 13250-6 | Information technology - Topic Maps - Compact syntax |
| GTM | ISO/IEC WD 13250-7 | Information technology - Topic Maps - Graphical notation |
| TMQL | ISO/IEC CD 18048 | Information technology - Topic Maps - Query language |
| TMCL | ISO/IEC CD 19756 | Information technology - Topic Maps - Constraint language |

와 정보자원의 값을 지정하는 <resourceData>의 두 형태가 있다. <resourceRef>는 XLink/XPointer 표기법으로 정보자원의 위치를 지정하는 것으로 웹상의 하이퍼링크(hyperlink)와 같은 역할을 하고, <resourceData>는 토픽에게 정보의 값을 부여하는 문자열이다. XTM 1.0에서는 정의된 데이터 형(data type)이 없기 때문에 정보의 값으로 부여된 문자열의 해석은 응용시스템에서 수행하여야 한다.

• 주제 관점 : <scope>

하나의 개념 요소(토픽맵에서는 주제)에 대하여 다양한 관점이 존재한다. 각 관점마다 기술하는 지식정보 형태가 다르기 때문에, 지식정보 모델링에서는 이러한 관점을 효과적으로 형식화할 수 있는 방법이 필요하며 XTM 1.0에서는 이를 <scope>로 모델화 하고 있다. <scope>는 토픽의 전문화된 세부개념을 정의하는데 사용할 수 있으며, 대형 토픽맵에서 불필요한 부분을 제거해 내는 여과(filter)장치로 활용할 수도 있다. 또한 토픽 명칭 충돌(name conflict)을 해결하기 위한 토픽 명명제한(naming constraint)으로도 활용된다.

2.2 기존 모델의 문제점

기존의 대표적인 토픽맵 관리 도구인 TopicMap Designer는 Topic, Association, Designer, Browser, Viewer의 Tab page로 구성되어져 있다[12]. <topic>과 <association>의 추가, 삭제, 수정의 기능을 제공하지만, 독자적 방식으로 단일 XTM 문서에 직접 접근하여 처리하는 방식으로서 대용량 토픽맵을 처리하는데 있어 저장시간, 조회시간, 관리방법 등 여러 가지 어려움이 있다. 토픽맵 관리 시스템으로 상용화된 시스템은 Ontopia의 Ontopia Knowledge Suite(OKS)와 Mondeca의 Modeca Intelligent Topic Manager(ITM)가 있으며, 이들은 토픽맵을 생성하고 저장하며 편집, 검색, 네비게이션 기능을 제공하는 것으로 지식관리 시스템 또는 콘텐츠 관리 시스템의 엔진으로 사용되고 있다[13].

하지만 XTM 문서 기반의 시스템은 저장 시 성능의 문제, 대용량 온톨로지 저장 및 관리 기능의 미흡, 다양한 관리도구 개발의 어려움, 기존 관계 데이터베이스와의 연동 미흡, 기존 데이터의 온톨로지 인스턴스화 기능 부족, 관계 데이터베이스 기술의 적용 및 활용에 어려움이 있고 기존 응용 시스템의 지식화 지원이 미흡하다. 이들 시스템의 성능 평가를 위하여, 2.5MB 용량의 '오페라 XTM 파일' (파일 내의 자료는 약10만개)을 파싱하는 동안 컴퓨터가 정지되는 현상이 나타났으며, 이들 시스템에서 대용량 토픽맵을 상용 목적으로 구축하는 것은 어렵다는 것을 발견하였다.

본 논문의 X-TOP 플랫폼은 일반 상용 DBMS 구조상에서 기존 모델의 기능을 포함하고, 대용량 토픽맵 저

장, 저장시간과 조회시간 단축, 편리한 관리방법 제공, 여러 저장소 지원, XTM 문서와 데이터베이스간의 상호호환, 외부 XTM 문서를 데이터베이스로 импорт 그리고 데이터베이스에서 XTM 문서로 익스포트할 수 있는 여러 향상된 기능까지 구현하였다.

3. X-TOP 설계 및 구현

본 장은 X-TOP 토픽맵 플랫폼의 시스템 아키텍처와 데이터베이스 모델링 그리고 이를 기반으로 한 설계 및 구현에 대해서 설명한다. 본 논문의 핵심은 XTM 1.0 표준 규격에 최적화 된 데이터베이스 모델링을 구축하고 빠른 속도로 대용량 토픽맵을 데이터베이스에 импорт할 수 있으며, 역으로 데이터베이스에서 대용량 토픽맵 문서로 생성하는 것이다. 뿐만 아니라 그래픽 뷰어의 시각화 기능을 통해 사용자가 효과적이고 손쉽게 토픽 검색 및 관리를 하고 등록된 모든 요소들에 대해서 통계정보를 제공받을 수 있도록 한다.

3.1 X-TOP 시스템 아키텍처

그림 1은 X-TOP의 시스템 아키텍처로서 XTM 기반 응용(XTM-based Application)과 데이터베이스 응용(Database-based Application)을 통합관리 할 수 있도록 X-TOP 시스템을 구현하였다.

각 기능을 보면 XTM 기반 응용은 본 연구에서 개발한 XTM Parsing API Adapter 모듈이 XTM 문서를 읽어 레코드 단위로 구분 해석하여 'TopicMap Ontology'에 저장하고, XTM File Generator는 레코드 단위로 저장된 토픽맵을 XTM File Generator 모듈을 통해 XTM 규격에 맞춰 XTM 파일을 생성한다. 데이터베이스 응용은 DBMS에 직접 접속하여 'TopicMap Ontology'의 신규 토픽맵을 레코드 단위로 생성 및 저장할 수 있고, GUI 기반으로 토픽의 추가, 삭제, 수정, 조회 기능 등 일반 SQL문을 통해서 토픽맵을 저장할 수 있는 기능을 제공한다.

3.2 X-TOP 데이터베이스 설계

X-TOP은 XTM 1.0 명세서 기반으로 설계 되었으며, 관계형 데이터베이스로 구현하였다. 데이터베이스관리 시스템(DBMS)은 Oracle을 활용함으로써 안정성과 대용량 처리에 있어서 보장받을 수 있는 기반을 마련하였으며 관계형 데이터베이스로 구축함으로써 어플리케이션 개발자와 사용자의 사용 편리성을 동시에 만족 시킬 수 있도록 설계하였다.

본 논문에서 설계한 E-R Diagram과 테이블은 개발자의 편의를 고려하여 가급적 XTM 1.0 DTD 에서 명명한 Element명과 Column명을 동일하게 사용하여 그림 2에 나타내었다.

3.3 X-TOP 시스템 구현

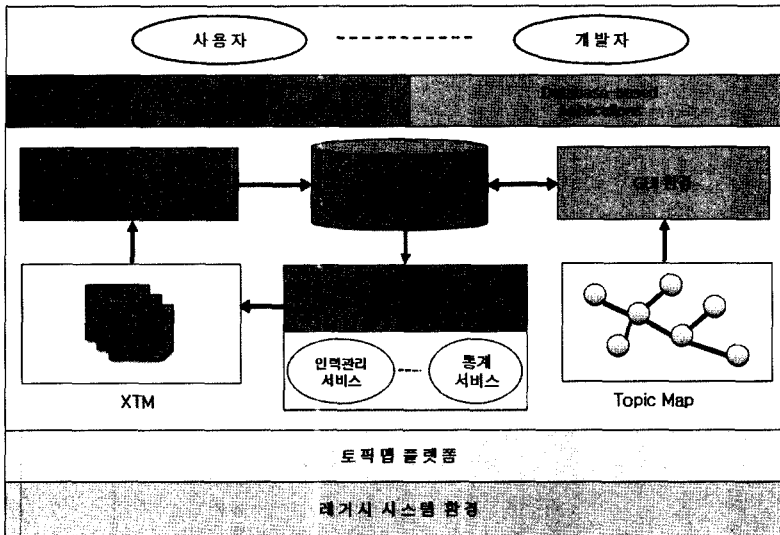


그림 1 X-TOP 시스템 아키텍처

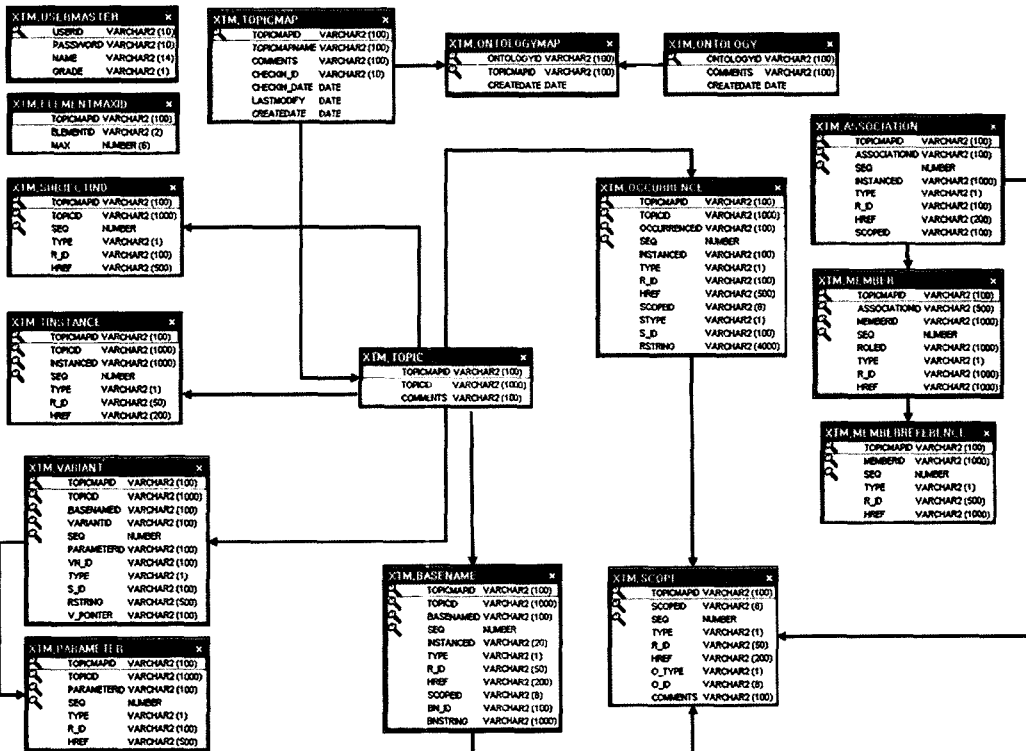


그림 2 X-TOP의 E-R Diagram

앞 절에서 설계한 내용에 따라 그림 1의 아키텍처 상에서 원활하게 구동될 수 있도록 X-TOP을 구현하였으며, 또한 X-TOP을 이용하여 누구나 토픽맵을 쉽게 접하고 실제 적용할 수 있도록 하기 위하여, 시스템 환경

을 표 2와 같은 기반으로 구현 하였다.

3.3.1 X-TOP 라이브러리 개발

본 논문에서 구현한 X-TOP의 시스템 설계는 향후 UI 변경과 다양한 DBMS를 지원하기 위해 유연한 구

표 2 X-TOP 구현 환경

| Item | Specification |
|-------------------|-------------------------------------|
| 운영체제 | Microsoft Windows XP Professional |
| 데이터베이스 | Oracle Database 10g Express Edition |
| 개발 Tool | Boland Delphi 7 |
| Graphic Component | Meta Tree Component |
| Grid Component | RealGrid Delphi 7 |
| XML Library | SimpleXML |

조를 갖는 3-계층 구조로 개발 하였다. 먼저 사용자 인터페이스를 담당하는 부분과 중간의 미들웨어 역할을 담당하는 부분, 마지막으로 데이터베이스를 직접 처리하기 위해 SQL을 담당하는 부분으로 나누어 개발 하였으며, 전체 시스템의 주요 모듈과 함수는 표 3과 같다.

3.3.2 시스템 메뉴 구성도

그림 3은 본 논문에서 구현한 시스템 메뉴 구성도로서, 주 메뉴는 File, Edit, View, Admin으로 구성되어 있다.

각 메뉴별 기능을 보면 Directory는 온톨로지 및 토픽맵을 신규 생성, 삭제, 수정 할 수 있는 기능을 담당한다. Load는 Directory에서 생성된 토픽맵과 외부 토픽맵을 X-TOP 시스템으로 적재(Load) 및 저장(Store)

하는 기능을 담당하는데, 파싱후 적재 시간을 줄이기 위해서 두 가지 형식으로 동작시킬 수 있다. 첫째는 화면으로 적재하여 작업할 수 있고 둘째는 필요시 적재할 수 있도록 하였다. Import는 Load의 기능과 유사하지만, Load에서는 외부 토픽맵을 파싱하여 Table View를 통해서 Store하는데 비하여, Import는 직접 Store하는 기능을 수행하며 Load와 마찬가지로 실무 응용의 빠른 동작을 위하여 사용자 편의성을 고려하였다. Export는 데이터베이스내의 각 테이블에 저장되어 관리되는 토픽맵 레코드를 XTM File로 생성하는 기능을 담당한다. Close XTM은 Open된 토픽맵을 종료하며, Exit는 X-TOP 프로그램을 종료하는 기능을 담당한다.

XTM Edit는 토픽맵의 <topic> 또는 <association> 등을 추가, 삭제, 수정을 할 수 있는 저작 도구기능을 제공한다. 저작자는 원하는 시점에 XTM 문서로 생성하여 볼 수 있으며, 사용자의 편의성을 고려한 Graphic View기능이 제공된다. XTM Delete는 Open된 토픽맵을 데이터베이스에서 삭제하는 기능을 수행하고, Table은 활성화된 토픽맵(X-TOP 시스템에 등록되어 Navigator를 통해서 Open된 토픽맵) 구조를 관계 데이터베이스의 테이블 스키마 형태로 표시한다. Navigator는

표 3 X-TOP 주요 모듈과 함수

| Module | Function | Comments |
|--------------------|---------------------------|--------------------------------------------|
| DB Handler | dbConnection() | 데이터베이스 연결을 담당하는 기능 |
| XTM Import | TopicmapLoad() | 외부 XTM 문서의 path정보를 파라미터로 갖고 메모리에 load하는 기능 |
| | read_Topic() | 토픽맵에서 topic을 추출하는 기능 |
| | read_InstanceOf() | topic의 instanceOf 추출하는 기능 |
| | read_SubjectIdentity() | topic의 subjectIdentity 추출하는 기능 |
| | read_BaseName() | topic의 baseName 추출하는 기능 |
| | read_Occurrence() | topic의 occurrence 추출하는 기능 |
| | read_Association() | 토픽맵에서 association과 관련 된 정보만을 추출 하는 기능 |
| | read_Scope() | association에서 scope 추출하는 기능 |
| XTM Export | read_Member() | association에서 member 추출하는 기능 |
| | write_Topic() | DB에서 Topic을 조회하여 메모리에 생성 |
| | write_TInstanceOf() | DB에서 instanceOf조회하여 메모리에 생성 |
| | write_SubjectIdentity() | DB에서 subjectIdentity조회하여 메모리에 생성 |
| | write_baseName() | DB에서 baseName조회하여 메모리에 생성 |
| | write_occurrence() | DB에서 occurrence조회하여 메모리에 생성 |
| | write_Association | DB에서 association을 조회하여 메모리에 생성 |
| | write_Scope() | DB에서 scope를 조회하여 메모리상 생성 |
| write_Member() | DB에서 member를 조회하여 메모리에 생성 | |
| Graphic Viewer | IXmlDocument.save() | 메모리상의 XTM 문서를 로컬영역에 파일로 저장 |
| | MetaTree.Items.Add() | 시작 Topic Node 생성 |
| | designTopicChild() | 하나의 Topic에 대해서 하위 엘리먼트를 화면에 모두 디자인하는 함수 |
| | designTInstanceOf() | topic의 instanceOf 정보를 디자인 |
| | designSubjectIdentity() | topic의 subjectIdentity 정보를 디자인 |
| | designBaseName() | topic의 BaseName 정보를 디자인 |
| designOccurrence() | topic의 occurrence 정보를 디자인 | |
| 기타 영역 | TopicmapidGenerate() | 토픽맵의 ID를 자동 생성 |

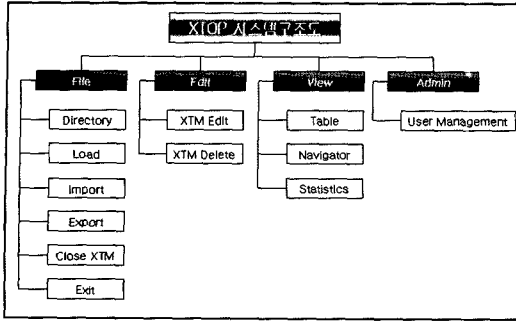


그림 3 X-TOP 시스템 메뉴 구성도

X-TOP 시스템에 등록된 모든 온톨로지 및 토픽맵을 조회 할 수 있으며, 각각의 토픽맵에 대한 사용자 접근 관리와 권한 설정 기능을 제공한다. Statistics는 토픽맵의 각 요소들의 등록 건수에 대한 통계정보를 제공하고, User Management는 X-TOP 시스템의 사용자 계정관리를 담당하는데 Admin, Author, Guest 등 3형태의 사용자 역할을 관리할 수 있다.

3.3.3 X-TOP 시스템 흐름 및 특징

본 논문에서 구현한 X-TOP 시스템의 처리 흐름은 그림 4와 같다. 먼저, 토픽맵 구축 작업 선택을 한 후 목적에 맞도록 토픽맵을 생성하게 되는데, 토픽맵 생성은 문서 편집기에서 새문서 만들기처럼 내용이 없는 문서 만들기과 유사하다. 사용자는 생성되는 토픽맵을 외부 XTM 문서를 임포트 할 것인가를 결정하여야 한다. 외부 토픽맵 문서를 임포트하는 경우에는 레코드 단위로 구문 분석하는 단계인 XTM 파싱 처리 과정을 거치게 된다. 다음에 데이터베이스에 저장(Store)할 것인지를 결정하게 된다. 외부 토픽맵 문서의 경우에는, 토픽

맵을 TopicMaps Open처리 과정을 거쳐 Editor 모드로 자동 전환되어 사용자가 저작할 수 있는 상태가 된다. 데이터베이스에 저장되는 토픽맵은 XTM Generator 모듈을 통해서 XTM 구문 구조에 맞는 파일을 생성하게 된다.

X-TOP은 온톨로지 구조를 관계 데이터베이스의 테이블 스키마로 모델링 하여 관리함으로써 대용량 DBMS (오라클, 사이베이스, 인포믹스, SQL서버 등)와 연동하여 시스템을 구현할 수 있다. 또한, 개발자나 사용자는 관계 데이터베이스의 개발 또는 응용 도구를 그대로 사용할 수 있으며, 상용 DBMS의 성능을 최대한으로 활용하여 온톨로지 문서의 고속 접근, 처리 시간 단축, 대용량 토픽맵 온톨로지 개발등을 효과적으로 수행할 수 있다. X-TOP은 기존의 관계 데이터베이스 기술을 토픽맵에 융합시킨 것으로, 이를 활용하여 통합된 온톨로지 기반 지식 관리 시스템 구축을 목적으로 개발된 시스템이다.

4. X-TOP 구현 결과

본 장에서는 헬스케어 분야의 암 토픽맵을 X-TOP상에서 구현한 사례에 대하여 기술한다. 또한, 기존 토픽맵 시스템에서 발생하는 저장용량, 속도 등 여러 성능 평가 요소를 비교하여, X-TOP의 유효성을 실증한다.

4.1 적용 사례

그림 5는 본 논문에서 구현한 X-TOP 시스템의 주 화면이다. 사용자 계정과 패스워드는 X-TOP 시스템의 User Management 을 통해서 관리되어지며 로그인 계정의 사용자 권한은 다음과 같다. Admin은 X-TOP의 최상위 관리자로서 모든 권한을 갖고 계정 추가 및 삭제, 수정을 할 수 있다. Author X-TOP시스템상에 등

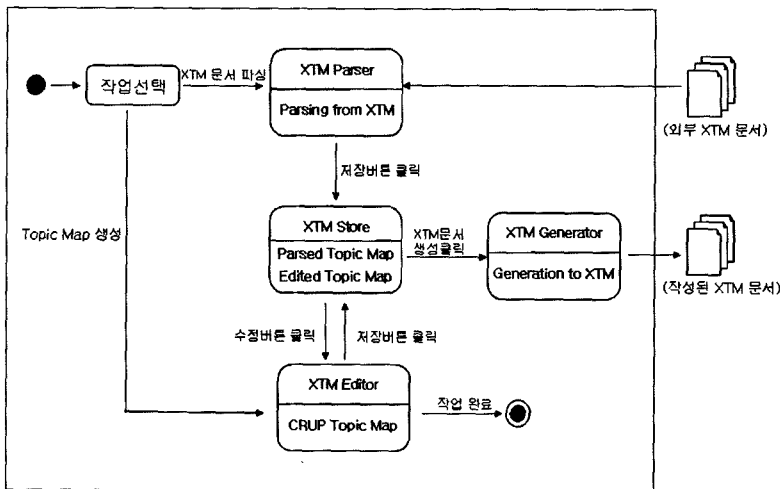


그림 4 X-TOP의 상태 다이어그램

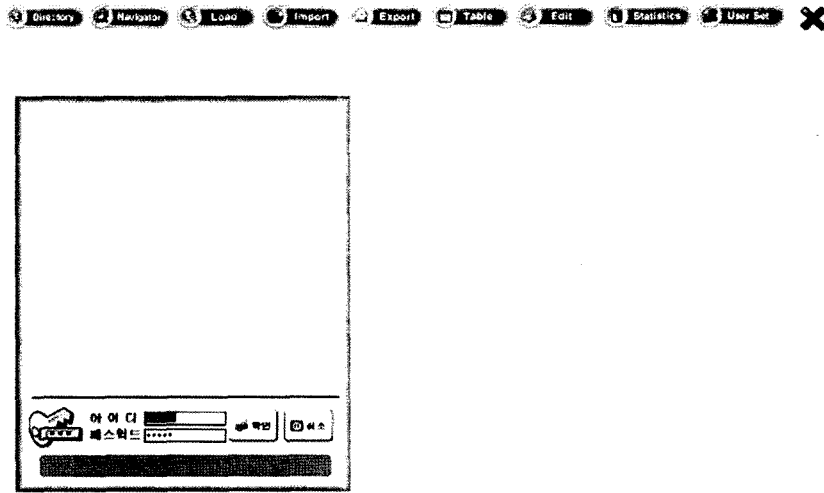


그림 5 X-TOP 시스템의 메인 화면

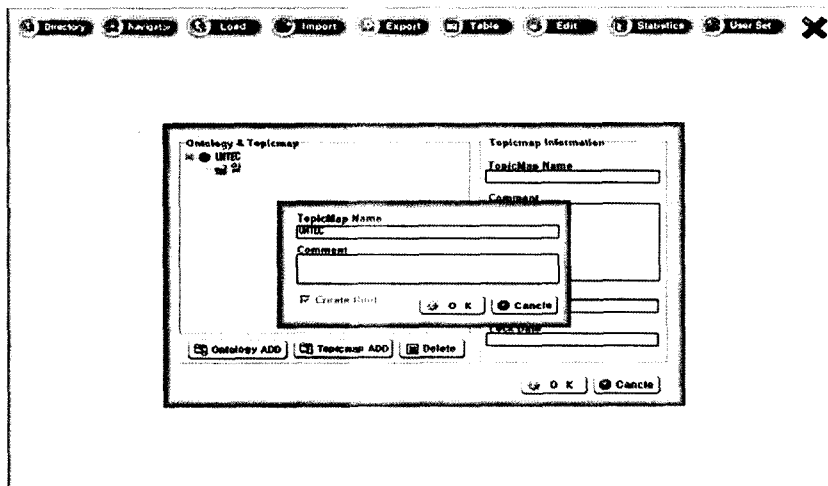


그림 6 X-TOP 시스템의 메인 화면

록된 토픽맵을 저장할 수 있는 권한을 갖는다. Guest는 X-TOP시스템에서 생성 및 관리되어지는 토픽맵들을 조회하는 최소 권한을 갖는다.

그림 6은 온톨로지 및 토픽맵을 신규 생성, 삭제, 수정할 수 있는 기능과 데이터베이스에 등록된 토픽맵의 적재 과정을 보여 주고 있다. 그림 6에서 암 토픽맵을 선택한다.

그림 7의 Table 구성 화면에서는 선택된 토픽맵 중에서 현재 작업 중인 토픽맵에 대해 Topic, InstanceOf, SubjectIdentity, BaseName, Variant, Parameter, Scope, Occurrence, Association, Member, Member_Refer-

ence의 정보 보여 주고 있다.

그림 8과 그림 9는 Editor화면으로서 토픽 저장 과정을 보여주고 있다. 주요 기능은 토픽을 추가, 삭제, 수정할 수 있으며 각각의 토픽에 대한 BaseName 및 Occurrence를 관리할 수 있는 기능이 제공된다. 그림 8의 좌측 트리에서는 “암”이라는 토픽맵에 등록된 모든 <topic> 및 <association> 정보를 보여 주며, 우측에서는 편집 기능이 제공된다. 토픽맵을 저장하는 사용자의 편의를 위해 마우스 우측 버튼을 활용하여 해당 요소 (Topic, Occurrence, Association등)를 클릭하면 그에 해당하는 팝업 메뉴를 제공하여 빠르고 손쉽게 편집할

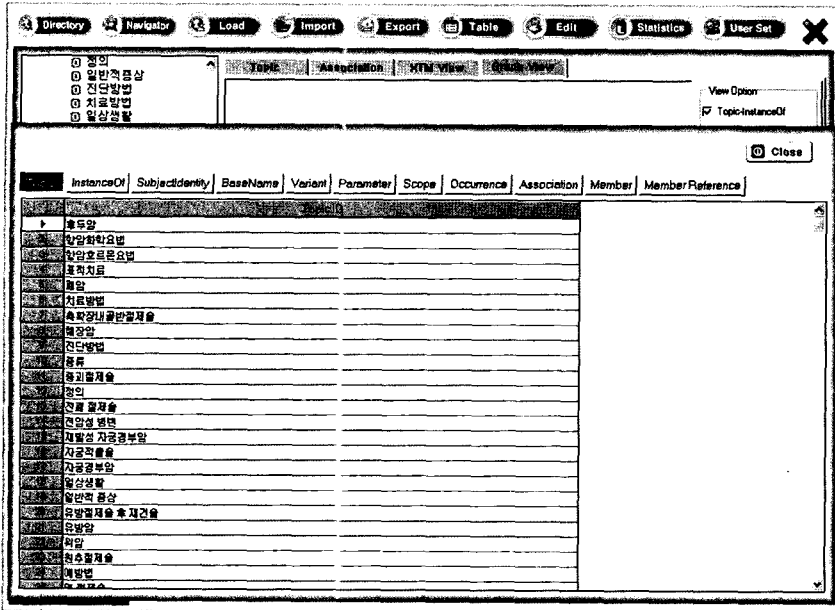


그림 7 Table 구성 화면

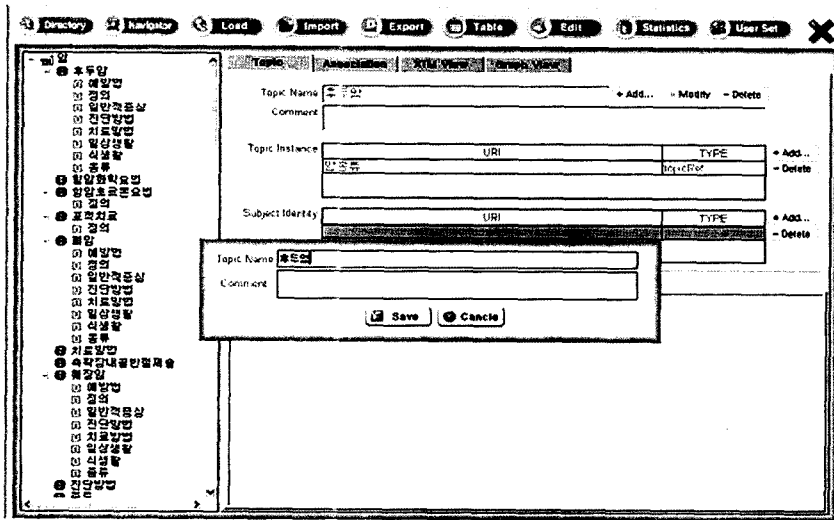


그림 8 Editor의 토픽 저작 화면

수 있다.

그림 10은 Editor 화면 중에 <association>을 추가, 삭제, 수정하는 과정을 보인 것이다. 좌측 트리에서 <association> 또는 <member> 엘리먼트를 클릭하면 그에 해당하는 정보가 우측의 각 항목에 값을 표시 된다. 또한, 좌측 트리에서 마우스 우측 버튼을 클릭하면 추가, 삭제, 수정할 수도 있다. X-TOP에서는 Association Name 또는 Member의 신규 생성 시, 정의할 데이터가 없으면 X-TOP 시스템 상에서 자동 생성되며,

추후 수정할 수 있도록 되어 있다.

그림 11은 Editor의 세 번째 탭으로 데이터베이스에 저장된 토픽맵을 읽어 들어 XTM 구문 구조 형식으로 변환하여 별도의 XTM 파일을 생성한다. 이것은 Export 작업 없이 시스템 상에서 실시간으로 XTM 파일을 생성하는 것이다.

그림 12는 Topic에 대한 연관 정보인 instanceOf, SubjectIdentity, BaseName, Occurrence의 정보를 시각화 하여 검색하는 Graphic Viewer 화면이다. 좌측

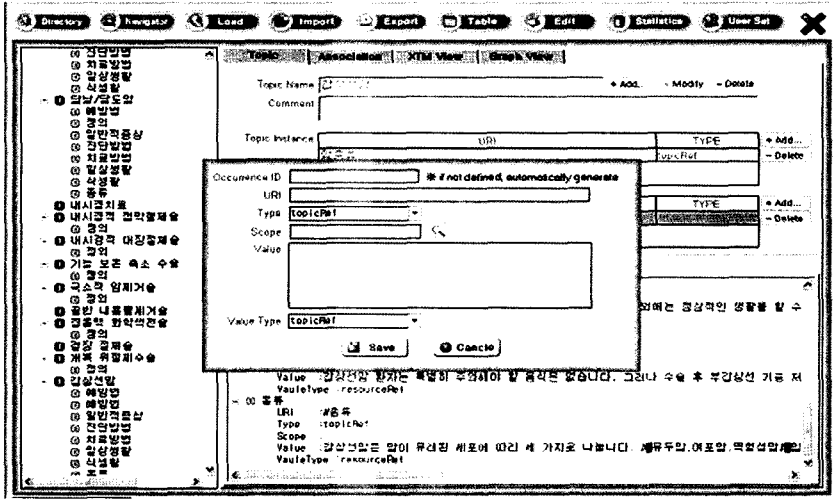


그림 9 Occurrence의 ID/URI/Type/Value 등을 등록하는 화면

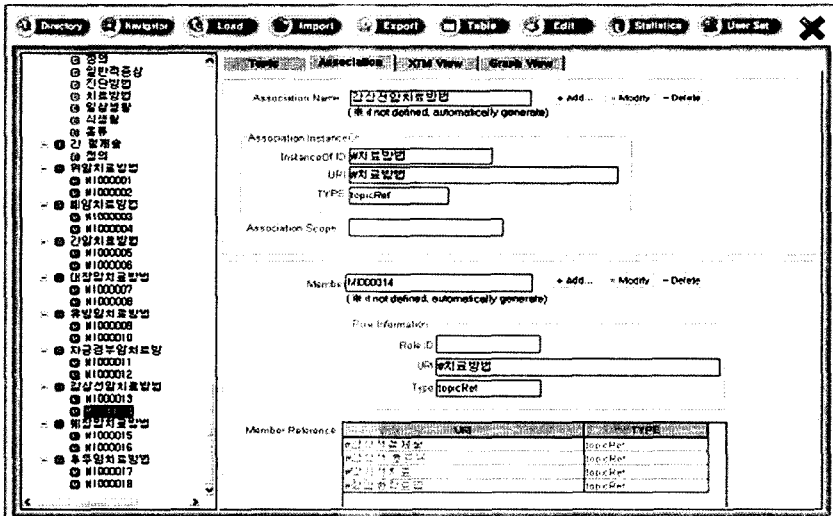


그림 10 Editor의 Association 저장 화면

트리에서 선택된 토픽을 중심으로 하여 연관된 토픽간의 관계가 우측 화면상에 표시된다. 또한 토픽의 관계 정보가 너무 많을 경우 화면이 너무 복잡해 질 수 있는 문제로 인하여 각 노드들에 대해서 확장/ 축소 기능, 각 노드 간의 간격을 설정 할 수 있는 기능, 이미지와 텍스트를 각각 또는 같이 보여줄 수 있는 옵션 기능을 두어 사용자가 효과적으로 검색할 수 있도록 하였다.

4.2 기능 비교

본 절에서는 X-TOP과 기존 시스템간의 성능 평가 결과에 대하여 서술한다. 실험 자료는 '오페라 XTM 파일'과 '암 XTM' 파일을 사용하였으나, 기존 시스템에서 '오페라 XTM 파일' (크기 2.5MB로 파일 내의 토픽 요

소 약10만개)을 적재하는 동안에 시스템이 정지하는 현상이 발생하여, 대용량 토픽맵 저장이 불가능함을 발견하였다. 하지만 X-TOP은 XTM 파일을 RDB 구조에 맞게 관리함으로써 접근 시간이 빨라졌고 컴퓨터가 다 운되는 현상도 없었다. 본 논문에서는 성능 평가를 위하여, W대학병원에서 헬스케어 정보를 토픽맵을 이용하여 서비스 하기위해 구축한 '암 XTM' 파일을 사용하였다. 암 XTM의 내용과 구축 과정은 전 장에서 상세히 기술하였으며, 표 4는 X-TOP과 기존 시스템의 기능을 비교한 것이다[14].

기존 시스템에 비해 X-TOP은 대용량 토픽맵 파일을 RDB 구조로 전환하여 저장 관리하므로 대규모의 토픽

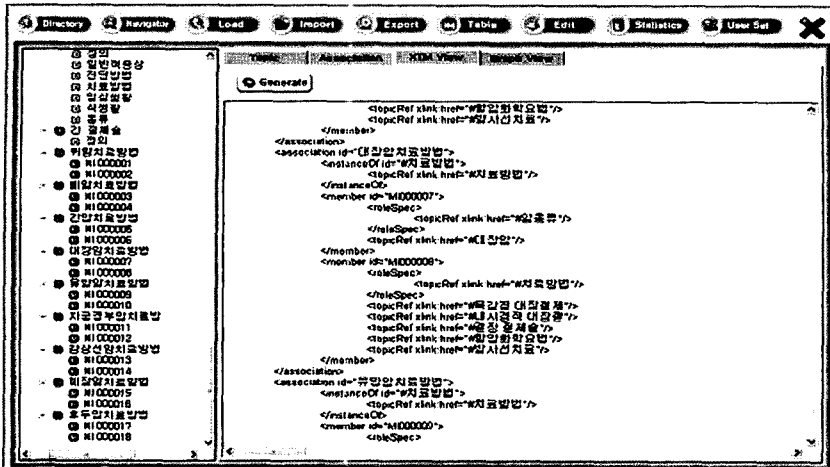


그림 11 Navigator 구현 화면

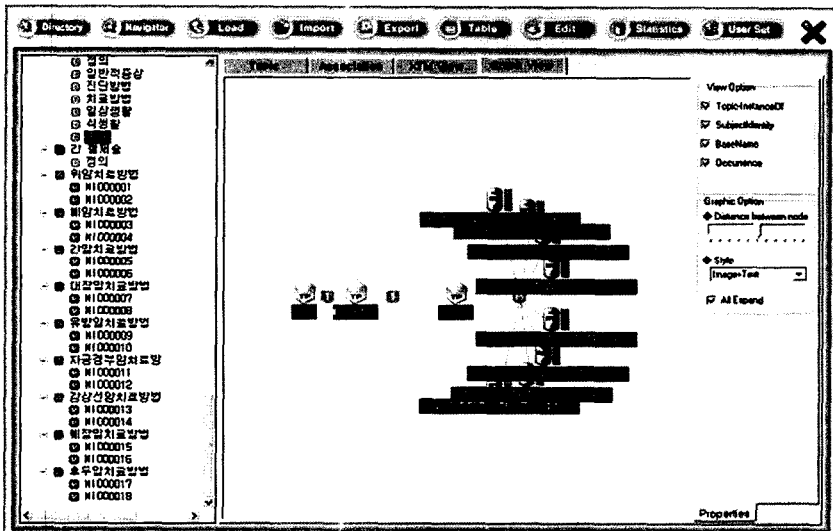


그림 12 Graphic Viewer 구현 화면

맵을 처리할 수 있고, 토픽맵 구축을 빠르고 쉽게 할 수 있는 개발 환경을 지원한다. 또한 멀티유저 기능을 지원하고 있어 대용량 토픽맵 구축 시 협업이 가능하고, 비전문 온톨로지 개발자들이 쉽게 접근 할 수 있으며, 기존의 사용하던 관계 데이터베이스 개발 도구들도 적용 가능하여 레거시 시스템 상에서 온톨로지 구축을 쉽게 할 수 있다.

5. 결론 및 향후 연구 과제

토픽맵은 초기에 서적의 인덱스, 용어, 시소러스, 목차 등의 기술을 위한 도구로 개발되었지만, 온톨로지 개념이 대두되면서 그 응용 영역이 크게 확대되었다. 토픽맵은 토픽중심의 도메인의미 정보 모델링을 지원하므로

다른 패러다임에서 출발한 온톨로지 언어라 할 수 있다. 토픽맵은 도메인의 지식정보를 온톨로지 형식으로 모델링하여 지식정보의 공유와 상호 교환을 실현해 주는 국제 표준 언어이다. 토픽맵은 RDF/S, OWL 등과 더불어 시맨틱웹을 실현하는 방법을 제공하며, 관계 데이터베이스의 정보 표현에도 활용된다. 데이터베이스는 정보와 정보객체사이의 관계만을 표현하지만, 토픽맵은 정보객체가 존재하는 다양한 위치를 연결할 수도 있다. 즉 개념간의 관계뿐만 아니라 개념과 정보자원을 연결하여 지식베이스를 구축하여 준다. 이것은 토픽맵이 기존 또는 앞으로 연구 개발되어질 지식 표현 방식까지 적극적으로 수용할 수 있도록 설계되어 있기 때문이며, 다양한 분야의 지식 표현에 적용할 수 있는 특성이 있다.

표 4 X-TOP과 기존 시스템과의 기능 비교

| 기능 \ 시스템 | X-TOP | Ontopia K.Suite | OntoEdit | KAON |
|---------------------------------|-------------------------------------|-----------------|------------------|--------------------------|
| 온톨로지 유형 | 토픽맵 | 토픽맵 | RDF | RDF |
| 저장장치 | In-memory, RDB, ODB, XML-DB, XTM 파일 | In-memory, RDB | 자체포맷의 파일 | In-memory, RDB, Textfile |
| 검색 방식 | 키워드 검색, SQL, 디렉토리, 그래프 | 키워드 검색, 디렉토리 | 키워드검색, 디렉토리, 그래프 | 검색기능 미약 전체 브라우징 |
| 모델링 | 릴레이션 (record, attribute) | XTM | Triple | Triple |
| 제약조건검사 | 지원 | 지원 | 지원 | 지원 |
| 추론 | 네비게이션 | 네비게이션 | 지원 | 지원 |
| 온톨로지 편집 | 지원 | 지원 | 지원 | 지원 |
| 외부텍스트포맷 | XTM | XTM, LTM | RDF, DAML | 없음 |
| Access 시간 (Opera XTM, 2.5MB 기준) | XTM파일:1hr이상 RDB파일:10sec | 관련없음 | 관련없음 | 관련없음 |

이러한 토픽맵 온톨로지의 실제응용을 위해서, 토픽맵의 구축 및 수정이 가능하고 검색 및 네비게이션 할 수 있는 토픽맵 관리 시스템이 요구되고 있다. 그러나, 기존에 개발된 도구들은 편집기 수준에 머물고 있으며 실무에서 필요로 하는 대용량 토픽맵 구축에는 많은 어려움이 있다. 본 논문은 XTM 1.0 규격에 기반한 대규모 토픽맵을 관계 데이터베이스 모델로 변환하여 관리할 수 있는 방법을 개발하였다. --- RDB 구조로 변환 후 처리 할 수 있는 데이터베이스 구조를 모델링하여 기존에 사용하고 있는 SQL를, 어플리케이션 개발 툴 등을 사용할 수 있게 하고, XTM 문서를 데이터베이스로, 데이터베이스에서 XTM 문서로 호환이 가능한 토픽맵 구축 툴을 개발하였다.

본 논문을 통해 얻을 수 있는 기술적 효과는 다음과 같다.

첫째, 대용량의 토픽맵 파일을 RDB 구조로 변환하여 저장하므로 대용량의 토픽맵을 처리할 수 있으며 또한 토픽맵을 빠르고 쉽게 구축할 수 있는 개발 환경이 구현되었다. 둘째, 독자 활용 방식이 아닌 네트워크 기반 시스템 구현으로 토픽맵 개발시 공간적 제약을 극복할 수 있다. 셋째, 멀티유저 기능을 지원하고 있어 대용량 토픽맵을 구축 시에 협업이 가능 하게 된다. 넷째, 안정적인 DBMS를 활용함으로써 토픽맵 기반 응용 시스템의 안정적인 운영 및 백업 관리가 가능하다. 다섯째, 토픽맵 파일을 RDB로 구현하므로 기존 프로그램 개발자들이 쉽게 접근 할 수 있으며 다른 개발 도구들도 활용 가능하고, 마지막으로 향후 사용자 인터페이스 변경과 다양한 DBMS를 지원할 수 있는 3-계층 구조를 갖고 있어 실무 적용이 가능하다.

향후, X-TOP 시스템을 타 시스템과의 연계를 위한 SOA기반의 웹서비스를 추가 개발하여 웹 포털, 통합

DB 서비스 등의 시스템과도 자동으로 상호 호환할 수 있도록 하고, 다양한 분야에서 응용할 수 있도록 범용 인터페이스 개발이 요청된다.

참 고 문 헌

- [1] 안기진(Kijin Ahn), 이재호(Jaeho Lee) 저, "Graphical TopicMaps Editor(GTM Editor)", 한국정보과학회 2006 가을 학술발표논문집, 제33권 제2호(B), 2006. 10.
- [2] <http://www.w3.org/TR/rdf-schema/>
- [3] <http://www.owl.org/index.jsp>
- [4] 김정민(Jungmin Kim), 박철만(Chulman Pak), 정준원(Junwon Jung), 이한준(Hanjun Lee), 민경섭(Ky-oungsub Min), 김형주(Hyoung-Joo Kim) 저. "K - Box : 토픽맵 기반의 온톨로지 관리 시스템", 정보과학회논문지 : 컴퓨팅의실체, 제10권 제1호, 2004. 2.
- [5] Resource Description Framework, <http://www.w3.org/RDF/>
- [6] 박영조(Youngjo Park), 박호병(Hobyung Park), 조용윤(Yongyoon Cho), 유재우(Chaewoo Yoo) 저, "XTM을 위한 다층적 시각화 방법", 한국정보과학회 2004년도 가을 학술발표논문집, 제31권 제2호(II), 2004. 10.
- [7] Moore, G. ; Ahmed, K. ; Brodie, A., "Topic Map Objects," TMRA, pp. 166-174, 2007.
- [8] 김정민(Jung-Min Kim), 신효필(Hyo-phil Shin), 김형주(Hyoung-Joo Kim), "부산 토픽맵의 다중 전략 매핑 기법", 한국정보과학회, 정보과학회논문지 : 소프트웨어 및 응용, 제33권 제1호, pp. 114-130, 2006. 1.
- [9] 오삼균(Sam-Gyun Oh), "A Comparison of Ontology Languages: Focusing on W3C OWL and ISO Topic Maps."
- [10] 박수민(Soomin Park), 김훈민(Hoonmin Kim), 양정진(Jungjin Yang), "온톨로지 저작도구를 이용한 OWL과 토픽맵의 비교", 한국정보과학회 2006 한국컴퓨터종합학술대회 논문집(B), 2006. 6.
- [11] TopicMaps.org, XML TopicMaps Specification, <http://www.topicmaps.org/xtm/1.0/>, 2001.
- [12] "TM-Design," <http://www.topicmap-design.com/en/>

topicmap-designer.htm.

- [13] Product White Paper, "The Ontopia Knowledge Suite," Ontopia, Inc.
- [14] Aimilia Magkanaraki, Grigoris Karvounarakis, Ta Tuan Anh, Vassilis Christophides, Dimitris Plexousakis, "Ontology Storage and Querying," Technical Report No 308, ICS-FORTH, 2002.
- [15] Michel Biezunski, Martin Bryan and Steve Newcomb, ISO/IEC 13250 TopicMaps
- [16] Herzum, P. Web Services and Service-Oriented Architectures, Executive Report, Vol.4, No.10. Cutter Distributed Enterprise Architecture Advisory Service, 2002.
- [17] Dicheva, B. ; Dicheva, D., "Visual Browsing and Editing of Topic Map-Based Learning Repositories," Lecture notes in computer science, v.4438, pp. 44-55, 2007.
- [18] Stevenson, A. ; Styron, E., "Topic Map Presentation Framework: An Approach to Delivering Newspaper Content Over the Web," Newspaper digitisation; Newspapers of the world online: U.S. and international perspectives, pp. 147-158, 2006.
- [19] Candan, K.S. ; Li, W.S., "Reasoning for Web document associations and its applications in site map construction," North-Holland, Elsevier Science Publishers.
- [20] Dong, Y. ; Li, M., "HyO-XTM: a set of hypergraph operations on XML Topic Map toward knowledge management," North-Holland.
- [21] http://www.idealliance.org/papers/dx_xmlc03/papers/02-03-02/02-03-02.html.
- [22] topicmap.com. "Tools," <http://www.topicmap.com/topicmap/tools.html>



한 성 국

1979년 인하대학교 정보전자공학(공학박사). 1984년~현재 원광대학교 컴퓨터공학부 교수. 1989년 University of Pennsylvania 방문교수. 2003년 DERI 방문교수. 2004년~현재 대한전자공학회 컴퓨터소사이어티 감사. 2005년~현재 한국정보과학회 부회장. 관심분야는 시맨틱 웹서비스, 온톨로지 공학, 웹서비스, 의료정보, e-Learning 등



박 여 삼

1989년 원광대학교 전자계산기공학과(공학사). 1997년 원광대학교 경영학과(경영학석사). 2005년~현재 전북대학교 컴퓨터학과 박사과정. 1991년~2001년 동양기전(주) 입사~전산팀장. 2001년~현재 (주)유엠텍 대표이사. 관심분야는 소프트웨어공학, 시맨틱 웹서비스, 온톨로지, 기술경영(MOT), 정보시스템관리, 의료정보 등



장 옥 배

1973년 고려대학교 졸업(학사, 석사). 1988년 산타바바라대 대학원(Ph. D.). 1974년~1980년 조지아 주립대. 오하이오 주립대 박사과정 수료. 1980년~현재 전북대학교 전자정보공학부 교수. 관심분야는 소프트웨어 공학, 전산교육, 수치해석, 인공지능 등