

XTM 기반의 지식맵

(Knowledge Map of XTM Base)

정 호 영*, 김 정 민*, 정 준 원*, 김 형 주**
(Ho-young Jeong*, Jung-min Kim*, Jun-Won Jung*, Hyoung-Joo Kim**)

요 약 지식관리시스템은 기업이나 조직의 지식을 저장하고 조직화하며 구성원들간에 지식을 공유할 수 있도록 지원하는 시스템으로서 여러 하위 시스템들로 구성되는데 그 중에서도 지식의 효율적인 검색과 네비게이션을 위한 지식맵이 중요한 부분을 차지한다. 그러나 기존 지식관리시스템들의 지식맵 구조를 보면 과거의 문서관리시스템에서 사용하던 디렉토리 구조의 색인을 사용하고 있기 때문에 지식 정보의 연관 검색이 불가능하며 지식관리시스템들 사이의 지식맵의 공유와 교환뿐만 아니라 여러 지식맵의 통합 및 분할 또한 어렵다는 문제를 가진다. 이러한 문제를 해결하기 위해 본 논문에서는 지식 자원들을 상호 연관성에 따라 연결하고 조직하여 지식 구조를 기술할 수 있도록 토픽맵 모델을 기반으로 하는 지식맵을 구축하기 위한 시스템을 제안하고 있다.

주제어 : 지식맵, 토픽맵, 지식관리

Abstract Knowledge Management System manages a process that helps organizations to become more competitive by enabling them to operate more effectively.

KMS consists of several sub-system knowledge portal, collaboration server, knowledge map, extractor, etc. Knowledge map is used to organize the knowledge index to search and navigate the knowledge base. But traditional KMS provides a simple directory-based knowledge map. This map structure is devel-

oped from the structure of the document index of Electronic Document Management System. So, this directory-based knowledge map can not provide the associative relationship between topics (or knowledge items). In this paper, we proposes a topic map based knowledge map with the semantic links between the topics in that map. Using this map we represents the hierarchical and associative relationships between the topics.

Keywords : Knowledge Map, Topic Map, Knowledge Management

1. 서 론

1.1 연구동기

지식관리는 기업이나 조직 내의 다양한 형태로 존재하는 지식을 중요한 자산으로 인식하는 데서 그 필요성이 대두되었으며 많은 기업들이 지식경영이라는 이름으로 기업의 지식을 기업경영에 적용하고 있다. 그러나 지식을 효율적으로 관리하고 활용하기 위해서는 무엇보다 지식을 관리하고 공유하도록 지원하는 시스템이 필수적인데 이를 위해 개발된 시스템이 지식관리시스템이다.

지식관리시스템은 지식을 저장할 수 있는 지식 베이스와 등록된 지식을 표현할 수 있는 지식표현, 지식을 검색하고 네비게이션할 수 있는 지식맵, 그리고 지식을 배포하고 활용하는 지식공유 및 지식

* 준회원 : 서울대학교 컴퓨터공학과 박사과정

** 정회원 : 서울대학교 컴퓨터공학과 교수

+ 논문접수 : 2003년 5월 27일

+ 심사완료 : 2003년 7월 18일

활동 평가 등 기업 및 조직의 활동에 지식을 활용할 수 있도록 지원하는 여러 하위 시스템들로 구성되어있다. 그러나 지식관리시스템은 처음부터 새롭게 개발된 시스템이라기보다는 기존의 그룹웨어, 워크플로우, 문서관리시스템, 웹포털 및 검색엔진 등의 개별 시스템들을 지식관리 프로세스에 적합하도록 통합해 놓은 시스템이라고 볼 수 있다. 이 하위 시스템들 중에 지식맵은 지식 리파지토리에 저장된 수많은 지식들을 유사성에 따라 분류하고 상호 연관관계를 맺음으로써 키워드 검색뿐만 아니라 분류 검색과 연관검색을 가능하도록 한다.

지식관리시스템이 구축되어 운용될 경우 적게는 수십만 건에서 많게는 수백만 건의 지식자료가 지식베이스에 저장되게 된다. 이 지식자료는 시스템을 활용할수록 더욱더 많이 증가하게 된다. 이 경우 지식맵이 없다면 지식 사용자들은 자신이 원하는 지식을 쉽게 획득하고 검색할 수 없게 되고 결국 시스템을 활용하지 않게 된다. 그러나 현재 지식관리시스템의 지식맵은 다음의 몇 가지 문제점을 가지고 있다. 이는 기존의 지식관리시스템이 위에서 기술한 것처럼 단위 시스템으로 개발된 여러 시스템들을 통합하여 구성되었기 때문에 지식맵 또한 문서관리시스템의 문서 색인이나 디렉토리 서비스와 같은 트리구조의 모델로 제공되기 때문이다.

- 지식항목간에 계층구조의 관계만을 제공한다

현재의 지식맵의 지식 데이터 분류구조는 트리구조를 기반으로 한다. 예를 들어 제품 설계와 관련된 지식의 경우 가전제품 > 세탁기 > 드럼세탁기 > 설계도면 > PART A 등으로 구성된다. 이 경우 특정 지식자료가 어떤 분류에 속하는지 그리고 같은 분류 내에 어떤 지식자료들이 있는지는 쉽게 알 수 있지만 서로 다른 분류에 속하는 또는 같은 분류에 속하는 것이라도 지식자료들간의 연관성에 대해서는 파악할 수 없는 구조이다.

- 지식맵의 지식내용이 주로 업무절차 중심적이다

기업 또는 조직의 지식은 크게 업무 프로세스를 어떻게 처리해 나가는지 절차에 대한 지식과 업무

별로 관리하고 생성해야 하는 자료에 대한 지식으로 나누어 볼 수 있다.

현재 구축되는 지식맵은 주로 업무 절차에 관한 지식맵이다. 즉 신규사원이 입사할 경우 그 사원이 해야 할 업무 절차가 분류별, 태스크별로 정리되어 있으므로 선임자의 업무 인수가 없더라도 지식맵을 통해 담당 업무를 처리할 수 있도록 되어 있다. 그러나 문서, 이미지, 동영상 등 데이터로서 존재하는 수많은 지식 데이터 또한 분류체계와 연관성에 따라 관리해야 할 필요가 있다.

- 지식관리시스템에 의존적이다

지식관리시스템마다 독자적인 맵 구조와 인터페이스를 제공한다. 이는 기업의 지식이 특정 지식관리시스템에 종속됨을 의미한다. 즉 지식관리시스템이 구축된 이후 뛰어난 성능의 다른 지식관리시스템으로 업그레이드를 하고자 할 경우에 지식맵을 새로 구축해야 한다는 것을 의미한다. 그러므로 지식맵을 구축하는 과정은 많은 시간과 노력이 소요되기 때문에 타 시스템으로 쉽게 변경할 수 없게 된다.

- 지식맵의 표준화된 데이터모델이 없다

지식맵의 내부 데이터 표현구조가 표준화되어 있지 않다. 지식관리시스템 마다 독자적인 데이터 구조로 맵을 표현하고 있으며 이는 지식맵을 여러 시스템과 공유하거나 통합 또는 분할 할 수 없다는 것을 의미한다. 즉 한 기업에서 해외에 여러 지사를 설치해서 각 지사에도 지식관리시스템을 구축할 경우 현재는 동일한 시스템을 사용하지 않는 이상 새로 구축해야 한다. 이는 본사에 구축된 지식맵을 활용할 수 없다는 것이다.

- 표준화된 질의언어가 없다

구축된 지식맵은 보통 웹페이지를 통해서 접근하게 되는데 이때 시스템에서 제공하는 인터페이스를 통해서만 접근이 가능하다. 이는 사용자가 지식맵 정보를 다양하게 검색할 수 있는 질의언어가 없기 때문인데 이로 인해 사용자의 요구는 제한되어 질 수밖에 없다.

지금까지 기술한 문제점들은 궁극적으로 지식맵의 표준화된 데이터 모델이 없다는 데에서 발생하는 문제라고 볼 수 있다.

본 논문에서는 이러한 문제를 해결하기 위해 XML에 기반을 둔 토픽맵(Topic map)이라는 데이터 모델을 소개하고 이 모델을 이용하여 지식맵을 구축할 수 있는 시스템으로 연구소의 세미나자료를 지식맵으로 표현, 관리하는 것에 적용시켰으며 이는 연구소 구성원들에게 소중한 자료(지식)인 세미나자료에 대한 지식기반의 콘텐츠 관리를 가능케 하며 기존의 검색 방법인 키워드 검색뿐만 아니라 지식리소스에 대한 분류검색과 연관검색과 같은 유용한 질의를 제공한다.

토픽맵은 정보 자원들을 상호 연관성에 따라 연결하고 조직하여 지식 구조를 기술할 수 있도록 제정된 ISO (International Organization for Standardization) 표준이다.

토픽맵은 대용량의 비구조화되고 비조직화된 정보를 효율적으로 검색하고 네비게이션하기 위한 해결책으로 제안되었다. 토픽맵은 대용량의 정보를 분류하고 구조화하며 의미론적인 연관관계를 설정할 수 있는 모델을 제시하고 있으며 원하는 지식을 쉽고 정확하게 찾을 수 있는 맵을 제시한다.

토픽맵은 지금까지의 전통적인 색인 방법뿐만 아니라 전자도서관, 지식관리시스템, 콘텐츠 관리 시스템과 함께 거의 모든 종류의 정보관리시스템에 정보 분류 및 검색을 위한 데이터 모델로 사용될 수 있다.

1.2 논문의 구성

2장에서는 정보 자원에 의미정보를 부여하여 구조화하는 여러 모델들에 대해 기술하고 있으며 3장에서는 토픽맵 모델에 대한 소개를 담고있다. 4장에서는 XTM(XML Topic Map)을 시스템에 적용시키는 방안에 대해서 소개하고 있다.

2. 관련 연구

색인(index), 용어해설(glossary), 시소러스 (thesaurus)

는 책이나 다른 정보자원에 담겨있는 지식 구조를 매핑하는 방법들이다. 인공지능 연구분야에서는 예전부터 인간과 기계사이에 의사전달을 위해 지식을 어떻게 표현할 것인가에 대해 연구를 해 왔다.

지식 표현에 널리 사용된 방법 중의 하나는 개념(concepts)과 개념들간의 관계 (conceptual relationship)로 구성된 개념도 (conceptual graphs)이다. 이 개념도는 실제 AI 시스템에서 'semantic network', 'associative network', 'partitioned network' 등 다양한 이름으로 구현되어 왔다.[4] 이 semantic network의 기본 모델은 전통적인 분류기법인 색인, 용어해설, 시소러스에서 나타나는 토픽, 연관관계와 매우 유사하기 때문에 이 두 기법을 접목함으로써 정보관리와 지식관리 분야에 많은 잇점을 제공할 수 있고 이것이 토픽맵이라는 새로운 표준을 만들게 할 수 있었던 토대가 되었다 할 수 있을 것이다.

토픽맵은 토픽/연관관계 모델에 토픽/어커런스(occurrence)를 추가함으로써 지식 표현과 정보관리 분야 사이에 연결고리 역할을 담당할 수 있게 되었다.

지식은 정보와 엄연히 다르다. 어떤 것을 알고 있다는 것과 그것에 대한 정보를 가지고 있다는 것은 다른 것이다. 그런 측면에서 볼 때 지식관리는 '생성', '형식화', '전달'의 3가지 지식 활동으로 요약할 수 있다. 토픽맵은 이 중 형식화를 위한 표준이며 생성과 전달을 위한 도구를 개발하는 데에도 필수 요소이다.

W3C(World Wide Web Consortium)에서는 토픽맵이 등장하기 이전부터 웹에 의미정보를 표현할 수 있는 방법을 연구해왔다. 즉 웹 페이지들에 그 페이지 자체를 설명하는 메타정보 -페이지에 담겨있는 내용의 주제, 만든 이, 만든 일자, 관련페이지 등- 를 기술함으로써 웹 페이지의 하이퍼링크뿐만 아니라 시맨틱 링크에 의해서 보다 정확한 검색을 하고자 하였다. 그 결과로 RDF (Resource Description Framework) 라는 모델이 제시되었으며 웹 자원의 시맨틱 맵을 구축하기 위한 도구로 사용되고 있다. 또한 토픽맵과의 모델 통합 및 상호 교환에 대해서도 연구가 수행되고 있다. 이는

양쪽 모델로 표현된 맵이 각각을 지원하는 시스템들 사이에 공유가 가능하다는 것을 의미한다.[15][16]

3. 토픽맵 모델

토픽맵 모델은 기본요소로서 토픽(Topic), 연관관계 (Association), 어커런스 (Occurrence) 로 나누어 볼 수 있다.[5]

3.1 토픽(Topic)

토픽은 일반적으로 어떠한 것이든 될 수 있다. 즉 사람, 사물, 개념, 의미 등 실제 존재하는 것, 또는 특정 속성이나 어떤 의미 등이 될 수 있다. 보통 명사형으로 표현된다. 즉 특정 문서의 토픽은 그 문서의 작성자가 나타내고자 하는 주제를 표현할 수 있는 단어들로 구성되어 있어야 한다.

XML에 관한 책에서 보면 'element', 'attribute', 'structured document', 'XML', 'DTD(Document Type Definition)' 등이 토픽이 될 수 있다. PDM(Product Data Management) 관련 지식관리 도메인에서의 토픽들을 분류하면 '설계 제품 모델명', '도면', '부품코드', '부품도', '설계자' 등이 토픽 타입이 될 수 있고 토픽은 각 토픽 타입의 실제 인스턴스가 된다.

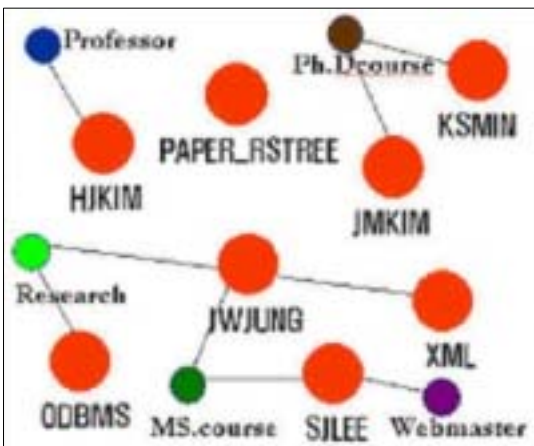
즉, '모델'이라는 토픽 타입에는냉장고 PDM 관련

지식관리 도메인에서의 토픽들을 분류하면 '설계 제품 모델명', '도면', '부품코드', '부품도', '설계자' 등이 토픽 타입이 될 수 있고 토픽은 각 토픽 타입의 실제 인스턴스가 된다. 즉, '모델'이라는 토픽 타입에는냉장고 모델번호 'DW2003'이라는 특정 객체를 가리키는 토픽과 다른 모델을 가리키는 토픽들이 있으며 '설계자' 라는 토픽타입에는 실제 설계를 담당하는 사람들 이름이 토픽이 된다. 그림 1은 서울대학교 객체지향 연구실 홈페이지에서 추출한 토픽들을 보여주는 그림이다.

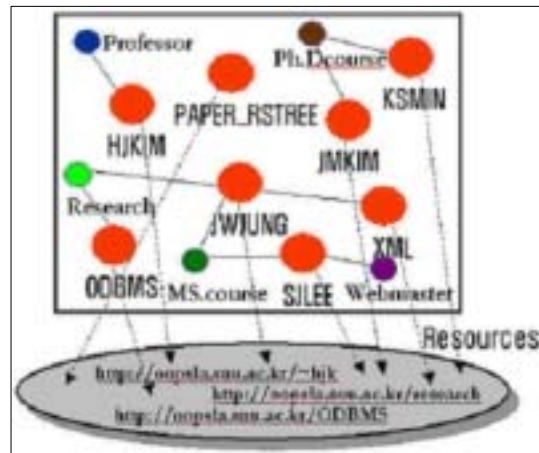
하나의 토픽맵에는 보통 수십만 개에서 수백만 개의 토픽들이 존재한다. 수많은 토픽들을 서로 연결짓기 위해서는 먼저 비슷한 유형에 속하는 것끼리 분류를 해야 하는데 이때 토픽 타입이라는 것을 사용한다.

토픽맵에 있는 토픽들은 토픽 타입이 없거나 있다면 1개 이상 여러 타입에 속할 수 있다. 토픽 타입과 토픽의 관계는 객체 지향 모델에서의 클래스와 객체(인스턴스)의 관계에 비유될 수 있다.

토픽맵 표준안 스펙에 따르면 토픽 타입은 그 자체가 또 하나의 토픽이다. 그러므로 토픽 타입들 간에도 상하관계와 연관관계가 존재할 수 있다. 예를 들어 '연구실멤버' 토픽 타입은 '교수', '박사과정', '석사과정' 등으로 나누어 질 수 있으며 연구실멤버와 이들 세부 멤버타입 사이에는 상하관계 또는 상속관계가 존재하게 된다. 이러한 관계는



(그림 1) 토픽타입과 토픽



(그림 2) 지식 항목을 가리키는 어커런스

객체 지향 모델에서 슈퍼클래스와 서브클래스간의 'isa' 관계에 해당하는 것이다.

토픽들은 '이름', '어커런스', 그리고 다른 토픽과의 연관관계에서의 '역할' 이라는 3가지 종류의 성격을 가진다. 일반적으로 모든 토픽은 최소 하나의 이름을 가진다.

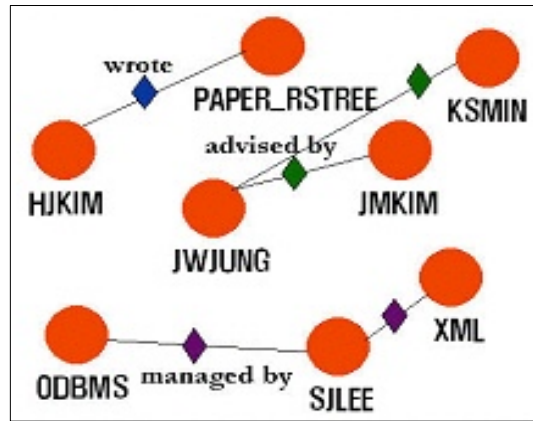
3.2 어커런스(Occurrence)

각 토픽은 자신이 참조하는 하나 이상의 실제적인 지식 항목(내용)과 연결될 수 있다. 예를 들어 '연구실멤버' 라는 토픽은 객체지향 연구실의 멤버들을 나열하고 있는 <http://oopsla.snu.ac.kr/member> 라는 페이지를 가리키는 것으로 자원(resource) 과 연결될 수 있다. 이러한 연결정보를 어커런스라고 한다. 어커런스는 여러 형태를 가질 수 있다. 문서 파일, 이미지 파일, 비디오 파일, 데이터베이스 내 특정 레코드 등으로 나타날 수 있다. 그림 2는 서울대학교 객체지향 연구실 홈페이지에서 각 토픽들이 지식항목을 가르키는 어커런스를 보여주는 그림이다. 이러한 어커런스에 의해 참조되는 자원들은 일반적으로 토픽맵과 별도로 저장된다. 토픽맵 내의 토픽들은 자신의 자원을 가리키기 위해 HyTime 어드레싱이나 Xpointer 와 같은 기법들을 사용하게 된다.

토픽맵과 실제 지식 자원 사이를 분리할 수 있다는 것은 실제 지식 항목의 이동없이 토픽맵만으로 지식 항목을 분류하고 색인 할 수 있는 강력한 기능을 제공하게 된다. 즉, 지식 리파지토리에 저장된 지식이 없더라도 토픽맵을 이용하여 지식맵을 구축할 수 있으며 지식 리파지토리와는 별도로 지식맵 만을 배포, 공유, 교환할 수 있다.

3.3 연관관계(Association)

토픽맵 표준안 스펙에는 토픽들간의 연관관계를 정의 할 수 있는 방법을 제공하고 있다. 토픽 연관관계는 둘 이상의 토픽들 사이에 상하관계가 아닌 의미적인 관계성을 정의하는 관계이다.



(그림 3) 토픽들간의 연관관계

JMKIM is a *member_of* OOPSLA
 SDLEE *wrote* a paper, 'SPY Tec'
 KSKIM *manage* OOPSLA homepages
 HJKIM is a professor
 A paper, 'The RS-tree', was *written_by* HJKIM

JMKIM과 OOPSLA는 토픽들이고 이 두 토픽 사이에는 'member_of' 라는 연관관계가 존재한다. 이처럼 토픽맵 내의 토픽들은 서로 독립적인 객체 이면서 동시에 특정 타입의 연관관계로 연결된 링크를 가진다.

토픽들이 토픽 타입이라는 것으로 분류가 되는 것처럼 토픽들간의 연관관계도 연관관계 타입이라는 것으로 분류된다. 위 예에서 'member_of', 'wrote', 'manage', 'is a', 'written_by' 등이 연관관계 타입이다. 즉 'RS Tree' 라는 논문이 HJKIM에 의해 쓰여졌듯이 다른 논문 도 다른 멤버에 의해 쓰여졌다는 관계를 설정할 수 있는 것이다. 'written_by' 라는 연관관계 타입은 하나이지만 실제 연결되는 토픽들에 따라 그 타입에 속하는 연관관계가 여러 개 나올 수 있는 것이다. 그림 3은 서울대학교 객체지향 연구실 홈페이지에서 각 토픽들간의 연관관계를 보여주는 그림이다.

토픽 타입과 연관관계 타입은 지식 및 정보 표현, 분류, 구조화를 위한 토픽맵에게 힘을 실어주는 기능이다. 직관적이면서 사용자에게 친숙한 인터페이스를 제공할 수 있을 뿐만 아니라 토픽맵을 네

비게이션 할 때 사용자가 관심 있는 분야를 쉽게 축소할 수 있고 검색 결과 또한 정확하게 만들어 낼 수 있다.

연관관계에는 타입외에도 '역할'이라는 중요한 정보가 있다. "JKIM is a *member_of* OOPSLA"라는 예를 보면 JKIM은 사람(person)이라는 역할을 가지고 OOPSLA는 연구실(laboratory)이라는 역할을 가진다. 토픽맵에서는 이 역할들 -□ person, laboratory- 도 하나의 토픽으로 구성되어 있다. 이 말은 연관관계에 어떤 제약조건을 줄 수 있다는 것을 의미한다. 즉, '*member_of*'라는 연관관계에서는 person 역할의 토픽과 laboratory 역할의 토픽이 있어야 하며 나오는 순서도 person이 laboratory보다 먼저 와야 한다는 것이다.

지금까지 살펴 본 토픽, 어커런스, 연관관계의 3가지 항목외에도 토픽맵 표준안 스펙에는 여러가지 지식을 관리하고 구조화하기 위한 여러가지 요소들을 나열하고 있다.

다음 장에서는 XTM을 이용한 연구실의 방대하고 구조화되어 있지 않은 온라인/오프라인 세미나 자료에 대해 지식맵을 구성하여 효율적인 콘텐츠 관리와 지식기반의 연관검색을 가능케 함을 보인다.

4. XTM 기반의 지식맵 구현

기업이나 조직의 지식과 마찬가지로 연구실의 세미나 자료는 연구원들에게는 좋은 지식이 될 수 있다. 주 단위로 수없이 쏟아지는 세미나 자료는 -이전까지의 연구실의 세미나 자료는 대개는 발표자 이외의 연구원들은 하드카피본으로 지나다가 분실하는 경우도 많고 소유하고 있는 경우라도 원하는 자료를 찾으면 그리 쉬운 일은 아니었다. 또한 세미나 자료는 주로 발표자의 로컬 컴퓨터에 저장되어 있는 경우가 많고, 공유와 검색에 있어서 효율적이지 못하다는 불편함이 있으며 이에 지식관리시스템을 필요로 한다. -연구소 구성원들에게 공유 및 검색을 지원하기 위해 주로 웹이나 인트라넷으로 지원되는데 효율적으로 관리하고 검색을 지원하기엔 이전까지의 디렉토리 색인에 의존하는

기존의 지식관리시스템은 1장에서 언급한 바와 같이 여러가지 문제점을 나타낼 수 있으며 특히 검색에 있어서는 세미나 자료의 키워드 검색을 통한 제목에 의존적이어서 분류검색이나 참고자료, 발표자 또는 다른 세미나 자료와의 연관성에 대한 파악이 어렵다.

본 논문에서는 이에 연구소의 세미나 자료에 대해 XTM을 사용하여 부가적인 메타 데이터-토픽, 어커런스, 토픽과 토픽간의 연관관계-를 기술함으로써 키워드 검색뿐 아니라 세미나 자료 지식에 대한 생성/유지/관리의 용이함을 지원하며 토픽들간의 분류/연관관계에 의한 검색을 지원하여 연구원이 필요로 하는 지식에 대한 접근을 쉽게 할 수 있는 방법을 제공한다.

그림 4는 XTM으로 기술된 세미나 자료를 보여주고 있다. 그림을 통해 알 수 있듯이 seminar1이라는 topic id를 가지는 토픽은 토픽타입이 seminar이고 base name으로 "Design and Implementation of a Three-Step Intentional Query Processing Scheme"을 가지며 토픽을 기술하는 어커런스는 "http://oopsla.snu.ac.kr/oopsla10/rsrch/images/3iqps.jpg"를 나타낸다.

그림 5는 세미나자료와 발표연도간의 연관관계를 나타낸다.

```

<!--세미나 자료-->
<topic id="seminar1">
  <instanceOf>
    <topicRef xlink:href="#seminar"/>
  </instanceOf>
  <baseName>
    <baseNameString>Design and
      Implementation of a Three-Step
      Intentional Query Processing
      Scheme</baseNameString>
  </baseName>
  <occurrence>
    <resourceRef
      xlink:href="http://oopsla.snu.ac.kr/oopsla10/rsrch/images/3iqps.jpg"/>
    </occurrence>
  </topic>

```

(그림 4) XTM으로 기술된 세미나 자료

```

<!--발표연도-->
<association id="#written_in1">
  <instanceOf>
    <topicRef xlink:href="#written_in"/>
  </instanceOf>
  <member>
    <roleSpec>
      <topicRef xlink:href="#r_seminar"/>
    </roleSpec>
    <topicRef xlink:href="#seminar1"/>
  </member>
  <member>
    <roleSpec>
      <topicRef xlink:href="#r_year"/>
    </roleSpec>
    <topicRef xlink:href="#year_1986"/>
  </member>
</association>

```

(그림 5) 세미나자료와 발표연도와의 연관관계

```

<!--발표-->
<association id="#written_by1">
  <instanceOf>
    <topicRef xlink:href="#written_by"/>
  </instanceOf>
  <member>
    <roleSpec>
      <topicRef xlink:href="#r_seminar"/>
    </roleSpec>
    <topicRef xlink:href="#seminar1"/>
  </member>
  <member>
    <roleSpec>
      <topicRef
        xlink:href="#r_professor"/>
    </roleSpec>
    <topicRef
      xlink:href="#professor_HJK"/>
  </member>
  <member>
    <roleSpec>
      <topicRef
        xlink:href="#r_researcher"/>
    </roleSpec>
    <topicRef xlink:href="#JHAHN"/>
  </member>
</association>

```

(그림 6) 세미나자료와 발표자와의 연관관계

그림에서 나타내는 연관관계는 세미나자료와 발표연도간의 연관관계를 나타내며 연관관계 타입으

로 "#written_in"을 가진다. 또한 "#written_in" 연관관계 타입에 대해 각 토픽 "#seminar1"과 "#year_1986"은 "#r_seminar"과 "#r_year"라는 역할을 하게 된다.

그림 6은 세미나자료와 발표자와의 연관관계를 나타낸다. 이 연관관계는 연관관계 타입으로 "#written_by" 가지며 세미나자료, 지도교수, 발표자(연구원) "#r_seminar", "#r_professor", "#r_researcher"의 역할을 가지게 된다. 4장에서는 위와 같이 비구조화되고 비조직화된 자원에 XTM을 이용하여 구조화된 메타데이터를 기술하고 상호 토픽들간에 연관관계를 기술함으로써 구조화된 지식맵을 표현하여 지식관리시스템에 대한 생성/유지/관리를 용이케 하며 분류검색과 연관검색과 같은 지식기반의 검색을 지원한다.

그림 7과 그림 8은 서울대학교 객체지향시스템 연구실에서 XTM을 이용하여 구현한 지식맵 엔진과 브라우징 시스템으로서 연구원들의 세미나자료에 대한 효율적인 관리를 해주며 토픽간의 지능화된 검색을 지원한다.

그림 7을 보면 좌측 프레임에서는 트리형태의 지식분류를 통한 탐색을 할 수 있으며, 지식맵 > 세미나자료 > "Design and Implementation of a Three-Step Intentional Query Processing Scheme"를 검색했을 때의 토픽에 대한 지식정보를 보여주고 있다. 우측 상단의 프레임은 토픽을 지칭하는 정보인 어커런스 "<http://oopsla.snu.ac.kr/oopsla10/rsrch/images/3iqps.jpg>"를 보여주는 프레임이고 우측 하단의 프레임은 토픽에 대한 지식정보 토픽 타입, 어커런스, 다른 토픽들간의 연관관계 등 - 를 보여주고 있다. 이 그림을 통해 알 수 있듯이 검색을 통해 얻은 결과는 그 토픽에 대한 정보뿐만 아니라 다른 토픽들간의 연관관계를 잘 나타내 주고 있다. 예를 들어 그림에서 얻는 검색의 결과를 통해 토픽을 지칭하는 정보인 어커런스 뿐 아니라 발표자, 주제, 출처, 연도 등 토픽에 연관된 다른 토픽에 대한 검색과 분류검색이 용이하다. 그림 8은 그림 7에서 얻은 검색의 결과중에 "written_by" 라는 연관관계를 가지는 즉 발표자(연구원)에 대한 검색결과이다. 이는 단지 토픽에 대한 키워드 검색으로 얻은



(그림 7) 토픽에 관한 지식 검색



(그림 8) 토픽의 연관관계에 의한 지식 검색

결과뿐 아니라 그 사람이 발표했던 다른 자료와 같은 연관검색이나 카테고리별 분류검색을 쉽게 할 수 있음을 알 수 있다. 즉 지식에 대한 시맨틱 링크를 제공하고 있다.

5. 결론

토픽맵 애플리케이션은 여러 분야에 활용될 수 있다. 콘텐츠 관리시스템의 경우 토픽맵 애플리케이션

이션을 이용하여 콘텐츠들의 연관관계를 설정하고 구조화하여 관리할 수 있으며 기업 애플리케이션 통합(EAI ? Enterprise Application Integration)에서는 토픽맵을 여러 분리된 데이터 소스를 통합하여 가상적인 뷰를 제공하는데 사용할 수 있다.

본 논문에서는 지식관리시스템에서 제품, 조직의 구성원, 문서 등 각종 지식 데이터들을 분류하고 연관관계를 설정하는 지식맵을 구축하기 위해 토픽맵 모델을 사용하는 것을 제시하고 연구소의 세미나자료에 대한 콘텐츠 관리시스템에 적용하였다. 기존의 지식맵은 표준화된 데이터 모델이 없기 때문에 여러 가지 문제점을 가지고 있다. 이러한 문제를 해결하기 위해서는 표준화된 데이터 모델을 기반으로 지식맵을 구축할 수 있는 시스템이 필요하며 본 논문에서는 이러한 시스템을 설계하는 과정을 설명하였다.

지식맵의 표준 데이터 모델로 사용되는 토픽맵은 ISO 에서 표준으로 제정한 것으로 주로 대용량의 정보 자원을 개념적 차원에서 분류하고 상호 연관관계를 정의하여 색인 구조를 구축하는 데 사용된다. 이 토픽맵을 사용함으로써 지식맵은 지식 데이터간에 시맨틱링크를 가지는 그래프 구조를 가지게 된다.

향후에는 토픽맵 검색과 네비게이션의 최적화를 위해 그래프 색인 기법의 향상에 대한 연구와 네비게이션 방법론에 대해 연구를 수행할 계획이며 토픽맵 리파지토리 또한 연구대상이다.

참고문헌

- [1] Michel Biezunski, Martin Bryan, Steve Newcomb, ISO/IEC 13250 TopicMaps.
- [2] S. Pepper, G. Moore, "XML Topic Maps(XTM) 1.0", TopicMaps.Org.
- [3] G. Moore, "Topic Map technology - the state of the art", XML 2000. Conference & Exposition.
- [4] Rafal Ksiezzyk, "Answer is just a question [of matching Topic Maps]", XML 2000 Conference & Exposition.
- [5] Steve Pepper, "The TAO of Topic Maps", XML 2000 Conference & Exposition.
- [6] Hans Holger Rath, "Making Topic Maps more colourful", XML 2000 Conference & Exposition.
- [7] Thomas Hofmann, "Probabilistic Topic Maps: Navigating through Large Text Collections", IDA, 1999, pp. 161-172.
- [8] Steve Pepper, "Navigating Haystacks and Discovering Needles: Introduce the New Topic Map Standard", Markup Languages 1(4), 1999, pp. 47-74.
- [9] Product White Paper, "The Ontopia Knowledge Suite", Ontopia, Inc.
- [10] Hans Holger Rath, "Technical Issues on Topic maps", 1999.
- [11] L. Olfman, J. Morrison, "Knowledge Management, Organizational Memory, & Organizational Learning - Introduction", HICSS, 2000.
- [12] John Kingston, Ann Macintosh, "Knowledge management through multi-perspective modeling: representing and distributing organizational memory", Knowledge Based Systems 13(2-3), pp. 121-131, 2000.
- [13] P. Mitra, G. Wiederhold, M. L. Kersten, "A Graph-Oriented Model for Articulation of Ontology Interdependencies", EDBT 2000, pp. 86-100.
- [14] Dieter Fensel, Ian Horrocks, Frank van Harmelen, Stefan Decker, Michael Erdmann, Michel C. A. Klein, "OIL in a Nutshell", EKAW 2000, pp. 1-16.
- [15] Ora Lassila, Ralph R. Swick, "Resource Description Framework (RDF) Model and Syntax Spec", W3C Recommendation, 22 Feb 1999, URL: <http://www.w3.org/TR/REV-rdf-syntax>.
- [16] Shawn Bowers, Lois Delcambre, "Representing and transforming model-based information", In Proceedings of the International Workshop of the Semantic Web (SemWeb), in conjunction

with ECDL 2000, Lisbon, Portugal, September 2000.



정 호 영

2000년 : 한국외국어 대학교 컴퓨터 공학과(학사)

2002년 : 한국외국어 대학교 컴퓨터 공학과(석사)

2003년~현재 : 서울대학교 컴퓨터 공학과 박사과정 재학 중

관심분야 : Semantic Web, XML

E-mail : hyjung@oopsla.snu.ac.kr



김 정 민

1992년 : 홍익대학교 전자계산학과(학사)

1994년 : 홍익대학교 대학원 전자계산학과(석사)

2002년 : 서울대학교 대학원 컴퓨터공학부 박사 수료

관심분야 : Semantic Web, Ontology, Graph Index

ontology, Graph Index



정 준 원

2000년 : 동국대학교 컴퓨터 공학과 (학사)

2003년 : 서울대학교 컴퓨터 공학과 (석사)

2003년~현재 : 서울대학교 컴퓨터 공학과 박사과정 재학 중

관심분야 : Semantic Web, Ontology, XML

ontology, XML



김 형 주

1982년 : 서울대학교 전산학과 (학사)

1985년 : Univ. of Texas at Austin (석사).

1988년 : Univ. of Texas at Austin (박사).

1998년 5월~1988년 9월 : Univ. of Texas at Austin. Post-Doc.

1988년 9월~1990년 12월 : Georgia Institute of Technology(부교수).

1991년 1월~현재 : 서울대학교 컴퓨터공학부 교수