

의미 네비게이션을 지원하는 온톨로지 기반 한의학 논문 검색 시스템 설계 연구

고유미 · 엄동명
한국한의학연구원

Abstract

The study on the design of Korean Medical Article Retrieval System Supporting Semantic Navigation based on Ontology

Ko Youmi, Eom Dongmyung
Korea Institute of Oriental Medicine

This study is to design a Semantic Navigation Retrieval System for Oriental Medicine Articles based on a XTM so that people can search and use them more effectively than before. Keywords extracted from articles are categorized 4 topics : herbs, prescription, disease, and action. Keywords analysis Ontology is modeled based on 4 topics and their relations, and then represented Topic maps. Next, Article analysis Ontology is consist of title, author, keywords, abstracts and organization Topics from metadata. Keywords and Article analysis Ontology were integrated through Keywords Topic. Korean Medical Article Retrieval System is optimistic in terms on search results supporting semantic navigation in the information service aspects and easier accessibility because all related information are semantically connected with each different DBs.

Keywords : ontology, Topic Maps, retrieval system

I. 서론

현재 한국한의학연구원에서는 2005년부터 전통지식의 권리 보호와 한의학 연구 활성화를 위하여 '전통지식 DB 구축 사업'을 추진해오고 있다. 이 사업에서는 기존의 방식과 같이 서지 정보 중심의 DB를 구축하고 추가적으로 논문의 내용을 분석하여 처방, 병증, 약재, 작용의 의미 카

테고리에 따라 DB를 구축하였으나, 키워드 기반 검색 기능만이 제공되기 때문에 구축된 DB의 분석 내용을 활용하는 검색을 할 수 없게 되어 있다.

키워드 기반 검색 방식은 이용자가 입력한 검색어를 가지고 단순한 문자열 매칭으로 하기 때문에 검색 결과로 관련이 없거나, 중복되는 정보가 검색되고, 대용량의 정보 검색에서는 검색 결과가 너무 많아 검색의 의미가 없게 된다. 이용자

가 직접 적합한 단어를 입력해야 하기 때문에 전문적인 분야에 대해서는 검색이 어렵고, 모호한 단어를 사용할 경우 검색의 정확성을 보장할 수 없다. 검색 결과는 다시 사람이 직접 확인하여 적합한 정보들을 필터링 해야 하기 때문에 검색 시간이 오래 걸리고, 적합한 정보를 찾기가 어렵다. 키워드 기반 검색의 문제점을 개선하기 위하여 다양한 랭킹 기법과 필드 제한 검색 등의 방법이 사용되고 있지만 정보 사이의 의미적인 관계 처리가 불가능하기 때문에 그 효과에 한계가 있다.

이러한 문제를 개선하기 위하여 의미 검색을 지원하는 온톨로지 기술이 등장하였다. 온톨로지는 정보 처리의 패러다임을 인간위주에서 컴퓨터 위주로 변화시키는 최신 기술로써, 실세계의 객체를 컴퓨터가 이해할 수 있도록 모델링하는 것이다. 따라서 온톨로지는 차세대 정보 기술의 핵심으로 부각되고 있으며, 다양한 분야에서 온톨로지 기반의 검색 시스템을 구축하기 위한 연구가 진행 중에 있다.

본 연구의 목적은 논문 분석 DB의 활용을 극대화하기 위하여 의미 네비게이션을 지원하는 온톨로지 기반 논문 검색 시스템을 토픽맵 기반으로 구현하는 것이다.

논문 분석 DB 활용을 극대화하기 위해서 현재 한국한의학연구원의 '전통지식 DB 구축 사업'을 참고하여 온톨로지를 모델링 하였다. 논문의 키워드를 추출하여 의미적으로 분류한 논문 분석 DB의 스키마에서 각각의 DB는 토픽이 되고, 그 관계를 어소시에이션으로 정의하여 논문 키워드 분석 온톨로지를 구축하였다. 키워드와 연계할 실제 논문 데이터는 논문의 서지사항을 분석하여 주요 검색 필드를 토픽으로 선정하고, 논문 온톨로지를 모델링하여 키워드 분석 온톨로지와의 연계하였다. 모델링 된 온톨로지는 토픽맵으로 구현하고, XTM으로 작성하여 검색 시스템을 구축하였다.

의미 네비게이션을 지원하는 온톨로지 기반 논

문검색 시스템은 토픽맵을 기반으로 하기 때문에 토픽간의 연결 관계를 명확하게 정의하고 있다. 이런 탄탄한 온톨로지 구조를 바탕으로 키워드 기반 검색의 문제점을 해결하고, 각각의 키워드 분석 DB를 의미적으로 연계하여 활용할 수 있다. 따라서 이용자는 의미 검색 기반의 편리하고 정확한 한의학 논문 검색 시스템을 이용할 수 있다.

II. 온톨로지 모델링

1. 전통지식 DB 구축 모델

본 연구에서는 정보 서비스의 효율을 극대화하기 위하여 단순한 형태적인 접근만이 아니라 논문의 내용을 분석하고 추출된 키워드를 의미적으로 분석하였으며, 분석된 내용에 따라 키워드를 분류하고 그 카테고리에 따라 DB를 구축하였다. 이러한 방식으로 논문, 병증, 처방, 약재, 생화학 정보, 고서 등의 데이터베이스를 구축하여 플렉스 트, 필드, 온톨로지 검색 등을 지원하는 검색 시스템이 가능하도록 하였다.

논문과 관련된 의미 정보는 토픽맵 기술을 통하여 연결하고, 의미 네비게이션을 지원하는 온톨로지 기반 논문 검색 시스템을 구축하기 위하여 DB 스키마를 분석하였다. 한의학 논문 분석 DB는 논문의 키워드를 병증, 처방, 약재 등으로 나누어 각각의 DB를 구축하였다. 이는 단순히 키워드를 분류하여 DB에 입력한 것이기 때문에 관리자의 측면에서 검색시스템 구축을 용이하게 하지만 키워드 기반 검색 환경에서 이용자는 검색 결과나 편의성 등의 개선을 느낄 수 없다. 따라서 현재의 논문 분석 DB에 온톨로지 기반 검색 시스템을 구축하게 되면, 의미적인 연관관계를 통하여 각각의 정보를 연계하기 때문에 이용자는 원하는 정보를 더 빠르게 찾아낼 수 있으며, DB 활용의 측면에서도 이용자에게 더 많은 정보를 제공할 수 있다.

2. 논문 키워드 분석 온톨로지 모델

본 연구에서는 논문의 내용을 처방, 약재, 병증의 세 가지 의미로 분석하였다. 즉, 논문을 분석하여 추출된 키워드를 각각 처방, 약재, 병증 DB에 구별하여 입력하고, 이를 키워드 기반 방식으로 내용을 검색한다. 각각 분석하여 카테고리화 된 처방과 약재, 병증의 키워드는 서로 연관관계를 가지고 있다. 그러나 키워드 기반 검색에서 이용자는 이러한 내용 분석 키워드 DB를 활용할 수 없다. 따라서 온톨로지를 적용하여 이러한 의미적인 연관관계를 연결한다면 이용자는 원하는 정보를 더욱

빠르고 쉽게 찾아낼 수 있다.

토픽맵 기반의 온톨로지 모델링을 위하여 논문 내용에서 키워드를 추출하여 분석한 결과와 전통 지식의 DB 카테고리를 중심으로 토픽을 선정하였다. 따라서 병증, 처방, 약재, 작용의 토픽에 키워드를 분류하였고, 이들은 의미적인 연결 관계를 가지고 있다. 예를 들어, 병증에 따라 처방을 하게 되고, 처방은 약재로 구성된다. 그리고 처방은 인체에 어떤 작용을 하게 된다. 아래에 예시한 표는 선정된 토픽과 토픽들 사이의 의미적인 관계를 어소시에이션과 롤로 정의한 것이다.

표 1. 논문 내용 분석 토픽, 어소시에이션, 롤의 관계

Topic 1	Role 1	Association	Role 2	Topic 2
처방	~을 처방하다	처방-병증 관계(~의 병증)	~의 처방이다	병증
처방	~을 구성한다	약재-처방 관계(~의 구성 약재)	~로 구성되다	약재
처방	~의 작용이다	처방-작용 관계(~의 효과)	~작용을 한다	작용
약재	~의 효과이다.	약재-작용 관계(~의 효과)	~효과가 있다	작용
약재	~을 이용하다	약재-병증 관계(~의 이용)	~에 이용되다.	병증

위의 표에서 정의된 토픽, 롤, 어소시에이션들의 관계를 기반으로 토픽맵 기반 온톨로지를 모델링하였다. 다음 그림은 논문에서 추출된 키워드를 기반으로 내용을 분석하고 이를 토픽맵으로 표현한 온톨로지 그림이다. 약재, 처방, 병증, 작

용이라는 토픽들은 논문에서 추출된 키워드로 의미적인 분석과정을 통해서 토픽으로 선정하였다. 각 토픽간의 연결관계는 어소시에이션으로 정의하고, 화살표를 통해서 롤을 정의하여 어소시에이션으로 연결하고 있다.

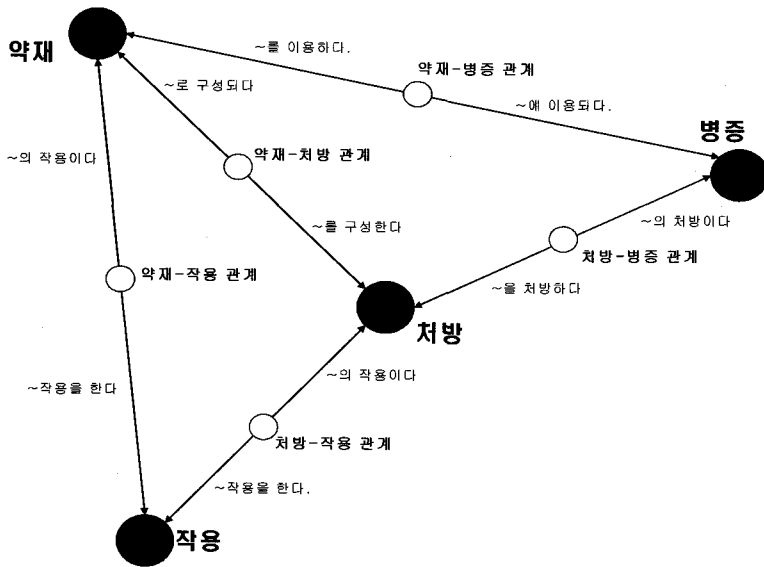


그림 1. 논문 내용 분석 온톨로지 모델의 토픽맵 표현

위의 그림에서 각각의 토픽들은 이용자가 검색을 하는데 의미적인 접근 포인트가 될 수 있다. 이는 전공 서적에서 제공하는 처방, 약재, 병증별 색인으로 각각의 키워드를 통해서 원하는 정보를 찾을 수 있는 것과 같은 의미이다. 기존의 키워드 검색은 각각의 DB가 있어도 이용자가 입력한 키워드에 맞는 결과만 검색 되거나 이용자가 직접 DB를 제한해야 했지만, 온톨로지 검색의 경우에는 각각의 토픽이 어소시에이션으로 연결되어 있기 때문에 이용자는 어떤 정보를 접하더라도 다른 토픽과 연결하여 적합한 정보를 찾을 수 있게 된다.

3. 논문 온톨로지 모델

논문 키워드 분석 온톨로지는 논문의 내용을 분석하여 추출한 키워드 간에 의미적인 관계를 맺어 준 것이기 때문에 우리가 실제적으로 검색

할 논문 정보를 연계해야 한다. 따라서 실제 논문 자체에 대한 정보를 가지고 있는 서지 정보 중심의 논문 온톨로지를 모델링하기 위해서 논문의 서지사항을 중심으로 토픽을 선정하였다. 아래의 표는 논문의 서지정보를 정리한 것이다. 논문은 기본적으로 ISBN 또는 ISSN을 가지고 있다. 논문의 공통적인 서지 정보로는 크게 저자정보, 내용정보, 발행 사항, 그리고 형태사항으로 구별할 수 있으며, 저자정보에는 한, 영문 저자명과 소속기관명이 있다. 내용정보에는 한글, 영어 제목과 국, 영문 초록, 그리고 한, 영 키워드가 해당된다. 다음 발행정보에는 발행처와 발행 날짜, 그리고 권호 사항이 있으며, 형태사항으로 페이지수 또는 수록 페이지가 필요하다. 그리고 논문의 발행 형태에 따라서 석사 논문에는 학위명이 필요하고, 학술 논문에는 학술지명이 서지정보에 해당된다.

표 1. 논문의 서지 사항

공통	출판 등록	ISBN/ISSN number
	저자 정보	한글 저자명
		영문 저자명
		한글 소속기관명
		영문 소속기관명
	내용 정보	국문 제목
		영문 제목
		한글 키워드
		영문 키워드
		국문 초록
		영문 초록
	발행사항	발행처
		발행날짜
		권호사항
형태	페이지 수(수록페이지)	
학위 논문	학위정보(석사/박사)	
학술 논문	학술지명	

위 표에서 제시된 서지사항 모두가 논문과 의미적으로 관계된 토픽들이 될 수 있지만 본 연구에서는 의미 정보 기반의 논문 온톨로지 검색 시스템 구축을 목표로 하기 때문에 내용과 관계된 논문의 제목과 키워드 초록, 그리고 여기에 저자 정보까지를 온톨로지 모델링 범위로 제한하였다. 따라서 논문 온톨로지의 토픽은 국문제목, 영문

제목, 키워드, 영문초록, 국문초록, 저자명, 영문 저자명, 소속기관명으로 선정하였고, 선정된 토픽들 중 주요 접근 포인트가 되는 제목, 저자, 키워드, 소속기관명을 중심으로 온톨로지의 관계를 정의하였다. 아래 표는 각각의 토픽과 어소시에이션, 롤을 정의한 표이다.

표 2. 논문 서지 분석 토픽, 어소시에이션, 롤의 관계

Topic 1	Role 1	Association	Role 2	Topic 2
논문	~의 키워드이다	논문-키워드 관계	~은 키워드가 된다	키워드
논문	~을 저작하다	논문-저자 관계	~에 의해 저작되다	저자
논문	~을 소유하고 있다	논문-기관 관계	~에서 소유되어 있다	기관
저자	~의 소속기관이다	저자-기관 관계	~에 소속되어 있다	기관

아래의 그림은 위의 표에서 정의한 관계를 바탕으로 논문과 저자명, 키워드, 기관명을 토픽으로 선정하고, 온톨로지 모델링 한 것을 토픽 맵 형식으로 표현한 것이다.

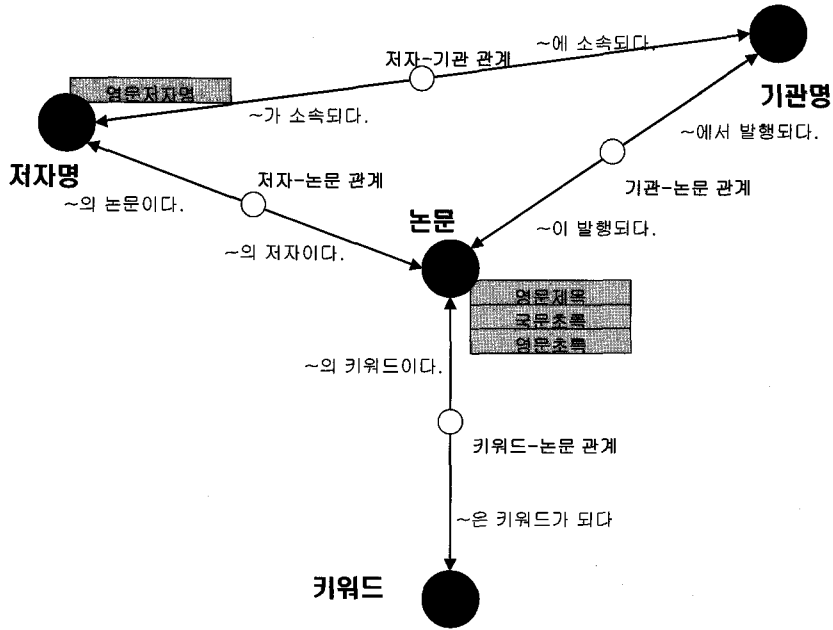


그림 2. 논문 온톨로지 모델의 토픽맵 표현

논문은 국문 제목을 기본이름(<baseName>)으로 가지며, 영문제목과 국문 초록, 영문초록을 어커런스를 통하여 연결하였다. 특히 ‘키워드’ 토픽은 앞의 그림 1.에서 모델링한 내용 분석 온톨로지의 해당 토픽과 의미적으로 연결되기 때문에 논문 검색에서 분석 키워드와 논문 온톨로지를 통합하기 위한 중요한 토픽이다.

4. 의미 기반 논문 검색 온톨로지 모델

그림 1.과 그림 2.의 온톨로지는 각각 키워드를 중심으로 연결하여 하나의 의미 네비게이션을 지원하는 논문 검색 온톨로지의 모델이 된다. 논문의 내용과 서지 정보를 중심으로 각각 모델링 한 두 개의 온톨로지를 연결한 논문 검색 온톨로지 모델을 토픽맵으로 표현하면 다음과 같다.

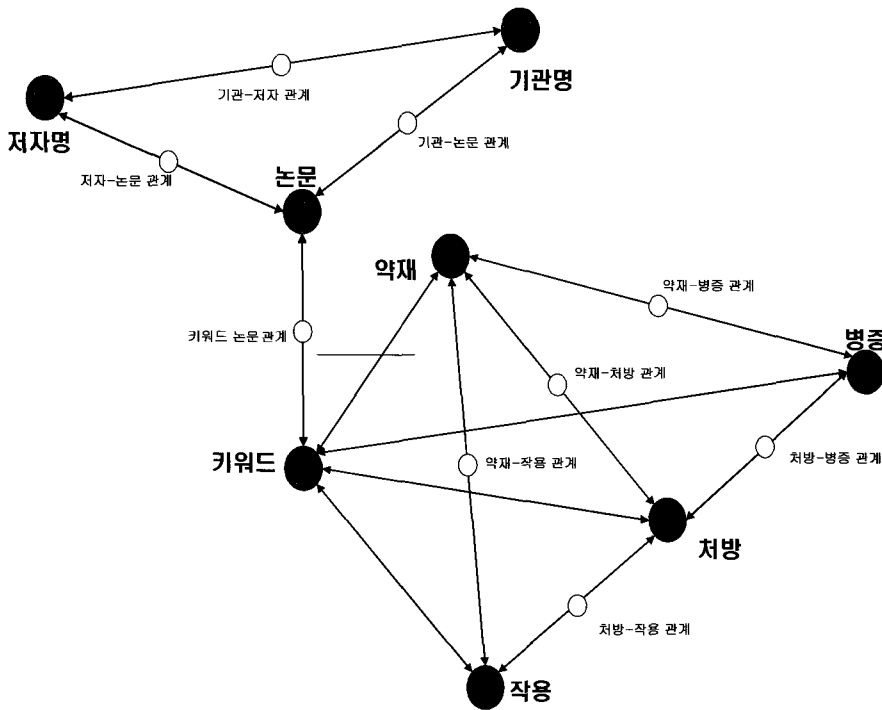


그림 3. 한의학 논문 검색 온톨로지 모델

이용자는 논문의 저자 정보인 제목과 저자명, 소속기관, 그리고 키워드를 통하여 논문을 찾을 수 있다. 뿐만 아니라 내용이 되는 처방, 병증, 약재, 작용의 네 가지 토픽은 키워드를 통해서 논문에 접근 할 수 있고, 이용자는 의미적으로 연결된 각각의 토픽을 네비게이션하면서 적합한 의미를 찾을 수 있기 때문에 검색의 정확성이 향상되고, 원하는 정보를 빠르게 찾을 수 있다.

게 표현할 수 있고, 따라서 여러 온톨로지를 의미나 구조적인 손상 없이 연결하거나 통합할 수 있다. 본 연구는 한의학 논문 DB를 기반으로 검색 분야에 기초적인 온톨로지를 적용하였고, 따라서 일반화된 몇 개의 개념을 중심으로 간단한 온톨로지를 모델링 하였다. 앞으로 다양한 분류체계나 도메인 중심의 온톨로지를 충분히 수용 가능한 확장성을 고려하여 토픽맵으로 표현하였다.

III. 모델링을 기반으로 한 토픽맵 작성

지식관리를 목적으로 하는 토픽맵은 복잡한 구조를 가지고 있기 때문에 정보의 구조를 명확하

1. 내용 분석 온톨로지의 XTM 작성

최종적으로 모델링 된 그림 3.의 온톨로지를 구현하기 위해서 XTM을 작성하였다.

표 4. 그림 1의 내용 분석 온톨로지의 토픽의 XTM 선언

```

<topic id="prescription">
  <baseName><baseNameString>처방</baseNameString></baseName>
</topic>
<topic id="medicine">
  <baseName><baseNameString>약재</baseNameString></baseName>
</topic>
<topic id="disease">
  <baseName><baseNameString>병증</baseNameString></baseName>
</topic>
<topic id="action">
  <baseName><baseNameString>작용</baseNameString></baseName>
</topic>

```

표 4는 그림 1의 내용 분석 온톨로지에서 선정한 처방, 약재, 병증, 작용 네가지 토픽을 XTM으로 선언한 것이다. <topic>은 토픽을 선언하는 역할을 하며, 기본이름(<baseName>)을 갖게 된

다. 각각의 토픽들은 인스턴스를 갖게 되고 실제 연결관계인 어소시에이션은 인스턴스 레벨에서 이루어지게 된다.

표 5. 내용 분석 온톨로지의 인스턴스 입력 예

```

<topic id="p00002">
  <instanceOf><topicRef xlink:href="#prescription" /></instanceOf>
  <baseName><baseNameString>육미지황탕</baseNameString></baseName>
</topic>
<topic id="m00001">
  <instanceOf><topicRef xlink:href="#medicine" /></instanceOf>
  <baseName><baseNameString>조구등</baseNameString></baseName>
</topic>
<topic id="d00003">
  <instanceOf><topicRef xlink:href="#disease" /></instanceOf>
  <baseName><baseNameString>부중</baseNameString></baseName>
</topic>
<topic id="a00002">
  <instanceOf><topicRef xlink:href="#action" /></instanceOf>
  <baseName><baseNameString>보신보혈</baseNameString></baseName>
</topic>

```

위의 표는 내용 분석 온톨로지의 처방, 약재, 병증, 작용에 대하여 하나의 인스턴스를 입력한 XTM의 예로, <instanceOf>를 사용하여 상위

토픽과 연결해준다. 처방의 인스턴스는 ‘육미지황탕’ 이고, 약재의 인스턴스로는 ‘조구등’, 병증은 ‘부중’, 그리고 작용에는 ‘보신 보혈’ 이

인스턴스로 연결되어 있다.

아래의 표는 토픽과 토픽 사이의 연결인 어소시이션을 작성하는 예로, 처방과 약제의 관계를 표현하고 있는 XTM이다. 어소시에이션은 먼저

토픽으로 선언을 하여 어소시에이션의 이름과 각각의 룰을 정의하고, 실제로 인스턴스 간의 연결을 할 때 <association>을 사용한다.

표 6. 처방과 약제의 어소시에이션 예

```

<!-- 처방과 약제의 관계 -->
<topic id="consistof">
  <baseName>
    <baseNameString>구성된다.</baseNameString>
  </baseName>
  <baseName>
    <scope><topicRef xlink:href="#medicine" /></scope>
    <baseNameString>을 구성한다.</baseNameString>
  </baseName>
  <baseName>
    <scope><topicRef xlink:href="#prescription" /></scope>
    <baseNameString>로 구성된다.</baseNameString>
  </baseName>
</topic>
<association>
  <instanceOf>
    <topicRef xlink:href="#consistof" />
  </instanceOf>
  <member>
    <roleSpec><topicRef xlink:href="#prescription" /></roleSpec>
    <topicRef xlink:href="#p00001" />
  </member>
  <member>
    <roleSpec><topicRef xlink:href="#medicine" /></roleSpec>
    <topicRef xlink:href="#m00001" />
  </member>
</association>
    
```

‘구성된다’ 라는 어소시에이션에는 ‘~을 구성한다’ 와 ‘~로 구성된다’ 라는 룰을 포함하고 있다. 그리고 연결되는 인스턴스는 ‘p00001’ 즉, ‘육미지황탕’ 과 ‘m00001’ 즉, ‘육미지황탕’ 의 구성 약제 중 하나가 ‘조구등’ 이라는 의미적인 연결 관계를 맺어주고 있다.

2. 서지 정보 중심의 논문 온톨로지 XTM 작성

다음의 표는 그림 2.의 논문 온톨로지 모델의 토픽을 XTM으로 선언한 것이다. 내용과 관련된 논문 제목, 영문 제목, 국문초록, 영문초록, 키워드 정보와 저자와 관련된 저자명, 영문저자명, 소속기관명을 각각 토픽으로 선언하였다.

표 7. 그림 2.의 논문 온톨로지 모델의 토픽 선언 예

```

<topic id="titlekor">
  <baseName><baseNameString>논문제목</baseNameString></baseName>
</topic>
<topic id="titleeng">
  <baseName><baseNameString>영문제목</baseNameString></baseName>
</topic>
<topic id="abstkor">
  <baseName><baseNameString>국문초록</baseNameString></baseName>
</topic>
<topic id="absteng">
  <baseName><baseNameString>영문초록</baseNameString></baseName>
</topic>
<topic id="aunamekor">
  <baseName><baseNameString>저자명</baseNameString></baseName>
</topic>
<topic id="aunameeng">
  <baseName><baseNameString>영문저자명</baseNameString></baseName>
</topic>
<topic id="organkor">
  <baseName><baseNameString>소속기관명</baseNameString></baseName>
</topic>
<topic id="keyword">
  <baseName><baseNameString>키워드</baseNameString></baseName>
</topic>

```

선언된 토픽은 각각 접근 포인트가 되는 논문제목, 영문 제목, 국문 초록, 영문 초록, 저자명, 영문 저자명, 소속기관명, 키워드 등 여 인스턴스를 입력하였다. 목과 키워드, 저자명과 소속기관명을 중심으로 하

표 8. 논문 온톨로지 인스턴스 입력 예

```

<topic id="titlekor_ART0000000001">
  <instanceOf><topicRef xlink:href="#titlekor" /></instanceOf>
  <subjectIdentity>
    <subjectIndicatorRef xlink:href="http://localhost:8080/fulltxt/ART0000000001.tif" />
  </subjectIdentity>
  <baseName><baseNameString>평간약(平肝藥)의 항고혈압작용에 관한 연구</baseNameString></baseName>
  <occurrence>
    <instanceOf><topicRef xlink:href="#titleeng" /></instanceOf>
    <resourceData>Study on the Anti-Hypertensive Effects of the Liver-Calming Drugs</resourceData>
  </occurrence>
  <occurrence>
    <instanceOf><topicRef xlink:href="#absteng" /></instanceOf>
    <resourceData>

```

Drugs whose principal effects are calming the liver to check endogenous wind and relieve convulsion and calming the liver to suppress hyperactivity of the liver-yang are called 'drugs for calming the liver'.

These drugs are chiefly indicated for dizziness and blurred vision due to hyperactivity of the liver-yang, and spasms and convulsions due to up-stirring of endogenous wind resulting from disorders of the liver.

Some drugs in this class also have the effects of clearing away heat from the liver and are therefore effective for blood-shot, swollen and painful eyes and headache due to intense heat in the liver.

We thought that the symptoms in hyperactivity of the liver-yang in thaditional Korean medicine is similar to the signs of hypertensive patients in some case.

We chose 7 kinds of drugs for calming the liver and observed the antihypertensive effect on male S.H.R. and S.D rats, direct relaxant activityon isolated rabbit aorta, and ACE inhibitory activity.

There was significant antihypertensive effect of Gastrodiae Rhizoma, Scorpio, Lumbricus and Scolopendra in S.H.R, and antihypertensive tendency of Uncariae Ramulus et Uncus, Cassiae Semen, Bombyx Batryticatus.

Therefore we concluded that there was some correlation between liver-calming and antihypertensive activity.

```

</resourceData>
</occurrence>
</topic>
<topic id="aunamekor01_ART0000000001">
  <instanceOf><topicRef xlink:href="#aunamekor" /></instanceOf>
  <baseName><baseNameString>김 호철</baseNameString></baseName>
  <occurrence>
    <instanceOf><topicRef xlink:href="#aunameeng" /></instanceOf>
    <resourceData>Kim Ho Cheol</resourceData>
  </occurrence>
</topic>
<topic id="organkor01_ART0000000001">
  <instanceOf><topicRef xlink:href="#organkor" /></instanceOf>
  <baseName><baseNameString>경희대학교</baseNameString></baseName>
</topic>

```

위의 표는 논문 온톨로지 인스턴스를 입력한 예를 보여주고 있다. 논문 인스턴스는 논문 토픽에서 <subjectIdentity>로 파일을 지정하고 있다. 위의 표에서 <subjectIndicatorRef>로 현재 서버에 저장되어 있는 파일의 위치 경로를 참조한다. 차후 이것은 범용성을 고려하여 URI로 지정해야 한다. 다음 기본이름(<baseName>)

은 논문 제목으로 입력하고 어커런스를 통해서 <resourceData>로 영문 제목과 영문 초록, 국문 초록을 입력한다. 저자명 역시 한글 저자명을 기본 이름으로 하고, 영문이름은 어커런스로 지정한다. 마지막으로 소속기관의 인스턴스인 '경희대학교' 를 작성한 예이다.

표 9. 논문 제목과 키워드 관계 예

```

<topic id="re_titlekor_keyword">
  <baseName>
    <baseNameString>키워드이다</baseNameString>
  </baseName>
  <baseName>
    <scope><topicRef xlink:href="#titlekor" /></scope>
    <baseNameString>의 키워드를 가지고 있다.</baseNameString>
  </baseName>
  <baseName>
    <scope><topicRef xlink:href="#keyword" /></scope>
    <baseNameString>의 키워드이다.</baseNameString>
  </baseName>
</topic>
<association>
  <instanceOf>
    <topicRef xlink:href="#re_titlekor_keyword" />
  </instanceOf>
  <member>
    <roleSpec><topicRef xlink:href="#titlekor" /></roleSpec>
    <topicRef xlink:href="#titlekor_ART0000000003" />
  </member>
  <member>
    <roleSpec><topicRef xlink:href="#action" /></roleSpec>
    <topicRef xlink:href="#a00004" />
  </member>
</association>

```

위의 표는 논문제목과 키워드의 관계를 나타낸 토픽맵 작성 예이다. 어소시에이션은 토픽을 선언하여 정의해주고, <association>으로 논문제목과 키워드를 <member>로 구성하여 앞에서 정의한 어소시에이션으로 연결한다. 어소시에이션 명은 '키워드이다'이며 각각의 롤은 '~의 키워드를 가지고 있다' 와 '~의 키워드이다' 로 정의하였다. 정의된 어소시에이션은 논문과 키워드를 연결하는 의미적인 연결 관계를 만든다.

IV. 토픽맵을 기반으로 한 프로토타입 구현

본 연구에서 진행된 토픽맵 기반 온톨로지 모델은 DB에 구축된 내용 중 일부 샘플을 대상으로 하여 인스턴스를 입력하고 검색 시스템의 프로토타입을 구현하였다. 샘플 데이터는 총 3개의 논문을 바탕으로 각각의 서지정보와 키워드를 추출하였고, 분석 결과 키워드는 약제가 22개, 처방이 4개, 병증이 5개, 작용이 4개로 총 35개가 분석되었다. 아래 표는 대상이 된 3개의 논문 제목과 해당 키워드를 나타낸 것이다. 이것은 실제 논문 온톨로지의 인스턴스로 입력된다.

표 10. 논문제목과 키워드 인스턴스

번호	논문제목	관련 키워드	
1	六味地黄湯 및 四物湯이Puromycin Aminonucleoside로 誘發된 白鼠의 腎症에 미치는 影響	처방	육미지황탕, 사물탕
		병증	부중, 수중, 신중후군
		약재	숙지황, 산약, 산수유, 백복령, 목단피, 택사, 백작약, 천궁, 당귀
		작용	보신 보혈
2	평간약(平肝藥)의 항고혈압작용에 관한 연구	처방	
		병증	간풍, 간양상항, 고혈압
		약재	조구등, 천마, 백강잠, 전갈, 오공, 구인, 결명자, 평간약
3	桂枝茯苓丸加味方の 抗癌活性과 免疫調節作用에 關한 研究	처방	계지복령환 가미방
		병증	어혈로 인한 부인과 질환
		약재	계지, 복령, 영지, 목단피, 도인, 백작약
		작용	항암활성, 면역조절

다음의 표는 논문의 키워드 분석 결과를 중심으로 인스턴스를 나열한 것이다. 키워드를 의미적인 구분에 따라 처방, 병증, 약재, 작용으로 구별하

여 표로 작성하였다. 각각의 인스턴스들은 XTM 작성시 토픽으로 선언되어, 어소시에이션으로 연결해 주었다.

표 10. 논문 키워드 분석 결과 인스턴스

번호	논문 분석	인스턴스
1	처방	육미지황탕, 계지복령환 가미방, 사물탕
2	병증	부중, 수중, 신중후군, 간풍, 간양상항, 고혈압, 어혈로 인한 부인과 질환
3	약재	숙지황, 산약, 산수유, 백복령, 목단피, 택사, 백작약, 천궁, 당귀, 조구등, 천마, 백강잠, 전갈, 오공, 구인, 결명자, 계지, 복령, 영지, 목단피, 도인, 백작약, 평간약
4	작용	보신 보혈, 항암활성, 면역조절

온톨로지에 기반하여 작성된 XTM은 Touchgraph Link Browser¹⁾의 공개 소스를 활용하여 논문, 처방, 약재, 병증에 따라서 검색이 화면을 구성하였고, 논문의 제목, 저자, 키워드로 필드를 한정하여 검색이 가능하도록 하였다.

검색 결과는 먼저 텍스트 화면에서 리스트로 제공된다. 리스트에는 검색 결과 토픽과 관련된 토픽들 그리고 그 의미 관계를 함께 보여주고 있다. 아래 그림은 ‘육미지황탕’에 대한 검색 결과 리스트이다. ‘육미지황탕’은 처방에 해당되며, 온톨로지 결과로 각각 관련된 토픽들과 어소

시에이션 명칭을 제공한다. 즉, ‘육미지황탕’을 구성하는 약재에는 ‘숙지황’ ‘산약’ ‘산수유’ ‘백복령’ ‘목단피’가 있고, 그 어소시에이션은 ‘~로 구성된다’이다. 관련된 병증은 ‘~을 치료한다’라는 어소시에이션으로 ‘부중’ ‘수중’ ‘신중후군’이 있다. ‘육미지황탕’의 작용은 ‘~에 효능이 있다’는 어소시에이션으로 ‘보신 보혈’과 연결되어 있다. 최종적으로 ‘육미지황탕’을 키워드로 갖고 있는 논문은 ‘~의 키워드를 가지고 있다’는 어소시에이션으로 논문을 연결한다.

1) Touchgraph Link Browser <<http://www.touchgraph.com/newLB/TGLinkBrowser.html>> (2005.10.26)

번호	검색어 결과	온톨로지 결과
1	처방 : 육미지황탕 ☼	약재 : 숙지황 [로 구성된다.] 약재 : 산약 [로 구성된다.] 약재 : 산수유 [로 구성된다.] 약재 : 백복령 [로 구성된다.] 약재 : 육단피 [로 구성된다.] 약재 : 맥사 [로 구성된다.] 병증 : 신경후군 [을 치료한다.] 병증 : 부종 [을 치료한다.] 병증 : 수종 [을 치료한다.] 작용 : 보신 보혈 [에 효능이 있다.] 논문제목 : 六味地黃湯 및 四物湯이 Puromycin Aminonucleoside로誘發된 白鼠의 腎症에 미치는 影響 [의 키워드를 가지고 있다.]

그림 4. ‘육미지황탕’ 검색 결과 리스트

리스트로 제공된 검색 결과는 검색 결과에 대하여 한눈에 보기 쉽게 정리가 되어 있지만, 온톨로지에서 제공하는 연결관계를 통한 의미 네비게이션이 불가능하다. 즉, 리스트 정보는 ‘육미지황탕’ 과 연관이 있는 의미들만 볼 수 있다. 따라서 온톨로지 모델을 이미지화한 터치그래프 툴로 토픽들의 연결관계를 의미관계를 동적으로 네비게이션하는 것이 가능하게 하였다. 사용된 툴은 자

바 기반으로 터치그래프를 제공하며, 좌측에는 온톨로지 모델을 제공하고 오른쪽에는 토픽에 대한 어커런스를 보여준다. 아래 그림은 ‘육미지황탕’ 을 중심으로 관련된 토픽들을 의미적으로 연결한 온톨로지 모델이다. 그림 4.에서 제공하는 것과 같이 구성 약재와 병증, 작용, 그리고 논문 정보를 연결하고 있다.

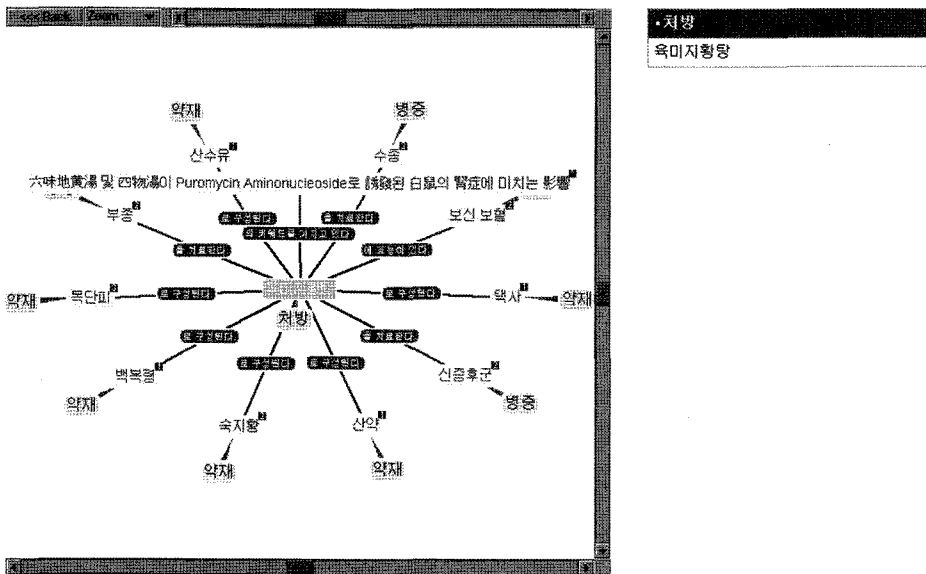


그림 5. ‘육미지황탕’ 의 온톨로지 이미지

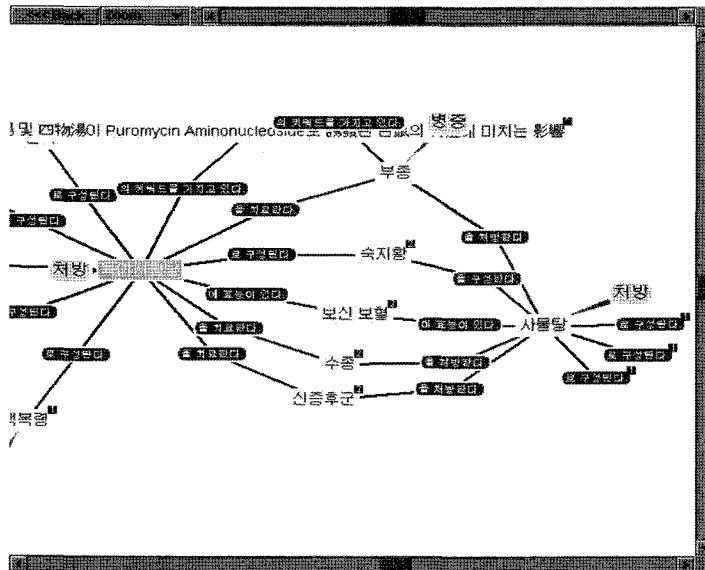


그림 7. '육미지황탕' 과 '사물탕' 의 연결 관계

위의 그림은 '육미지황탕' 과 '사물탕' 의 의미적인 연관관계를 나타낸 온톨로지이다. '부종' 이라는 병증에 관계된 처방으로 '육미지황탕' 과 '사물탕' 은 병증인 '수종' 과 '신증후

군' 에도 관계된 처방이며, 두 처방에는 공통적으로 '속지황' 이라는 약재가 사용된다. 또한 각 처방은 '보신보혈' 작용을 통해서 병증을 치료한다는 것을 알 수 있다.

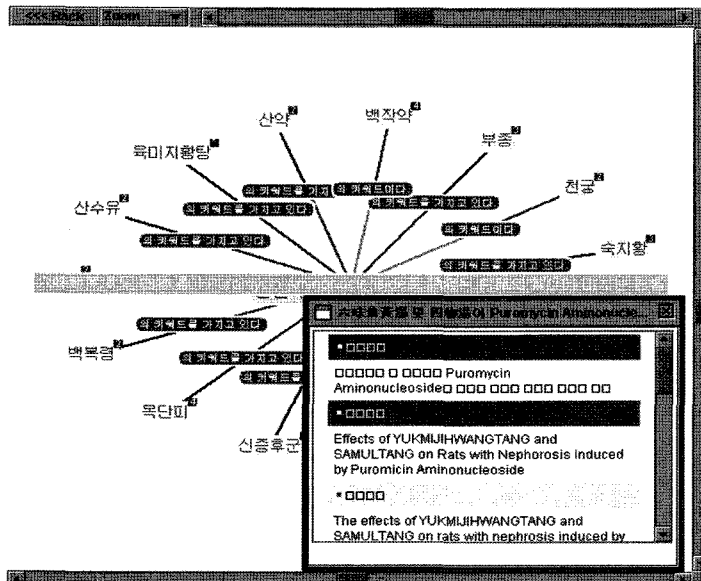


그림 8. 논문 중심의 온톨로지 모델

위의 그림은 ‘육미지황탕’ 과 ‘부종’ 등을 키워드로 갖고 있던 논문을 중심으로 온톨로지를 확장한 것이다. 논문이 가지고 있는 연결관계는 키워드와 저자, 저자와 소속기관명이며, 어커런스를 통해서 영문제목과 국문, 영문 초록정보를 보여준다. 앞에서 어커런스를 우측의 다른 화면에

보여주고 있는데 위 그림에서는 새 창을 띄워서 보여주고 있다. 현재의 화면에서는 논문과 관련된 키워드 및 서지 정보를 연결하고 있다. 이용자는 논문에 연결된 키워드를 선택하여 키워드간의 관계를 파악할 수 있고, 키워드와 연결된 다른 논문에도 의미적으로 접근할 수 있다.

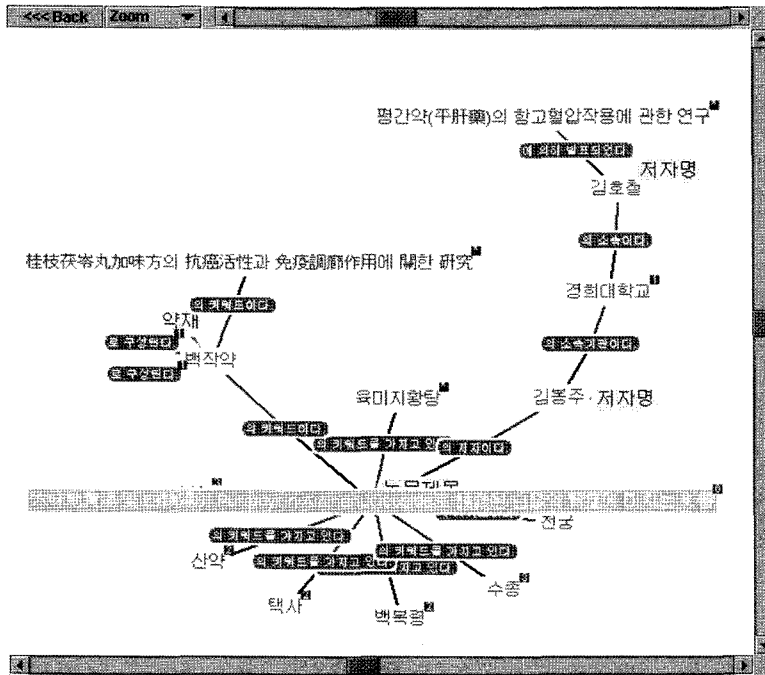


그림 9. 논문의 키워드 및 저자 확장 온톨로지

위의 그림은 논문에서 키워드와 저자를 통해서 다른 논문을 찾는 확장 온톨로지 모델이다. 논문의 키워드가 되는 약재인 ‘백작약’ 은 다른 논문 ‘桂枝茯苓丸加味方の 抗癌活性和 免疫調節作用에 關한 研究’ 의 키워드도 될 수 있기 때문에 ‘백작약’ 을 통해서 논문을 찾을 수 있다. 그리고 논문의 저자 ‘김보주’ 는 ‘경희대학교’ 에 소속되어 있으며, 경희대학교에 소속된 다른 사람 ‘김호철’ 은 ‘평간약(平肝藥)의 항고혈압작용에 관한 연구’ 를 발표하였다는 관계까지 알 수 있다.

VI. 결론

현대의 정보화 사회에서는 잘 구축된 정보를 지식으로서 활용하는 것이 중요하다. 본 연구에서는 가장 대표적인 전통지식이라 할 수 있는 한의학 분야의 연구를 활성화하고, 이를 지식으로서 활용하기 위하여 온톨로지를 적용해 보았다.

온톨로지 모델은 실제적으로 논문의 내용을 분석하여 추출된 키워드를 의미적인 카테고리나 누고, 그 의미에 따라서 토픽을 선정하고 의미적

인 연결을 해주었기 때문에 향후 확장된 논문 검색 시스템 등에 충분히 활용이 가능하다. 온톨로지 모델은 명확한 연관 관계와 구조화가 가능한 토픽맵으로 작성하였고, 이용자의 자유로운 의미 네비게이션이 가능하도록 터치 그래프를 사용하여 프로토타입을 구현하였다.

온톨로지를 적용한 의미 네비게이션 논문검색 시스템은 논문 키워드 분석에 따른 처방, 병증, 약재, 작용 DB의 의미적인 연관 관계를 정의하고 연결하였기 때문에 이용자는 원하는 정보를 찾기 위해 정확한 의미를 파악하지 않아도 네비게이션을 통하여 원하는 정보를 빠르게 검색할 수 있다. 각각 토픽은 정보 검색의 검색 포인트가 되고, 이것은 관련된 정보만을 연결하고 있기 때문에 정보의 필터링이 용이하며, 자신이 원하는 목적에

따라서 정보를 네비게이션 할 수 있어서 검색의 결과가 더 정확하다. 또한 온톨로지의 토픽과 어소시에이션은 다른 온톨로지나 개념들을 통합하고 확장하기 용이한 구조를 제공한다. 따라서 앞으로 다른 분류체계나 시소러스 등을 토픽맵으로 작성하여 온톨로지와 연계, 통합하는 것이 용이하여 확장성 있는 검색 시스템 구축이 가능하다.

차후에는 학회나 관련 특허, 한의학 문헌, 고서와 같은 한의학 관련 정보뿐만 아니라 분류체계, 시소러스 등 주제 분야에 대한 세부적인 접근이 가능한 온톨로지로 확장하여 한의학 분야의 모든 정보를 연계하는 한의학 통합 온톨로지 정보 시스템을 구축하기 위한 연구가 진행되어야 할 것이다.

검색어 : 온톨로지, 토픽맵, 검색 시스템

참 고 문 헌

1. 고유미. 『토픽맵 기반의 특허정보 서비스를 위한 시스템 구축에 관한 연구 (항체이용기술 분야를 중심으로)』 석사학위논문:숙명여자대학교. 2004:9-34
2. 김성혁. 『과학기술 정보자원의 차세대 웹 서비스 체계 구축 연구』 최종보고서:KISTI. 2003:12-50
3. 이병욱, 권영규. 「한국과 중국의 MeSH 호환성 연구」. 『대한한의정보학회지』. 2005;11(2호):65-82
4. 이은아. 『XML 토픽맵(XTM)을 이용한 시멘틱 네비게이션 시스템 구현』 석사학위논문: 숙명여자대학교. 2003:7-27
5. 이정은. 『의미 네비게이션을 지원하는 온톨로지 기반 포털 구축 (인맥 정보와 영화정보를 중심으로)』 석사학위논문:숙명여자대학교. 2004:7-32
6. Michel C. Daconta, LeoJ. Obrst, Kevin T, Smith. 『The Semantic Web : A Guide to the Future of XML, Web Services, and Knowledge Management』 Wiley, Inc., 2003.
7. Steve Pepper. 『Topic Maps : The Tutorial』 2000-2004 Ontopia AS. 2004:25-30
8. Steve Pepper. 『The TAO Topic Maps』 <<http://www.gca.org/papers/xml europe2000/papers/s11-01.html>>(2005.10.22)
9. Touchgraph Link Browser <<http://www.touchgraph.com/newLB/TGLinkBrowser.html>> (2005.10.26)