한국학 연구 논문의 텍스트 구조 기반 메타데이터 검색 시스템 개발 연구*

A Study on Developing a Metadata Search System Based on the Text Structure of Korean Studies Research Articles

송민선 (Min-Sun Song)**
고영만 (Young Man Ko)***
이승준 (Seung-Jun Lee)****

초 록

본 연구는 한국학 연구 논문 텍스트의 의미 구조를 기반으로 하는 메타데이터를 적용한 학술정보시스템을 구축하여 기존 유사 시스템과의 비교를 통해, 텍스트 구조 기반 메타데이터의 활용 가능성을 확인해 보고자 하는 것을 목적으로 한다. 이를 위해 한국학술지인용색인(Korea Citation Index, KCI)에서 일정 기준을 충족하는 한국학 분야 연구 논문 데이터를 대상으로 의미 구조 메타데이터 항목을 적용한 시범적 검색 시스템(Korean Studies Metadata Database, KMD)을 구축하였으며, 동일한 검색 키워드를 적용하여 기존의 KCI 시스템과 비교했을 때 어떤 특징과 차이점을 갖는지 비교해 보았다. 연구 결과, KMD 시스템이 KCI에 비해 이용자의 검색 의도에 맞는 결과를 보다 효율적으로 보여주는 것으로 확인되었다. 즉 검색하고자 하는 키워드의 조합이나 조건식이 기존 시스템과 동일하더라도 검색 결과를 통해 최종적으로 연구 진행과 관련해 찾고자 하는 연구 목적, 연구의 대상 데이터나 시공간적 배경 등에 따른 검색 결과를 다양하게 보여줄 수 있는 것으로 나타났다.

ABSTRACT

This study aims to develope a scholarly metadata information system based on conceptual elements of text structure of Korean studies research articles and to identify the applicability of text structure based metadata as compared with the existing similar system. For the study, we constructed a database(Korean Studies Metadata Database, KMD) with text structure based on metadata of Korean Studies journal articles selected from the Korea Citation Index(KCI). Then we verified differences between KCI system and KMD system through search results using same keywords. As a result, KMD system shows the search results which meet the users' intention of searching more efficiently in comparison with the KCI system. In other words, even if keyword combinations and conditional expressions of searching execution are same, KMD system can directly present the content of research purposes, research data, and spatial-temporal contexts of research *et cetera* as search results through the search procedure.

키워드: 한국학, 연구 논문, 텍스트 구조, 의미 구조, 메타데이터 검색 시스템
Korean studies, research articles, text structure, semantic structure, metadata search system

---

* 이 논문은 2014년 정부(미래창조부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2014R1A2A1A11053124).
** 성균관대학교 정보관리연구소 연구원(songsery@skku.edu) (제1저자)
*** 성균관대학교 문과대학 문헌정보학과 교수(ymko@skku.edu) (교신저자)
**** 성균관대학교 정보관리연구소 연구원(dino@skku.edu) (공동저자)
■ 논문접수일자 : 2016년 8월 17일 ■ 최초심사일자 : 2016년 9월 2일 ■ 게재확정일자 : 2016년 9월 20일
■ 정보관리학회지, 33(3), 155-176, 2016. [http://dx.doi.org/10.3743/KOSIM.2016.33.3.155]

1. 서 론

연구 논문은 연구자의 분석과 논증을 통해 작성되며, 학술커뮤니케이션을 통해 공유되고, 이후 관련 분야에서 생산된 또 다른 연구를 통해 지속적으로 검증과 평가가 이루어진다. 연구 논문은 동일한 학문 분야에 속한다 하더라도 연구 문제를 다루는 다양한 관점이 존재하기도 하고, 학문 분야가 다르더라도 연구 대상 데이터나 연구 방법이 동일하거나 유사할 수 있다. 특히 '지역학'의 경우 해당 지역을 주제로 하는 모든 영역의 학문 분야를 포괄하는 다학제적 특성으로 인해, 포함되는 연구 주제의 범위와 연구 문제, 연구 대상, 시대 및 지역 배경 등이 매우 광범위하고 다양하게 관련을 맺고 있다는 특성을 보인다.

따라서 다학제 분야의 성공적인 연구 수행을 위해서는 무엇보다 관심 있는 주제에 대한 다양한 연구 관점, 연구 방법, 논증 방식 등 해당 주제 연구의 학술적 의미를 파악하고 연구 진행에 필요한 요소의 조망, 즉 학술적 이해 형성(scholarly sense-making)이 선행되어야 한다. 이러한 조망을 통해 해당 주제와 관련된 연구 성과의 유사성과 차별성 등을 이해하고 평가할 수 있게 되면 훨씬 효율적으로 연구를 수행할 수 있을 것이다. 그러나 현재 제공되고 있는 대부분의 학술정보시스템은 검색에 사용된 용어가 내포하고 있는 개념이 해당 검색 결과에서 어떤 의미를 갖는지 보여주지 못하며, 단순히 제목이나 초록 등의 색인 대상 항목에 같은 단어가 포함된 결과만 제시한다. 그 결과 연구의 생산성이 높아지고 연구 논문의 절대량이 기하급수적으로 증가하고 있는 오늘날의 연구 환경에서는 단순히 관련 단어의 검색에 의한 결과를 통해 학술적 이해형성을 위한 정보 요구를 충족시키는 것이 점점 더 힘들어지고 있다.

학술적 이해형성을 위한 정보요구를 충족시키기 위해서는 연구 논문 간의 의미 관계를 식별하고 파악할 수 있는 형식으로 학술정보시스템을 개선하는 것이 필요하다. 그리고 이를 위해서는 연구 논문 텍스트의 의미 구조를 구성하는 개념 요소를 파악할 수 있는 새로운 형식의 지식조직방법을 토대로 학술정보시스템이 구축될 필요가 있다. 이러한 시스템을 통해 연구자들은 연구 과정의 각 단계에서 발생할 수 있는 정보 요구를 충족시키는 데 필요한 내용들을 용이하게 파악할 수 있을 것이다.

본 연구는 다양한 학문 분야를 포괄하는 '한국학' 분야 연구 논문 간의 학술적 이해형성에 필요한 연관 관계를 파악할 수 있는 새로운 형식의 연구 논문 지식조직 방법, 즉 논문의 연구 과정에서 생산된 텍스트 구조에 기반을 둔 메타데이터 스키마를 토대로 시범적 데이터베이스를 구축한 후 기존 유사 서비스에서의 검색 결과와 어떤 차이를 보이는지 비교해 봄으로써 텍스트 구조 기반 메타데이터의 활용 가능성을 확인하고자 하는 것을 목적으로 한다. 즉 연구 논문의 텍스트 구조 기반 메타데이터 시스템이 기존의 논문 형식 구조에 기반한 검색 시스템에 비해 연구자들이 필요로 하는 학술적 정보 요구에 적합한 결과를 얼마나 잘 보여줄 수 있을지의 여부를 실제 시스템의 구현을 통해 확인해 보고자 하는 것이다.

2. 텍스트 구조 분석과 지식조직체계

2.1 텍스트 구조와 텍스트 유형

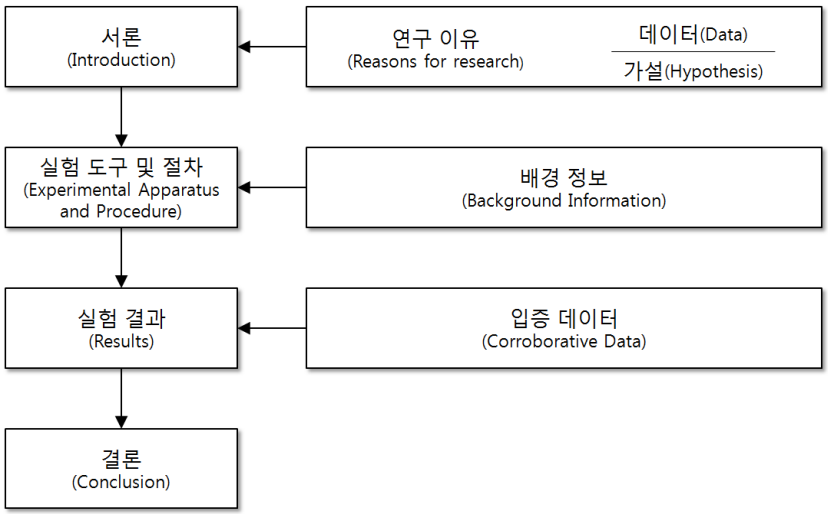
텍스트에 관한 연구의 역사는 고대 그리스의 수사학까지 거슬러 올라가지만, 현대적 의미의 텍스트 구조 분석 연구는 1970년대 이후 기존의 문장 언어학에 대한 비판과 더불어 텍스트의 의미와 구조를 분석함으로써 텍스트의 이해 과정을 밝히기 위한 목적에서 비롯되었다. 최근까지도 학자들에 따라 '텍스트'에 대한 이해와 관점이 다양하게 나타나고 있으나, 텍스트 언어학에서 규정하는 '텍스트'의 일반적 개념은 결속성(cohesion)과 응집성(coherence)을 갖춘 일련의 발화 혹은 문장의 연속체, 즉 서로 관련되어 있는 문장들의 연속체가 구성하고 있는 '문장의 상위 단위(unit beyond the sentence)'라 할 수 있다(박진용, 1997; 정여훈, 2013; Weigand, 2009, p. 151). 이때 결속성은 텍스트 표층 요소들의 통사적 의존 관계와 관련되는 형식적 특성을 말하며, 응집성은 텍스트가 표상하는 의미 내용의 연결 관계와 관련된 내용적 특질을 말한다(de Beaugrande & Dressler, 1981).

텍스트 구조에 관한 연구는 초창기 텍스트 언어학 분야와 인지심리학 분야에서 이루어진 텍스트의 의미 구조 분석 연구 성과를 토대로 발전하였다. 텍스트의 본질 즉 텍스트성에 관해 천착한 텍스트 언어학자의 연구 성과에 따르면(Halladay & Hasan, 1976; de Beaugrande & Dressler, 1981; Vater, 1994), 텍스트의 본질은 텍스트 기저를 의미적으로 연결하는 '의미 구조' 즉 응집성이며, '의미 구조'는 텍스트 기저의 의미 내용을 연결하는 '의미 관계'에 의해 이루어진다. 따라서 텍스트 의미 구조의 분석은 텍스트 구성요소 간의 '의미 관계'를 분석하는 작업이다. 인지심리학 분야의 연구는 독자의 이해형성에 영향을 미치는 텍스트 요인을 규명하기 위해 텍스트를 구성하는 단위들 간의 관계를 표상하는 일반적 체계를 수립하고자 하였으며, 대체로 텍스트의 구조를 의미(내용)의 위계 구조로 파악하였다(Kintsch, 1974; Frederiksen, 1975; Meyer, 1975).

텍스트의 구조 분석에 관한 연구에 있어서 텍스트 언어학 분야와 인지심리학 분야 모두 형식적 특성보다는 의미(내용)의 관계 구조로 파악하는 점에서 일치를 보인다. 이는 텍스트의 본질적 성격은 텍스트의 구성 요소(문장)들이 서로 관련되어 연속된다는 데 있으며, 텍스트 구조의 분석은 텍스트 기저에 자리 잡고 있는 의미의 연결성 즉 의미 구조의 개념요소를 파악하는 것이라는 점을 말해주는 것이라 할 수 있다.

텍스트 구조 분석의 대상이 되는 텍스트 유형(text type)은 일정한 언어적, 구조적 특징과 기능을 가지고 있는 텍스트 집단을 말하며, 이에 관한 논의도 관점의 차이에 따라 다양하게 전개되어 왔다. Werlich(1976)는 언어학적 특질에 기초하여 서술(descriptive), 이야기(narrative), 설명(expository), 논증(argumentative) 텍스트로 분류하였으며, Brewer(1980)는 텍스트 기저의 인지구조에 따라 서술, 이야기, 설명 텍스트로 구분하였고, Brinker(1985)는 서술, 이야기, 논증 텍스트로 나누었다.

이들의 구분에 따르면 학술적 논문은 대체로 논증 텍스트 혹은 설명 텍스트에 속하며, 이때의



〈그림 1〉 Superceanu의 과학논문 초구조(superstructure of scientific articles)

논증은 논리학에 기반한 필연적 관계가 아니라 일상적 혹은 학문적 논증에서의 가정과 결론 사이의 개연적 추론 관계로 이루어지는 것을 말한다. van Dijk(1980)는 논증 텍스트로부터 학술논문(scholarly paper)을 분리하여 서론(문제제기), 배경(제안이유, 분석이론, 서술, 실험), 결론 등으로 구성된 학술논문의 초구조 스키마(superstructure schema)를 제시하였으며, 이 구조는 이후 학술논문의 텍스트 구조 분석 연구의 기본 모형으로 많은 영향을 미쳤다. Superceanu(1998)는 van Dijk의 학술논문 초구조 스키마를 발전시켜 과학논문(scientific articles)의 초구조 모형(superstructure model)을 제시하였다(〈그림 1〉 참조).

2.2 연구 논문의 텍스트 구조 기반 메타데이터

연구 문제를 해결하기 위해 연구자들은 다양한 방법론을 사용해 데이터를 분석하고 결과를 해석하는 연구 과정을 거치며, 이 때 경우에 따라서는 가설을 수립하고 필요한 데이터를 수집해서 가공한다. 그리고 이러한 과정을 통해 생성되는 연구논문, 연구보고서, 학술회보(proceedings) 등은 일정한 형식의 문서구조와 서술 내용의 의미구조를 통해 지식을 구성하고 표현한다(〈표 1〉 참조).

연구 문헌에 내포된 내용 및 의미와 관련하여 많은 연구자들이 연구 문헌이 갖는 텍스트 구조의 중요성에 주목하였다. 관련된 초기의 연구는 대체로 텍스트 분석을 토대로 텍스트의 의미 구조를 도출하기 위한 연구와 의미 구조를 구성하는 개념요소를 기반으로 하는 지식조직체계를 개발하기 위한 연구로 나누어진다.

텍스트 구조에 관한 대표적인 연구로는 Kando(1997, 1999)와 Harmsze(2000)의 연구를 들 수 있다. Kando는 연구 문헌의 내용을 분석해 배경(background), 목적, 방법 등 연구 문헌의 전형적이고 기능적인 의미 요소를 정리한 세트

〈표 1〉 연구 문헌 텍스트의 문서구조와 의미구조(고영만, 송인석 2011)

구조	문서 구조 (Syntactic structure)		⇒	의미 구조 (Semantic structure)	
	형식 구조 (Formal structure)			개념 구조 (Conceptual structure)	논증 구조 (Argumentation structure)
구성	문서의 서술형식			서술내용의 의미	서술내용의 논증적 맥락
요소	제목, 장, 절, 표, 그림, 참고문헌 등			연구목적, 연구필요성, 연구방법, 연구내용, 연구결과, 결론 등	이론, 관점, 관찰수단, 측정도구, 데이터처리, 검증방법 등

를 'text-level 구조'로 정의해 제시하고, 해당 구조에 의해 연구 문헌을 검색하는 것이 검색의 정확성을 높일 수 있음을 강조하였으며, Harmsze는 정보통신 환경에서 이루어지는 학술 커뮤니케이션 과정에서 연구자들이 본인의 연구 자료를 보다 효율적으로 공유하고 검색할 수 있도록 하기 위해, 텍스트의 의미 관계를 기반으로 연구 문헌을 분석한 모형을 제시하였다.

지식조직체계 개발과 관련해서는 다양한 방식들이 제시되었다. Sibun(1993)은 검색 시스템을 만들 때는 대체로 텍스트 구조(text structure)에 근거하지만, 실제 연구자들은 연구 문헌과 관련된 주제(domain) 및 수사 구조 등 다양한 관점에서도 연관된 내용을 찾고자 하므로, 연구 주제와 관련된 전체적인 내용의 윤곽들을 보다 쉽게 이해하고 그것을 조절하는 도구를 개발하는 것이 필요하다고 하였다. Ono, Sumita, & Miike (1994)는 단어 빈도 또는 스크립트 기반의 전통적 초록 산출 시스템은 다양한 주제의 문서들에 일관된 적용이 어렵기 때문에, 연구 문헌의 의미 구조를 표현하는 언어적 단서(linguistic clues)에서 도출한 논리적 관계를 토대로 초록 산출 시스템을 제시하고, 해당 방식이 원자료 내에서 중요한 의미를 내포하는 문장을 최대 74%까지 보여줄 수 있음을 검증하였다. Bouayad-Agha, Power, & Scott(2000)는 문헌의 의미 구조에 대한 색인 체계가 의미 기반 검색에서 중요함을 강조하고, 연구 문헌을 재현하는 모든 최소 요소들을 추출해 다른 연구 문헌과 구조적 호환성을 가질 수 있는 지식조직체계의 구축이 필요하다고 하였으며, Beissel-Durrant(2004) 또한 사회과학 분야에서 연구 방법의 우선순위나, 최신 연구 파악, 앞으로 필요한 미래 연구 분야의 선별, 연구 프로젝트 분류와 같은 다양한 목적으로 향후 연구방법 유형학(research method typology)이 중요해질 것으로 판단하고, 이를 포괄할 수 있는 구조화된 시소러스와 같은 색인어 조직체계가 필요하다고 하였다.

새로운 정보기술의 발전과 웹 문서의 기술 표준이 진화됨에 따라 1990년대 중반 이후부터는 인터넷 환경에서 정보문제를 보다 효율적으로 처리하기 위한 의미 검색 기술에 관한 연구들이 진행되기 시작하였다. 메릴랜드 대학의 SHOE (Simple HTML Ontology Extensions) 프로젝트는 추가 정보 표현을 위한 개념을 온톨로지로 만들어 기존 HTML 언어로 작성된 웹 문서에 삽입함으로써, 해당 온톨로지를 통해 간단한 추론 규칙을 정의해 의미 검색을 수행할

수 있는 구조를 제안하였다(Luke, Spector, & Rager, 1996; Heflin, Hendler, & Luke, 1998; Kampa, 2002). 또한 2000년대 중반 이후에는 웹 기반의 학술커뮤니케이션을 지원하기 위해 개발된 지식 체계인 ScholOnto(Scholarly Ontology)와 인용색인관계를 웹 온톨로지 언어인 RDF 형식으로 기술하는 Cito(Citation Typing Ontology) 개발이 이루어졌다(Buckingham Shum, Motta, & Dominique, 2000; Shotton, 2009).

지금까지 수행된 연구들을 살펴보면 학술 연구 자료 텍스트의 의미 구조에 대한 분석, 그리고 해당 연구 자료에 내포된 의미를 보다 효율적으로 검색하기 위한 지식조직체계와 이를 정보기술에 응용하기 위한 연구들이 주로 진행되어 왔음을 확인할 수 있다. 하지만 연구자의 논증 구조를 반영하는 연구 논문 텍스트의 의미 구조에 기반한 메타데이터를 토대로 검색 시스템을 구축해 의미 검색의 가능성을 검증한 연구는 이루어진 바가 없다. 따라서 본 연구는 지금까지 이루어진 선행연구에서 개발된 한국학 연구 논문 텍스트의 의미 구조 메타데이터 항목을 적용하여 실제 논문 검색을 위한 시범시스템을 구축하고, 기존 검색 시스템과 동일한 검색어를 적용해 실시한 검색 결과를 비교해 봄으로써 그 차이를 분석해 보고자 한다.

3. 연구 논문의 의미 구조 탐색 시스템 구축

3.1 데이터베이스 구축 대상 자료의 범위

본 연구에서는 한국학 범위에 속하는 일정한 수준 이상의 논문이 수록된 학술지를 대상으로 프로토타입 데이터베이스를 구축하여 실험을 하였다. 한국학의 학문적 정의는 학자마다 상이하며, 넓은 의미에서는 '한국'을 주제로 하는 인문, 사회, 자연과학 분야 모두를 가리키기도 하나, 대체로 인문학과 사회과학 영역의 학문으로 받아들여지는 경우가 많다. 따라서 본 연구에서는 인문학과 사회과학 분야의 2011-2013년도 한국학술지인용색인(Korea Citation Index: 이하 KCI) 영향력지수(Impact Factor, 이하 IF) 평균값인 0.5와 0.84 이상의 학술지를 데이터베이스 구축 대상으로 선정하고자 하였다. 그러나 '한국학'과 관련된 검색어를 통해 검색한 대부분의 한국학 분야 학술지가 대분류 기준 인문학 분야에 속해있는 것으로 나타남에 따라 인문학 분야의 IF값 0.5 이상인 학술지를 대상으로 데이터베이스를 구축하였다. 다만 KCI에는 2004년 이후부터 논문 데이터가 본격적으로 축적되기 시작하였으며 2015년 하반기를 기준으로 2014년도 발행 논문이 KCI에 탑재되기 때문에, 2004년부터 2014년까지의 논문을 수록하고 있는 학술지를 모두 데이터베이스 구축 대상으로 포함시켰다.

이러한 기준에 따라 KCI에서 검색어 '한국학', 'Korean Studies' 및 한국학 분야 학술지명에 '한국'의 이칭으로 많이 사용되는 '국학', '대동', '태동', '극동', '조선', '동학', '동방' 키워드로 검색된 29종의 한국학 관련 학술지 중 IF값이 0.5 이상이며, 2004년부터 2014년까지의 11년치 누적 논문 데이터를 갖고 있는 학술지 14종에 수록된 논문 중 2009년, 2014년 발간 논문 894편이 데이터베이스로 구축되었다(〈표 2〉 참조). 2009년, 2014년의 5년 단위 발간 논문을

〈표 2〉 데이터베이스 구축 대상 학술지 및 논문 수

no.	학술지명	발행기관	논문 수
1	대동한문학	대동한문학회	51
2	대동문화연구	성균관대학교 대동문화연구원	120
3	한국학논총	국민대학교 한국학연구소	43
4	정신문화연구	한국학중앙연구원	95
5	한국문화	서울대학교 규장각한국학연구원	82
6	동방학지	연세대학교 국학연구원	69
7	조선시대사학보	조선시대사학회	56
8	고조선단군학	고조선단군학회	29
9	한국과 국제정치	경남대학교 극동문제연구소	59
10	동학학보	동학학회	46
11	동방한문학	동방한문학회	78
12	북한연구학회보	북한연구학회	24
13	민족문화연구	고려대학교 민족문화연구원	67
14	한국민족문화	부산대학교 한국민족문화연구소	75
합계			894

데이터베이스 구축 대상으로 삼은 이유는 논문의 발간 시기에 따라 한국학의 연구 분야나 특성 등에 차이가 있을 수도 있다는 점을 고려한 것이다. 다만 2015년 6월 말 기준 시점에 구축 대상 데이터 목록을 정리해 입력 작업을 수행하였기 때문에 KCI에서의 논문 정보 등록이 늦어진 일부 연구 논문의 경우에는 누락이 있을 수도 있다.

3.2 메타데이터 항목의 구성

한국학 연구 논문의 의미적 탐색 시스템을 구현하기 위한 메타데이터 요소의 개발은 유사라, 이혜원, 송인석(2009), 고영만과 송인석(2011), 송민선과 고영만(2015) 등의 연구를 통해 이루어진 바 있다. 가장 최근 이루어진 송민선과 고영만(2015)의 연구는 연구 문헌의 텍스트 구조를 분석해 정리한 Kando(1999)와 유사라, 이혜원, 송인석(2009), 고영만과 송인석(2011)의 연구를 토대로 추가적으로 필요한 의미 구조 항목을 포함시켜 16개의 메타데이터 스키마를 제시하였다.

본 연구에서는 송민선과 고영만(2015)이 제안한 16가지 텍스트 구조 메타데이터 항목 중 일부를 조정하여 15개 항목을 적용하고, KCI를 비롯해 기존의 일반적인 학술정보시스템에서 제공되는 형식 구조 메타데이터 35개 항목을 함께 반영해 데이터베이스를 구축하였다. 의미 구조 메타데이터 항목의 수를 조정한 이유는, 실제 데이터베이스 구축 과정에서 '연구 문제'와 '가설'로 구분한 세부 항목의 내용이 한국학 관련 논문 내에서 명확히 분리되어 나타나는 경우가 많지 않아, 해당 항목을 '연구문제/가설'로 합쳐서 구축하는 것이 효율적이었기 때문이다(〈표 3〉 참조).

데이터베이스 구축 작업은 문헌정보학 분야 석사 과정 이상의 연구 인력들이 구축 대상 논

〈표 3〉 한국학 연구 논문의 의미 구조 검색을 위한 메타데이터

구분	메타데이터 항목	메타데이터 설명
의미 구조 메타데이터	Research Objectives	연구 목적
	Research Significance	연구의 의의
	Research Limitation	연구의 제한점
	Research Type	연구 유형
	Research Problem / Hypotheses*	**연구 문제 / 가설***
	Research Method	연구 방법
	Data	연구대상 데이터
	Temporal Background	연구대상의 배경이 되는 시간적 범위
	Spatial Background	연구대상의 배경이 되는 공간적 범위
	Related Models	연구와 관련되거나 차용한 이론 및 모형 등
	Method of Data Collection	데이터 수집 방법
	Data Size	데이터 수집 규모
	Analysis Method	연구 데이터 분석 방법
	Research Outcome	연구 결과
	Further Research	후속 연구 제언
형식 구조 메타데이터	Authors_Kr, Authors_Eng, Authors_Chi, ArticleTitle_Kr, ArticleTitle_Eng, Article Title_etc, JournalTitle_Kr, JournalTitle_Eng, JournalTitle_Chi, Language, Author Keywords_Kr, Author Keywords_Eng_etc, Abstract_Kr, Abstract_Eng, AuthorAddresses, AuthorID_NRF, CitedCount, Publisher, Publisher_Eng, ISSN, Publication_Month, Year_Published, Volume, Issue(No), Beginning_Page, Ending_Page, All_Pages, Subject(NRF), IF(2011), IF(2012), UCI, Article_ID, KCI_URL, KCI_URL_Link, etc_URL_Link	KCI 및 일반적인 논문 검색 서비스에서 제공되는 형식 구조 메타데이터 항목 35개

* 통합 조정한 메타데이터 항목

문을 직접 읽고 해당 항목을 확인해 입력하는 수작업 방식으로 이루어졌다. 다만 형식 구조 메타데이터의 대다수는 KCI에 입력된 내용을 엑셀 형태로 직접 다운로드 받을 수 있기 때문에 해당 데이터를 그대로 사용하였으며, KCI에서 누락되거나 잘못 입력된 내용은 보정하였다. 그리고 연구 논문 텍스트의 의미 구조 메타데이터 항목은 데이터 입력 과정에서의 내용 왜곡을 방지하기 위해 가급적 연구자가 논문에 기술한 문장을 변형 없이 그대로 기재하는 것을 원칙으로 삼았으며, 일부 연구 방법이나 연구 유형과 같이 논문에서 명시적으로 나타나지 않는 내용들은 통일성 유지를 위해 〈표 4〉의 입력 지침에 따라 입력 작업을 수행하였다.

〈표 4〉 한국학 연구 논문의 의미 구조 메타데이터 항목 입력 지침

메타데이터 항목	메타데이터 설명	입력 내용 및 방법
Research Objectives	연구 목적	논문에 제시된 연구 목적으로 가급적 논문에 사용된 문장을 변형 없이 그대로 기재
Research Significance	연구의 의의	논문에 제시된 연구의 의의로 가급적 논문에 사용된 문장을 변형 없이 그대로 기재
Research Limitation	연구의 제한점	논문에 제시된 제한점으로 가급적 논문에 사용된 문장을 변형 없이 그대로 기재
Research Types	연구 유형	연구 유형 구분(질적연구/양적연구/혼합연구 등)
Research Problem/ Hypotheses	연구 문제/가설	논문에 제시된 연구 문제로 가급적 논문에 사용된 문장을 변형 없이 그대로 기재
Research Methods	연구 방법	논문에 연구자가 구체적으로 제시한 방법을 기술
Data	연구대상 데이터	논문에 연구자가 연구를 위해 수집하고 분석한 데이터(문헌명/인명/사회현상/정책자료 등)
Temporal Background	연구대상의 배경이 되는 시간적 범위	논문의 연구 대상과 관련된 시간적 배경 범위(조선시대/1970년대/일제시대 등)
Spatial Background	연구대상의 배경이 되는 공간적 범위	논문의 연구 대상과 관련된 공간적 배경 범위(경북 지역/안동 지역/중국 등)
Related Models	연구와 관련되거나 차용한 이론 및 모형 등	해당 연구를 위해 다른 연구에서 차용하거나 관련된 이론 및 모형
Methods of Data Collection	데이터 수집 방법	연구 데이터를 수집한 방법(설문조사/면접/관찰/기초자료 활용 등)
Data Size	데이터 수집 규모	수집한 데이터의 규모(설문 대상자 수, 분석 문헌의 건수 등)
Analysis Methods	연구 데이터 분석 방법	통계 방법이나 연구 데이터의 분석 방법(ANOVA/구술기록 분석 등)
Research Outcome	연구 결과	논문에 제시된 연구결과로 가급적 논문에 사용된 문장을 변형 없이 그대로 기재
Further Research	후속 연구 제언	논문에 제시된 후속 연구에 관한 내용으로 가급적 논문에 사용된 문장을 변형 없이 그대로 기재

3.3 시스템 구축

3.3.1 시스템 구성

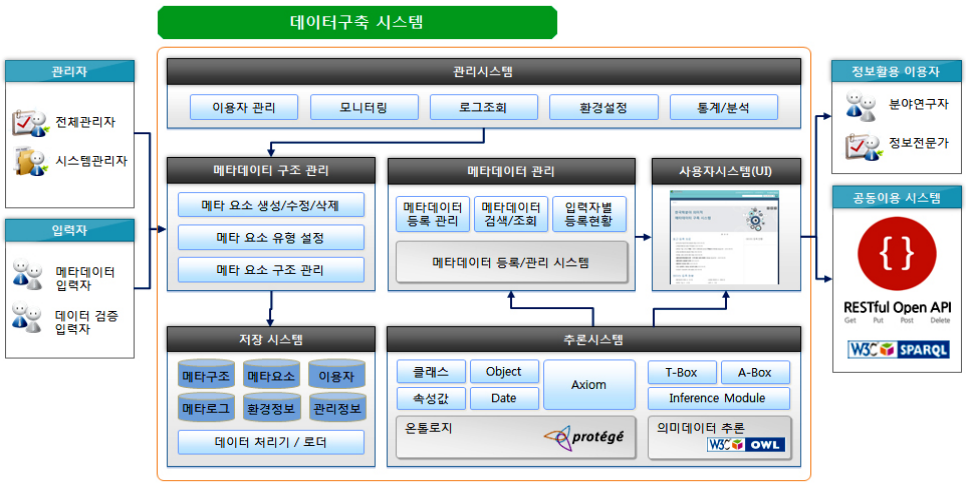
본 연구에서 구축한 한국학 메타데이터 데이터베이스(Korean Studies Metadata Database: 이하 KMD, http://14.63.171.94:8090/)는 Apache 프로젝트의 HTTP 서버를 기반으로, WAS(Web Application Server) 운영 환경과 Apache 프로젝트의 Tomcat을 갖추어 웹 서비스를 위한 기본환경을 조성하였다. KMD의 DB 개발환경은 기능 확장 및 수정이 용이하고 MVC(Model View Controller) 패턴을 지원하는 Struts2를 기반으로 하였으며, 한국연구재단 및 기초학문자료센터에서 사용하고 있는 관계형 DBMS Oracle을 사용하여 데이터베이스를 관리하였다(〈표 5〉 참조).

KMD 시스템은 통합관리시스템, 메타데이터 구조 관리 시스템, 사용자 시스템, 저장 시스템, 추론 시스템으로 구성된다(〈그림 2〉, 〈그림 3〉 참조).

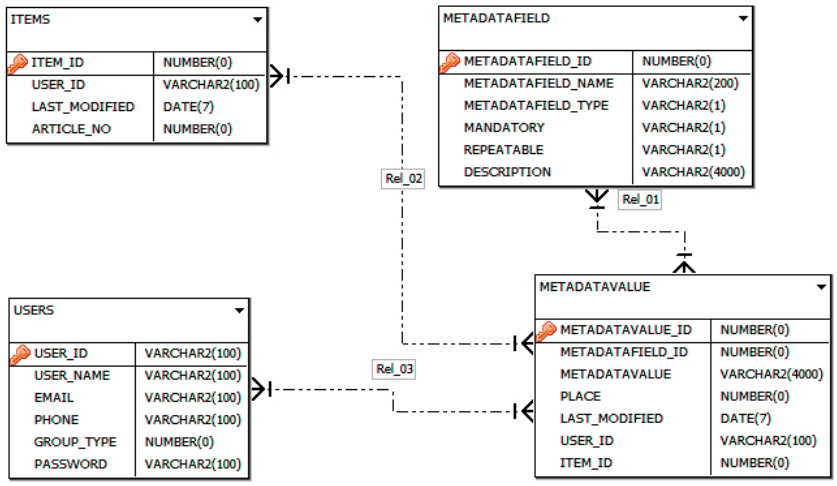
통합관리시스템과 메타데이터 구조 관리 시스템은 전체 관리자만 이용할 수 있는 모듈이다. 통합관리시스템은 이용자 관리, 데이터 등록 현황 모니터링, 각종 로그 정보 조회, 시스템 환경설정, 통계 조회 및 데이터 분석 관리 기능을 제공하며, 메타데이터 구조 관리 시스템은 KMD 시스템에서 사용하는 메타데이터 요소들을 생성/수정/삭제 할 수 있고, 각 메타데이터 유형별 필수/반복 여부와 입력 형태 등을 편집하고 관

〈표 5〉 시스템 구축 환경

WEB/WAS Server	Database
• 1대의 서버에 WEB/WAS 서버를 동시에 설치하여 사용 • CPU: Intel 2vCore • Memory: 4GByte • Disk: 100GByte • OS: centos64-64-140930 • WEB: Apache HTTPD • WAS: Apache Tomcat	• WEB/WAS 서버와 동일한 서버 사용 • CPU: Intel 2vCore • Memory: 4GByte • Disk: 100GByte • OS: centos64-64-140930 • DBMS: Oracle 11G



〈그림 2〉 시스템 구성도



〈그림 3〉 KMD DB의 핵심 테이블 ERD(Entity-Relationship Diagram)

리할 수 있다.

사용자 시스템은 시스템 관리자가 주로 사용하는 모듈이며, UI(User Interface) 생성 모듈로 관리자 및 데이터 입력자의 UI를 생성할 수 있다. 저장 시스템에는 메타데이터 및 입력 데이터 정보, 이용자 정보, 로그 정보, 환경 설정 및 관리 정보가 저장되고, 추론 시스템은 향후 KMD 시스템에 등록된 데이터를 기반으로 추론 서비스를 제공하기 위한 공리와 추론엔진으로 구성된다.

3.3.2 인터페이스 구성

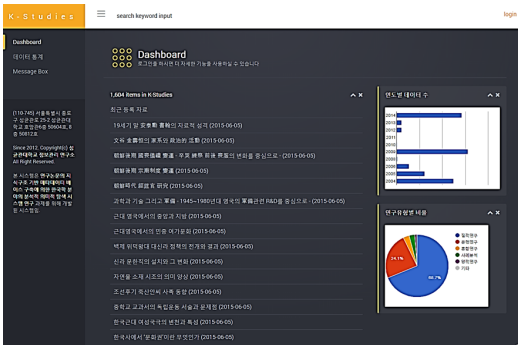
KMD 시범시스템의 기능별 인터페이스는 크게 메인 페이지, 메타데이터 구조 관리, 입력 데이터 조회, 메타데이터 항목별 데이터 입력 및 관리의 4개 페이지로 구성된다(<그림 4> 참조).

1) 메인 페이지

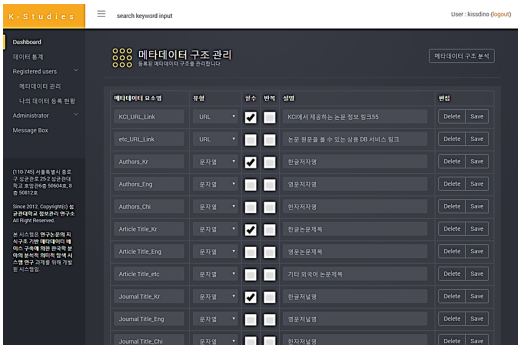
KMD 시범시스템 접근 시 출력되는 화면으로 최근 등록된 자료 목록과 함께 등록된 데이터의 현황을 시각화 한 도표를 보여주며, 시스템 로그인 시 데이터 검색 기능 및 주요 메뉴에 접근할 수 있는 기능을 제공한다.

2) 메타데이터 구조 관리 페이지

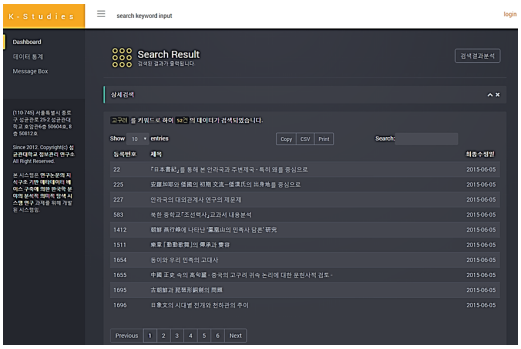
현재 KMD 시스템에 등록되어 있는 메타데



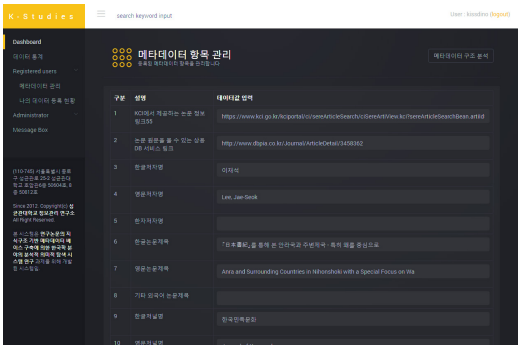
[메인페이지 화면]



[메타데이터 구조 관리 화면]



[입력데이터 조회 화면]



[메타데이터 항목별 데이터의 입력과 관리 화면]

<그림 4> KMD 시스템 인터페이스

이터의 구조를 관리할 수 있는 화면으로 신규 메타데이터 요소를 등록하고, 기존에 등록되어 있는 메타데이터 요소를 수정하고 삭제할 수 있으며, 각 메타데이터 요소의 입력 형태(숫자, 문자, 객체, URL 등)와 해당 항목의 필수/반복 등록 여부 등의 설정 기능을 제공한다.

3) 입력 데이터 조회 페이지

등록된 데이터를 조회할 수 있는 화면으로 기존의 다른 논문 검색 서비스들에서 제공하고 있는 것과 동일한 메타데이터 항목(제목, 저자, 출판사, 키워드, 초록 등) 외에 본 연구에서 확장 설계한 텍스트의 의미구조 기반 메타데이터 항목(연구의 의의, 연구의 제한점, 연구방법, 연구문제/가설, 연구대상 데이터 등)을 포함한 키워드 검색 기능을 제공한다.

4) 메타데이터 항목별 데이터의 입력과 관리 페이지

전체 관리자가 메타데이터 구조 관리 화면에서 설정한 메타데이터 요소를 토대로 논문 1편에 대한 각 항목별 데이터를 등록하고 수정할 수 있는 화면으로 필수/반복 여부에 따라 입력 구조가 제공되고, 잘못 입력된 데이터에 대해서는 구조적 타당성을 판별해 에러 메시지를 출력하는 기능을 제공한다.

4. 한국학 연구 논문의 의미 구조 검색

본 연구에서는 동일한 키워드를 적용하여 KMD 시스템과 KCI 시스템에서의 검색 결과를 비교하였다. 이때 KCI의 경우 KMD와 같은 메타데이터 항목 값을 따로 가지고 있지 않기 때문에 검색 대상 항목을 키워드로 하였다. 또한 KMD에 탑재된 논문과 동일한 논문의 검색 결과를 KCI로부터 얻기 위해 학술지와 발행년도를 2009년과 2014년 발행 KMD 구축 대상 학술지로 제한하여 검색 결과를 출력하였다.

두 시스템에서 나타나는 검색 결과의 차이를 분석할 수 있도록 동일한 검색어를 사용하였다. 본 연구의 주요 목적이 텍스트 구조 기반 메타데이터 시스템의 텍스트의 의미 구조 검색 기능을 확인하는 데 있으므로, 두 시스템으로부터 비교 가능한 검색 결과를 얻는 것이 중요하며 검색 결과를 기대할 수 없는 임의의 검색어를 입력하는 것은 의미가 없다. 따라서 검색어 선정을 위해 먼저 KMD 데이터베이스 구축 대상 논문으로부터 추출한 저자키워드를 정제해 각 키워드의 출현빈도를 정리하였다. 키워드 출현빈도의 경우 멱함수 형태로 나타나기 때문에 출현빈도가 낮아질수록 키워드 수가 급격하게 늘어난다. 이에 비해 동시에 출현하는 키워드쌍의 수는 상대적으로 적어지는 경향을 띠어 일정 수의 검색 결과값을 얻기 어려워진다. 따라서 검색 결과값의 비교가 용이한 출현빈도 상위 순위의 키워드 쌍을 적용하기 위해 10회 이상 출현한 키워드 중 3순위에 해당하는 '동학', '시천주', '정조'까지를 예비 검색어로 선정하였다. 그리고 예비 검색어와 동시에 출현한 키워드쌍을 정리하여 그 중 2회 이상 동시 출현한 쌍 중에서 '동학'과 '시천주', '정조'와 '심환지'를 검색어로 최종 선정하였다(〈표 6〉 참조). KMD 시스템과 KCI 시스템 모두 'AND' 연산을 적용하여 검색어 쌍에 대한 검색을 실시하였다.

〈표 6〉 데이터베이스 구축 논문의 출현빈도 및 동시 출현 키워드쌍

출현빈도 상위 키워드			동시출현 키워드	
순위	키워드	출현횟수	키워드	동시출현 횟수
1	**동학***	18	**시천주***	9
			전봉준	3
			무극대도, 성경신, 영성, 종교체험, 지구생명공동체, 최제우	2
			공공성, 한울님.....	1
2	**시천주***	13	**동학***	9
			무극대도, 수심정기, 양천주, 지구생명공동체, 한울님, 후천개벽	2
			성경신, 영성....	1
3	**정조***	10	**심환지***	2
			노론벽파, 건륭제, 남인....	1
4	고조선	9	단군	3
			고인돌, 기자조선, 신시, 옥기문화, 홍산문화, 북한	2
			지명, 서술, 환웅.....	1
5	북한	8	시장화	3
			고조선	2
			선군사상, 숙청, 장성택.....	1
6	정체성	8	공간성, 오리엔탈리즘.....	1
. .	. .	. .	. .	. .

* 검색에 사용한 키워드 쌍

4.1 검색 결과 비교

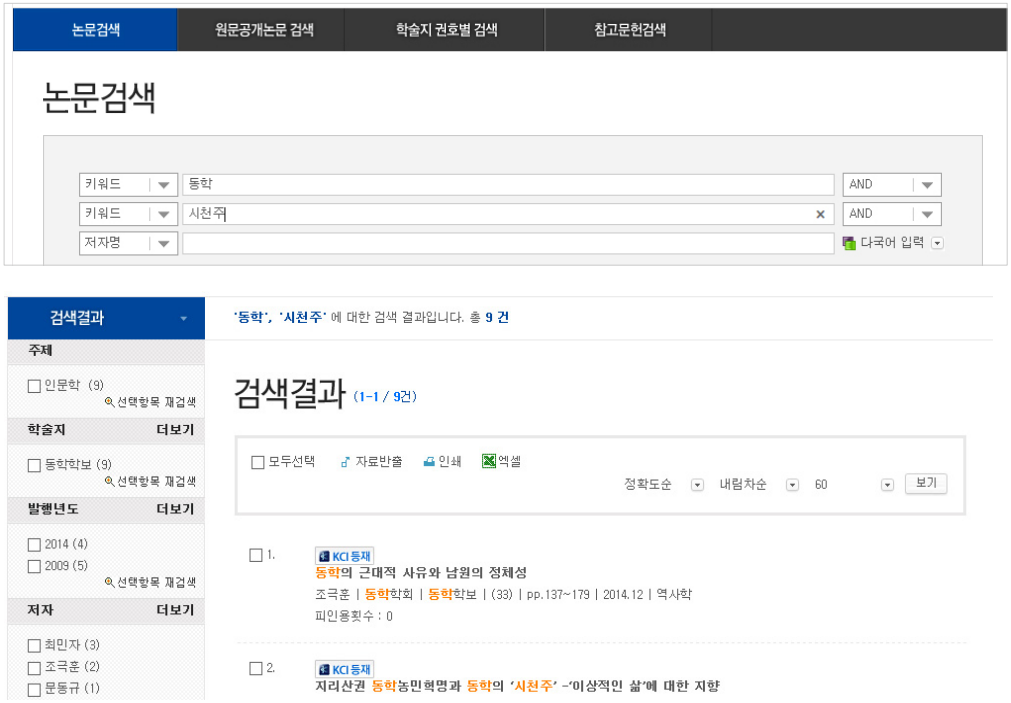
1) '동학 AND 시천주'에 대한 KMD 및 KCI 검색 결과

'동학'과 '시천주'를 검색어로 하여 KMD와 KCI를 검색한 결과는 두 시스템 모두 9건으로 나타났다. KCI 검색 결과 화면은, 시스템 간의 인터페이스의 차이는 있으나 보통 다른 논문 검색 시스템에서도 볼 수 있는 논문제목, 저자, 논문 출처 등에 따라 정렬되어 나타난다(〈그림 5〉 참조).

KMD 역시 검색결과에 대해 별도의 조건을 두지 않았기 때문에 〈그림 6〉과 같이 논문 제목에 따라 정렬되었다. 하지만 KMD의 경우 의미구조 메타데이터 항목을 기준으로 검색 결과값을 선택해 볼 수 있기 때문에 '동학 AND 시천주' 키워드를 포함하는 논문들의 연구 목적, 연구 대상 데이터, 공간적 배경 등을 결과값으로 따로 선택하여 정렬해서 볼 수 있다(〈그림 7〉 참조).

2) '정조 AND 심환지'에 대한 KMD 및 KCI 검색 결과

'정조'와 '심환지'를 검색어로 하여 KMD와 KCI를 검색한 결과는 두 시스템 모두 2건으로 나타났다. '동학'과 '시천주'의 검색 결과와 마찬가지로 AND 조건 검색에 의한 KCI의 검색



〈그림 5〉 '동학 AND 시천주'에 대한 KCI 검색 결과

〈그림 6〉 '동학 AND 시천주'에 대한 KMD 검색 결과

[연구 목적]

동학,시천주 를 키워드로 하여 9건 의 데이터가 검색되었습니다.

Show 10 entries Copy CSV Print Search:

등록번호	연구목적
1038	지리산권 동학농민혁명과 동학의 '시천주' 사상이 '이상적인 삶'에 대한 지향임을 밝힌다.
1048	동학 사상의 깊은 곳에 담겨있는 후현대의 경향성을 살펴본다.
1098	기독교와 천도교의 핵심개념인 '하나님 나라'와 '시천주'를 수양론적 시각에서 탐구한다.

[연구 대상 데이터]

동학,시천주 를 키워드로 하여 9건 의 데이터가 검색되었습니다.

Show 10 entries Copy CSV Print Search:

등록번호	연구대상데이터
1038	지리산권 동학농민혁명
1048	동학의 후현대 경향성
1098	하나님 나라, 시천주

[연구의 공간적 배경]

동학,시천주 를 키워드로 하여 4건 의 데이터가 검색되었습니다.

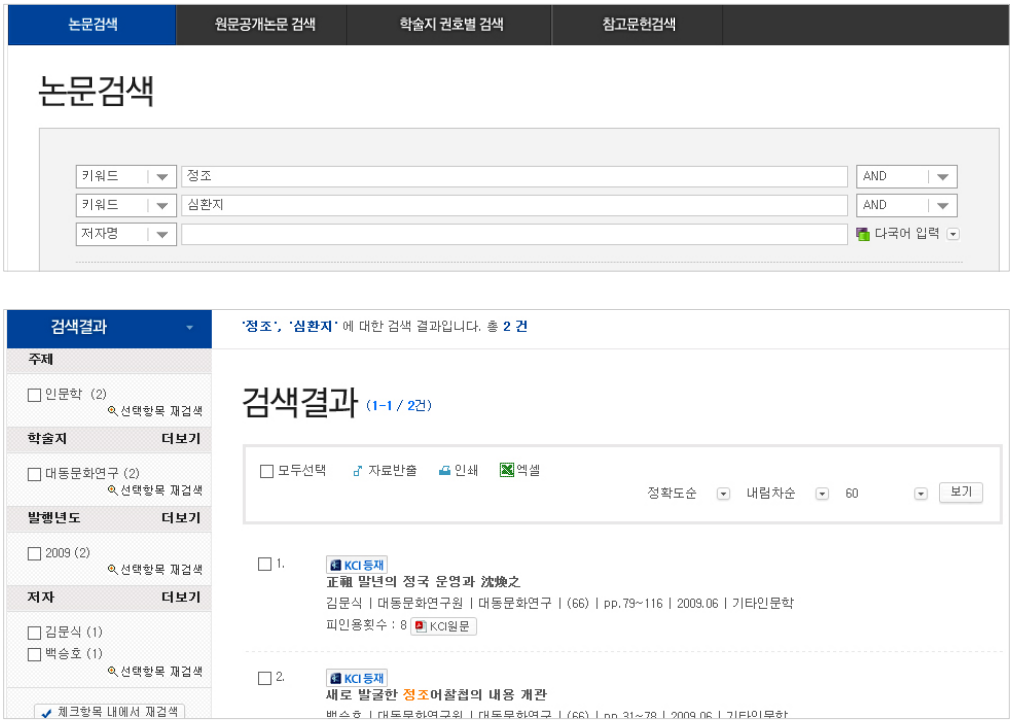
Show 10 entries Copy CSV Print Search:

등록번호	연구대상의 배경이 되는 공간적 범위
1038	지리산 권역
1194	러시아
1194	조선

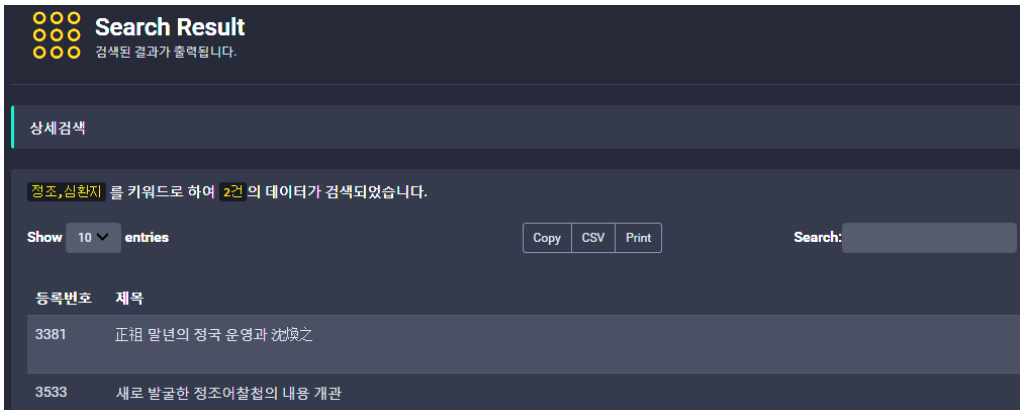
〈그림 7〉 '동학'과 '시천주'를 키워드로 갖는 논문의 연구목적, 연구대상 데이터, 공간적 배경에 대한 KMD 검색 결과

결과 화면은 논문제목을 중심으로 정렬되어 있어 논문 제목이나 저자명, 출처 외의 다른 내용들을 확인하기 위해서는 상세 정보 화면으로 이동해 초록이나 원문 파일을 재확인해야 한다(〈그림 8〉 참조).

KMD의 경우 '동학'과 '시천주'의 검색 사례와 마찬가지로 검색결과에 대해 별도의 조건을 두지 않으면 〈그림 9〉와 같이 논문 제목에 의해 정렬되며, 키워드를 포함하는 논문들의 의미구조 메타데이터 항목 결과 값을 따로 선택하여 정렬해서 볼 수 있다. 해당 조건에 대한 시간적 배경과 연구 결과, 후속 연구 제언에 대한 값을 출력하면 〈그림 10〉과 같다.

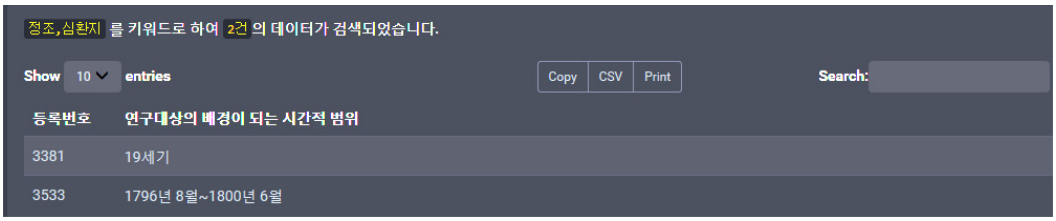


〈그림 8〉 '정조 AND 심환지'에 대한 KCI 검색 결과

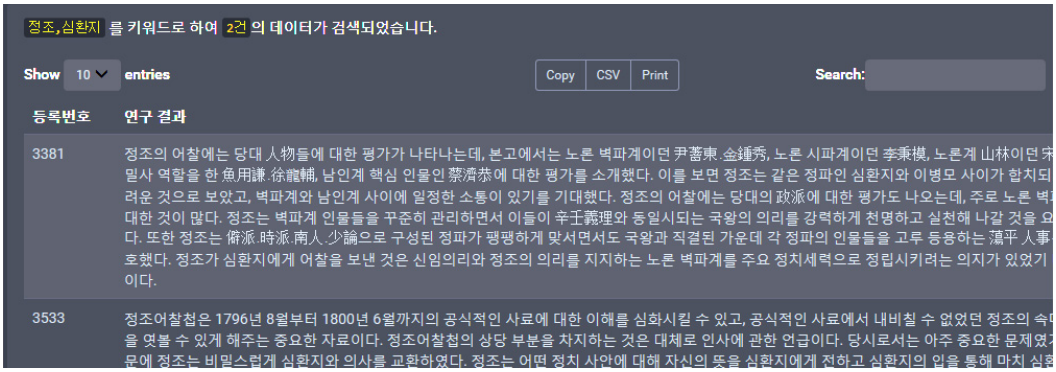


〈그림 9〉 '정조 AND 심환지'에 대한 KMD 검색 결과

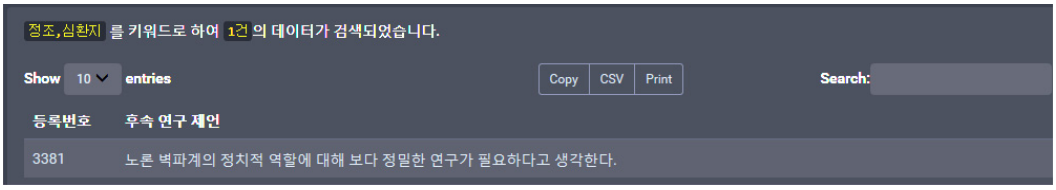
[시간적 배경]



[연구 결과]



[후속 연구 제언]



〈그림 10〉 '정조'와 '심환지'를 키워드로 갖는 논문들의 시간적 배경, 연구 결과, 후속 연구 제언에 대한 KMD 검색 결과

4.2 검색 결과 비교 평가

KCI와 같은 기존의 논문 검색 서비스를 이용할 경우에는 이용자가 의도한 검색 키워드와 연관된 '연구 대상 데이터'나 '연구 결과' 등의 내용을 구체적으로 확인하기 위해 검색 결과로 출력된 각 논문의 초록이나 원문을 다시 한 번 훑어봐야 하는 번거로움이 생긴다. 이에 비해 텍스트 구조 기반 메타데이터 시스템을 구축할 경우 관련된 메타데이터 항목들이 있기 때문에 연구 논문의 의미구조에 해당하는 내용을 바로 확인할 수 있는 것으로 나타났다.

연구 논문에 수록된 다양한 연구 내용들을 검색 결과 화면에서 바로 보여줄 수 있게 될 경우, 이용자 인터페이스(UI)의 구성 방식이나 적용하는 시각화 도구에 따라 이용자들이 검색 결과를 보다 직관적으로 이해할 수 있도록 제공할 수 있는 활용 가능성 또한 훨씬 높다. 현재

데이터 통계를 그래프 형태로 시각화하여 이용자들이 쉽게 인지할 수 있도록 제공하고 있는 KMD 시범시스템의 메인 페이지(〈그림 4〉 메인페이지 화면 인터페이스 참조)를 발전시켜 향후 보다 다양한 시각화 도구들을 검색 결과 통계에 적용해 제공한다면, 이용자들이 본인이 검색한 내용과 관련해 어떤 방식의 연구가 많이 이루어졌는지(연구 유형), 연구에 적용된 방법론은 무엇인지(연구 방법), 연구 대상 데이터는 어떤 것이었는지(연구 대상 데이터), 어느 시기의 어떤 지역을 배경으로 한 연구들이 주로 이루어졌는지(연구의 시간적・공간적 배경) 등에 대해 보다 직관적으로 확인할 수 있게 될 것이다.

또한 본 연구의 목적에서 벗어나 연구 논문의 의미 구조 내용을 반영하는 KMD의 메타데이터 검색 기능 사례를 모두 서술하지는 않았으나, 텍스트의 의미 구조 메타데이터 항목을 기반으로 구축된 KMD 시스템은 키워드 검색 결과값의 필터링 뿐 아니라, 검색을 실시하는 조건 역시 이용자가 원하는 의도에 따라 연구 대상 데이터(예: '삼국유사'에 대해 연구한 논문), 연구의 시대적 배경(예: '18세기'와 관련한 논문)이나 공간적 배경(예: '안동' 지역과 관련한 논문) 등으로 구체적으로 제한해 수행할 수도 있다.

5. 결론 및 제언

본 연구에서는 한국학 분야 연구 논문을 대상으로 논문에 서술된 텍스트의 의미 구조를 파악할 수 있는 메타데이터 요소를 추출하여, 이를 토대로 시범적 검색 시스템을 구축하였다. 그리고 연구 논문 텍스트의 의미 구조를 기반으로 구축한 메타데이터 시스템이 기존의 논문 검색 시스템과 비교했을 때 어떤 특징과 차이점을 갖는지 확인하고, 연구자가 연구의 진행과정에서 필요로 하는 연구 논문의 텍스트 구조에 대한 의미 검색 기능 구현 가능성을 대해 검토하였다.

연구 결과, 텍스트 구조를 기반으로 하는 메타데이터 항목을 적용해 구축한 KMD 시스템이 기존의 일반적인 논문 검색 시스템에 비해 연구 논문 자료를 이용하고자 하는 이용자들의 정보 의도에 맞는 검색 결과를 보다 효율적으로 보여줄 수 있는 것으로 확인되었다. 즉 검색하고자 하는 키워드의 조합이나 조건식이 동일하더라도 검색 결과를 통해 최종적으로 연구 진행과 관련해 어떤 내용을 찾고자 하는 것인지, 연구 목적, 연구 대상 데이터, 연구의 시공간적 배경 등에 따라 검색 결과를 다양하게 보여줄 수 있었다.

본 연구는 연구 논문 텍스트의 의미 구조를 기반으로 하는 메타데이터 항목을 적용한 시범적인 검색 시스템을 구축하고, 기존 유사 시스템과의 비교를 통해 검색 기능과 효과를 확인해 보기 위한 목적으로 진행되었기 때문에 실험 대상으로 삼은 테스트 데이터의 수가 적다. 또한 검색 결과 값을 보여주는 인터페이스의 구성 방식이나 적절한 시각화 도구에 대해서는 논외로 했기 때문에 해당 내용과 관련된 시각적 효율성에 대해서는 평가하지 않았다는 제한점을 갖는다. 그렇지만 한국학 분야의 연구자들이 관심을 갖는 주제와 관련된 연구성과물의 내용을 보다 효율적으로 탐색할 수 있도록 텍스트의 의미 구조 기반 탐색 시스템의 프로토

타입을 구축하고, 텍스트의 의미 구조 검색 가능성을 직접 확인해 보았다는 점에서 그 의의를 찾을 수 있다. 따라서 본 연구를 발전시켜 각각의 의미 구조 메타데이터 요소 간에 성립될 수 있는 논리적 의미 관계의 특성들을 정리하고, 이를 통해 관련 자료의 의미 검색을 실현하기 위한 온톨로지(ontology), 공리(axiom), 추론 규칙(inference rules) 등을 도출해 활용하게 된다면, 앞으로 한국학 분야 연구 성과물의 의미 검색과 관련된 연구와 응용 영역에 많은 영향을 끼칠 수 있을 것이다.

본 연구에서는 직접 논문을 읽고 해당 항목들에 대한 내용을 입력하는 수작업 방식으로 데이터베이스 구축 작업을 진행하였으나, 향후 보다 경쟁력 있는 시스템 구축을 위해서는 각 메타데이터 항목에 입력되는 값들을 분석해 논문에 삽입된 문장의 수사적 기법이나 맥락의 구성 등을 파악해 의미 구조 메타데이터와 매칭하여 시스템에서 메타데이터의 자동 식별 처리를 할 수 있어야 한다. 예를 들면 연구 목적의 기술에 있어서 '~의 목적은', '~의 목적에서 작성되었다', '~의 목적에서 이루어졌다', '~을 살펴보고자 한다' 등의 어구가 많이 사용되는 경향을 분석하여 자동적으로 처리하는 기능이 해당될 수 있을 것이다. 또한 메타데이터의 식별 처리에 관한 연구와 더불어 본 연구 결과를 발전시켜 더 많은 관련 데이터를 지속적으로 구축하고, 실질적으로 연구자들이 연구 진행과 관련해 검색에 활용하는 검색어의 내용이나 표현 방식 및 주요 메타데이터 요소들을 보다 면밀히 분석해 다듬어 나가는 후속 연구들이 지속적으로 진행될 필요가 있다.

참 고 문 헌

고영만, 송인석 (2011). 연구 문헌의 지식구조를 반영하는 의미기반의 지식조직체계에 관한 연구. 정보관리학회지, 28(1), 145-170. http://dx.doi.org/10.3743/kosim.2011.28.1.145

박진용 (1997). 텍스트 의미 구조의 과정 중심 분석 방법 연구. 석사학위논문, 한국교원대학교 대학원, 국어교육학과 국어교육전공.

송민선, 고영만 (2015). 한국학 연구 논문의 의미구조 기반 메타데이터 연구. 한국도서관·정보학회지, 46(3), 277-299. http://dx.doi.org/10.16981/kliss.46.3.201509.277

유사라, 이혜원, 송인석 (2009). 연구자 중심 연구성과물 의미검색을 위한 인문사회 학술용어 온톨로지 적용 및 유지관리 체계 연구. 한국문헌정보학회지, 43(2), 277-298. http://dx.doi.org/10.4275/kslis.2009.43.2.277

정여훈 (2013). 수사구조이론과 한국어 텍스트 분석의 실제. 언어사실과 관점, 32, 261-288.

Beissel-Durrant, G. (2004). A typology of research methods within the social sciences. NCRM (National Centre for Research Methods) working paper, NCRM.

Bouayad-Agha, N., Power, R., & Scott, D. (2000). Can text structure be incompatible with

rhetorical structure? Proceedings of the First International Conference on Natural Language Generation. 12-16 Jun 2000, Mitzpe Ramon, Israel. http://dx.doi.org/10.3115/1118253.1118280

Brewer, W. F. (1980). Literary theory, rhetoric, and stylistics: Implications for psychology. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education. Hillsdale, NJ: Lawrence Erlbaum Associates.

Brinker, K. (1985). Linguistische textanalyse. Berlin: Erich Schmitdt Verlag.

Buckingham Shum, S., Motta, S. E., & Dominique, J. (2000). Scholonto: Ontology-based digital library server for research document and discourse. International Journal on Digital Libraries, 3, 237-248. http://dx.doi.org/10.1007/s007990000034

de Beaugrande, R., & Dressler, W. U. (1981). Einführung in die textlinguistik. Tübingen: Niemeyer.

Frederiksen, C. H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 7(3), 371-458. http://dx.doi.org/10.1016/0010-0285(75)90016-x

Halladay, M. A. K., & Hasan, R. (1976). Cohension in English. London: Longman.

Harmsze, F. A. P. (2000). A modular structure for scientific articles in an electronic environment. Ph.D. Dissertation, Van der Waals-Zeeman Institute, University of Amsterdam.

Heflin, J., Hendler, J., & Luke, S. (1998). Reading between the lines: Using SHOE to discover implicit knowledge from the web. AAAI-98 Workshop on AI and Information Integration (Vol. 297).

Kampa, S. R. (2002). Who are the expert? e-scholars in the semantic web. Ph. D. dissertation, University of Southampton.

Kando, N. (1997). Text-level structure of research papers: Implications for text-based information processing systems. Proceedings of the 19th Annual BCS-IRSG Colloquium on IR Research, Aberdeen, Scotland, UK, 68-81.

Kando, N. (1999). Text structure analysis as a tool to make retrieved documents sable. Proceedings of the 4th International Workshop on Information Retrieval with Asian Languages. 126-135.

Kintsch, W. (1974). The representation of meaning in memory. Oxford, England: Lawrence Erlbaum.

Luke, S., Spector, L., & Rager, D. (1996). Ontology-based knowledge discovery on the world-wide web. Working Notes of the Workshop on Internet-Based Information Systems at the 13th National Conference on Artificial Intelligence (AAAI96), 96-102.

Meyer, B. J. F. (1975). The organization of prose and its effects on memory. Amsterdam: North-Holland Publishing Co.

Ono, K., Sumita, K., & Miike, S. (1994). Abstract generation based on rhetorical structure extraction. Proceedings of the 15th conference on Computational linguistics, 1, 344-348.

Shotton, D. (2009). Cito, the citation typing ontology. Journal of Biomedical Semantics, 1(Suppl 1), S6. http://dx.doi.org/10.1186/2041-1480-1-s1-s6

Sibun, P. (1993). Domain structure, rhetorical structure, and text structure. Proceedings of the Workshop on Intentionality and Structure in Discourse Relations on the 31st ACL. Ohio State University Columbus, Ohio, USA.

Superceanu R. (1998). The rhetoric of scientific articles: A genre study. Timisoara, Rumania: Orizonturi Universitare.

van Dijk, T. A. (1980). Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition. Hillsdale, NJ: Lawrence Erlbaum.

Vater, H. (1994). Einführung in die textlinguistik. München: Wilhelm Fink Verlag.

Weigand, E. (2009). Language as dialogue: From rules to principles of probability. (Ed. Feller, S.), Philadelphia, PA: John Benjamins Publishing Company.

Werlich, E. (1976). A text grammar of English. Heidelberg: Quelle und Meyer.

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

Ko, Young-Man, & Song, Inseok (2011). A study on the knowledge organizing system of research papers based on semantic relation of the knowledge structure. Journal of the Korean Society for Information Management, 28(1), 145-170. http://dx.doi.org/10.3743/kosim.2011.28.1.145

Park, Chin-Yong (1997). A study on the method of process-driven analysis in text structure. Master's Thesis, Graduate School of the Korea National University of Education, Department of Korean Language Education.

Song, Min-Sun, & Ko, Young Man (2015). A study on the metadata based on the semantic structure of the Korean studies research articles. Journal of Korean Library and Information Science Society, 46(3), 277-299. http://dx.doi.org/10.16981/kliss.46.3.201509.277

Yoo, Sa-Rah, Lee, Hye-Won, & Song, Inseok (2009). A study on the application and management framework of social science scholarly ontology for semantic information navigation. Journal

of the Korean Society for Library and Information Science, 43(2), 277-298. http://dx.doi.org/10.4275/kslis.2009.43.2.277

Jeong, Yeo Hoon (2013). Rhetorical structure theory and the practice of Korean text analysis. Language Facts and Perspective, 32, 261-288.