

인문학 분야 연구데이터를 위한 연구논문의 의미구조 기반 메타데이터 연구*

A Study on the Metadata based on the Semantic Structure of the Humanities Research Articles for Research Data

안 병 군 (Byoung-Goon An)**

고 영 만 (Young Man Ko)***

초 록

본 연구의 목적은 인문학 분야 연구데이터와 연구논문을 연계할 수 있는 '인문학 분야 연구데이터를 위한 연구논문 의미구조 기반 메타데이터'를 개발하는 것이다. 인문학 분야 연구자의 관심 정보를 파악하기 위해 키워드 유형화를 진행하고 이를 일반적인 연구논문의 의미구조를 기준으로 분석함으로써 인문학 분야의 특성을 반영하기 위해 추가될 요소를 확인하였다. 이를 기반으로 인문학 분야 연구논문의 의미구조 기반 메타데이터를 구성하였으며 해당 메타데이터가 인문학 분야 연구데이터의 메타데이터로서 적합함을 평가하기 위하여 인문학 분야 연구자를 대상으로 설문을 진행하였다. 인문학 분야 연구논문의 의미구조를 기반으로 도출된 17개의 메타데이터 요소 중 인문학 분야 연구데이터 메타데이터로 활용하기에 적합하다고 판단된 요소는 7개이며, 그 외 10개 요소는 사용 가능한 요소로 확인되었다. 평가 결과를 바탕으로 인문학 분야 연구데이터를 위한 연구논문 의미구조기반 메타데이터를 구성하였다.

ABSTRACT

The purpose of this study is to build a metadata set based on the semantic structure of the humanities field research articles for research data that can associate research data and research articles. To understand the interest of Humanities researchers, keywords of research articles were categorized. By proceeding the keywords classification and analyzing based on the semantic structure of general research papers, additional elements to be added were identified for reflecting the characteristics of the humanities field. Based on it, 17 metadata elements based on the semantic structure of research papers in the humanities field was determined. In order to evaluate whether the metadata is appropriate as a metadata for research data in the humanities field, a survey was conducted on researchers in the humanities field. As a result of the survey, 7 elements were judged to be suitable for use as metadata for research data in the field of humanities, and 10 other elements were identified as usable elements. Drawing on the survey results, the metadata based on the semantic structure of the humanities field research articles for research data was constructed.

키워드: 연구데이터, 인문학, 의미구조, 메타데이터, 학술적 이해 형성

Research data, Humanities, Semantic structure, Metadata, Scholarly sensemaking

* 이 논문은 2021년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2021S1A5C2A02086918).

** 성균관대학교 인문학연구원(rns0305@skku.edu) (제1저자)

*** 성균관대학교 문헌정보학과 교수(ymko@skku.edu) (교신저자)

논문접수일자 : 2022년 2월 23일 논문심사일자 : 2022년 2월 25일 게재확정일자 : 2022년 3월 8일
한국비블리아학회지, 33(1): 345-369, 2022. <http://dx.doi.org/10.14699/kbiblia.2022.33.1.345>

※ Copyright © 2022 Korean Biblia Society for Library and Information Science

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

1.1 연구의 배경 및 필요성

연구자는 학술 연구 수행을 위하여 관련 분야의 선행 연구를 참고하고 다른 연구자와 연구성과를 공유하며, 이와 같은 과정을 통해 자신의 연구 분야에 대한 학술적 이해를 형성한다(고영만, 송인석, 2011). 과거에는 연구 문제에 결론을 도출하기 위해 참고하고 검토해야 할 기존의 연구 성과가 학술논문, 연구보고서 등 연구문헌에 한정되어 있었으나, 연구데이터의 공유와 활용에 대한 관심이 높아지고 관련 정책이 마련되면서 연구자가 참고할 수 있는 연구 성과의 범위가 연구데이터까지 확대되었다. 연구 과정에서 참고할 수 있는 자료가 다양해지는 것은 분명히 긍정적인 현상이지만, 연구데이터 자체는 고유성을 담보하는 요소가 분명하지 않은 경우가 많고 생산 배경이나 과정, 의도 등 구체적인 배경 정보가 제공되지 않을 경우 데이터를 정확하게 이해하기 어렵다. 따라서 충분한 정보 없이 제공되는 연구데이터는 오히려 연구자의 학술적 이해 형성에 어려움을 초래할 수 있다.

이러한 문제는 특히 인문학 분야 연구데이터에서 더욱 부각된다. 인문학 분야의 연구 방법은 세부 분야에 따라 굉장히 다양하고 정형화되어 있지 않으며, 생산되는 연구데이터의 종류도 다양하다. 또한 기존에 존재하던 자료에 대한 분석 및 해석 연구가 다수 존재하므로 다른 학문 분야와는 달리 기존에 존재하던 자료에 대한 정보도 중요하게 다뤄진다. 이와 같은 인문학 분야의 특성을 고려한다면, 인문학 분

야 연구데이터에 대해서는 단순히 연구데이터 자체에 대한 정보만을 제공하는 것만으로는 충분하지 않고 연구데이터가 생산된 연구 수행 과정과 배경 정보, 맥락까지 서술적인 방식으로 함께 제공될 수 있어야 한다(심원식, 안혜연, 변제연, 2015; Burrows, 2011).

인문학 분야 연구데이터를 제공할 때 연구데이터가 생산된 연구 과정에 대한 정보도 함께 제공하기 위한 방법으로 본 연구에서는 연구데이터와 연구논문 의미구조 연계를 제안하였다. 연구데이터 메타데이터에 연구논문 의미구조를 기반으로 구성된 메타데이터를 포함함으로써 연구의 생산 과정과 배경에 대한 정보도 더불어 제공하고자 하는 것이다. 연구문헌의 의미구조를 기반으로 연구자의 학술적 이해 형성을 도모하는 연구는 지속적으로 진행되어 왔으므로(송민선, 고영만, 2015; de Waard et al., 2006; Kando, 1997; van Dijk, 1980), 연구논문에 대한 정보를 기반으로 이해되어야 하는 연구데이터에도 연구논문의 의미구조를 기반으로 개발한 메타데이터가 유효하게 작용할 수 있다는 것이다.

인문학 분야 연구데이터를 위한 연구문헌 의미 구조 기반 메타데이터를 구성하고자 본 연구에서는 우선 연구데이터가 생산된 인문학 분야 연구논문을 대상으로 저자키워드를 수집하여 유형별로 분류함으로써 인문학 분야 연구자들이 필요로 하는 정보를 확인하였다. 이후 선행연구에서 제시된 연구문헌의 의미구조를 기준으로 유형화된 저자키워드를 분석하여 인문학 분야 연구문헌의 의미구조를 구성하였다. 인문학 분야 연구문헌의 의미구조를 기반으로 메타데이터 요소를 도출한 후, 각 요소가 인문

학 분야 연구데이터 메타데이터로 사용되기에 적합한지를 확인하고자 연구데이터 관리 경험이 있는 인문학 분야 연구자들을 대상으로 적합도 평가를 실시하였다. 적합도 평가 결과를 바탕으로 연구데이터를 위한 인문학 분야 연구 문헌의 의미 구조 기반 메타데이터를 구성하였다.

2. 이론적 배경 및 선행연구

2.1 연구문헌의 지식구조와 의미구조

지식구조는 연구자가 속한 학문 분야와 연구자가 요구하는 지식의 종류에 따라 개념 정의가 달라질 수 있으므로 지식구조 자체에 대한 합의된 정의는 존재하지 않는다. 문헌정보학에서는 지식의 범위가 넓은 점과 지식을 이해하기 위한 개념이 지속적으로 변화하고 발전한다는 점에 주목한다. 지식구조를 지식을 분류하고 조직함으로써 구성된 체계(송민선, 2015)로 간주하고 지식구조를 파악하기 위하여 지식구조를 구성하는 요소를 분석·조직화하는 연구를 진행해 왔다.

연구문헌은 해당 분야 고유의 지식구조를 통해 반영되는 연구자의 논증적 주장이며, 연구자는 논증적 주장을 수립하기 위해 연구 활동을 진행하는 과정에서 다양한 정보문제를 해결해 나간다(고영만, 송인석, 2011). 그러므로 해당 분야 고유의 지식구조는 연구문헌의 기틀이 되고 연구문헌의 집합은 지식구조를 구성하는 관계를 가지며, 연구문헌간의 의미관계를 파악하고 연구문헌 내에서 드러나는 지식구조의 구

성 요소와 구성 요소 간 의미구조를 분석함으로써 지식구조를 추론할 수 있게 된다.

연구문헌의 의미구조와 지식구조에 대한 연구는 다양한 측면에서 진행되어 왔다. van Dijk (1980)은 연구문헌의 한 종류인 학술논문을 분석하여 서론(문제 제기), 배경(연구배경, 이론, 실험 등), 결론으로 구성된 학술논문의 초구조 스키마(superstructure schema)를 제시하였으며, Superceanu(1998)는 van Dijk의 초구조 스키마를 과학 분야에 적용하여 과학 분야 논문의 초구조 모형을 고안하였다. 이후에도 과학 분야 연구문헌의 의미구조와 지식구조에 대한 연구는 지속적으로 진행되었다. de Waard et al.(2006)은 과학 연구문헌에서 발견되는 수사적 구조를 파악하고 이를 기반으로 과학 논문을 구조화하였다. 이를 바탕으로 문헌 내부의 수사적 요소와 요소들 간 논증 관계, 데이터 요소와 엔티티의 외부 소스와 연결을 구현할 수 있는 새로운 모델을 제시하고 이를 세포 생물학과 고생물학에 적용하였다. 후속연구인 de Waard(2007)에서는 과학 연구 문헌에 포함된 논증 구조를 표현하는 방법으로 텍스트의 수사 구조와 지식요소를 태깅하는 방법을 제시하였으며, 네 가지 마커(Fact, Hypothesis, Problem, Goal, Method, Result, Implication)를 세포생물학 분야 연구문헌에 실제로 태깅하고 문헌에서 각 태깅이 드러나는 양상과 관계를 분석하였다.

지식구조의 구성요소를 분석함으로써 연구문헌 검색의 정확성을 향상시키기 위한 연구도 진행되었다. Kando(1997)는 텍스트 구조를 활용한 텍스트 기반 정보 처리 시스템의 가능성을 확인하고자 연구를 진행하면서 연구문헌의 구

조를 분석하였으며, 이 과정에서 논문의 일반적인 기능 구성 요소를 밝혔다. 후속연구로 진행된 Kando(1999)에서는 연구문헌 전문 데이터베이스에 구조화된 태깅을 한 후 이를 기반으로 연구문헌의 텍스트 구조를 파악하고 텍스트 구조를 통해 검색된 문서의 활용을 돕는 정보시스템을 개발하였다.

고영만, 송인석(2011)에서는 연구문헌의 지식구조를 문서구조와 의미구조로 구분하고 의미구조는 다시 개념구조와 논증구조로 세분하였으며, 연구문헌 내 학술적 개념요소들 간에 어떠한 논증적인 의미관계가 성립되는지를 분석하였다. 이후 송민선, 고영만(2015)은 고영만, 송인석(2011)의 연구를 비롯한 선행연구에서 제시된 연구문헌 의미구조의 공통 의미를 정리하였다. 송민선, 고영만(2015)은 이를 바탕으로 논문의 의미구조를 반영하는 메타데이터를 개발하였고 송민선, 고영만, 이승준(2016)은 개발된 메타데이터를 활용하여 이용자의 검색 의도를 반영하는 검색 시스템을 구축함으로써 논문의 의미구조를 반영한 메타데이터의 유용성을 검증하였다.

2.2 인문학 분야 연구데이터

연구데이터에 대한 관심이 점차 높아지고 있으나 주로 논의가 이루어지는 분야는 자연과학이나 기술과학, 의학 분야였으며 근래에 들어서야 비로소 인문사회분야 연구데이터에 대한 관심이 생기고 있는 상황이다. 인문사회분야 연구데이터와 관련된 선행연구의 대부분은 연구데이터 리포지터리에 대한 분석 연구이다. 신영란, 정연경(2012)은 국내 가장 대표적인 인문사

회분야 리포지터리인 기초학문자료센터(Korea Research Memory, 이하 KRM)에 대한 분석과 연구자와의 면담을 통해 연구데이터 아카이브의 개선방안을 제시하였다. 조재인, 박종도(2019)는 re3data에서 정보를 제공하는 해외 인문사회분야 연구데이터 리포지터리 305개를 분석하여 인문사회분야 연구데이터의 전체적인 현황을 파악하고 이를 유형화하였다. 특히 인문학 분야와 관련하여 역사학 분야의 연구데이터는 대부분 이미지 자료인데 반해 언어학 분야는 시청각 자료 및 텍스트가 다수를 차지한다는 점을 밝히며 세부 학문분야에 따라 수집되는 데이터 유형에 차이가 있음을 밝혔다. Stigler와 Steiner(2018)는 예술 분야와 인문학 분야의 연구데이터를 대상으로 하는 리포지터리인 Geisteswissenschaftliches Asset Management System을 분석함으로써 연구데이터의 장기 보존과 데이터 큐레이션 정책을 확인하였다. 이처럼 인문학 분야를 다른 분야와 접목하여 통합적으로 관리하는 체계에 대한 연구는 존재하지만, 인문학 분야 연구데이터에만 집중된 연구는 극히 드물다. 심원식, 안혜연, 변제연(2015)은 KRM 시스템 운영자의 입장에서 시스템 구성과 연구데이터 구축 현황을 설명하고 인문학 분야에 주목하여 연구데이터 수집과 활용성을 높이기 위한 방안을 모색하였다. Burrows(2011)은 인문학 분야 연구데이터의 특징을 조사하고 Linked Open Data(이하 LOD)를 이용하여 인문학 분야 연구데이터 아카이브 구축을 위한 프레임워크를 분석하였다.

인문학 분야에서 특히 연구데이터에 대한 관심이 낮은 것은 인문학 분야의 특수한 연구 행위와 연구 방법 때문이라고 할 수 있다. 인문학

분야의 연구방법론과 데이터 접근 방식은 다른 학문분야와는 차이가 있고, 따라서 인문학 분야의 연구데이터에 사회과학이나 자연과학 분야의 연구데이터 구분을 적용하기 어렵다(Burrows, 2011)는 것이다. Burrows(2011)는 특히 인문학 분야의 경우 자연과학분야와는 달리 연구방법론이 명확하게 문서화되어있지 않고 연구방법도 다양하기 때문에 일반적인 연구 프로세스를 모델링하기 힘들다는 점을 지적하였다. 이와 더불어 인문학 분야에서는 연구 수행 과정을 반영하는 연구데이터 큐레이션이 중요하며 이를 위하여 대상에 대한 정보를 보다 서술적으로 기술할 필요가 있다고 설명하면서 그 방안으로 LOD의 이용을 제안하였다. 심원식, 안혜연, 변제연(2015)도 인문학 분야는 타 학문분야에 비해 세부 분야가 다양하여 통일된 연구 방법을 제시하기 어렵다는 점을 지적하며 Burrows(2011)와 견해를 같이했다. 이와 더불어 인문학 분야의 경우 기존에 존재하던 자료에 대한 분석 및 비교 연구가 다수 존재한다는 특징을 이야기하면서 연구 과정에서 생산되는 연구데이터 뿐만 아니라 연구에서 핵심적으로 참고·활용한 기존의 자료 또한 함께 관리되어야 한다고 설명하였다. 다양한 연구 방법을 포괄하고 기존에 존재한 자료에 대한 정보를 더불어 제공할 수 있는 연구데이터 큐레이션을 위하여 디지털 지식지도 구축의 필요성을 이야기하였다. Burrows(2011)와 심원식, 안혜연, 변제연(2015)의 연구를 종합적으로 고려함으로써 인문학 분야는 그 특성상 연구의 방법이 다양하고 정형화 되어있지 않으며 연구 과정에서 기존에 존재하던 자료가 주요하게 다루어지는 경우가 많다는 점을 파악할 수 있었다. 또한 이러

한 연구데이터의 특성을 고려하여 연구데이터 큐레이션을 제공할 때는 연구데이터와 관련된 연구 수행 과정과 인문학 분야의 지식구조를 반영할 수 있는 서술적인 기술 방식이 필요하다는 사실을 확인하였다.

3. 인문학 분야 연구논문의 의미구조 기반 메타데이터 도출

3.1 인문학 분야 연구논문의 키워드 유형 분석

연구문헌의 의미구조를 반영하는 메타데이터를 구성하기 위해서는 연구자가 요구하는 정보를 구체화하는 과정이 요구된다. 이를 위하여 인문학 분야 연구자들이 저술한 논문에서 수록된 저자키워드를 추출하고 이를 유형화하였다. 분석 대상 논문은 한국 연구재단 인문학 분야 연구지원사업의 결과로서 저술된 학술논문 중 연구데이터가 생산된 바 있는 논문으로 제한하였다. 이는 인문학 분야 연구데이터를 위한 연구문헌 의미구조 기반 메타데이터 개발이라는 본 연구의 목적을 고려한 것이다. 구체적으로, 2015년부터 2020년 사이에 KRM시스템에 연구데이터가 구축된 연구 사업의 학술논문을 분석의 대상으로 선정하였다. 동일한 해에 연구지원사업으로 선정되더라도 사업의 기간에 따라 연구 종료 일자가 상이하고 이에 따라 연구데이터 수집 및 구축 기간도 달라질 수 있으므로 연구지원사업 선정 연도나 종료 연도가 아닌 연구데이터 구축 연도를 기준으로 수집 대상 기간을 결정하였다.

2015년부터 2020년 사이에 KRM에 연구데

이터가 구축된 연구사업은 총 2,556개였으나, 데이터 수집 시점(2021년 5월)을 기준으로 연구 결과물로서 저자키워드가 존재하는 논문이 산출되지 않은 연구사업이 350개 존재하였다. 따라서 2,206개 연구 사업에서 산출된 2,873편의 논문을 대상으로 키워드를 수집하였다. <표 1>은 분석 대상 논문을 KCI의 학술지 구분에 따라 세부 학문분야별로 구분한 것이다. 인문학 분야 연구지원을 받은 연구사업에서 산출된 논문임에도 불구하고 KCI 학술지 구분상 인문학 분야에 해당하지 않는 논문(비인문학)도 232건 존재하였다. 비인문학으로 구분된 논문은 주로 교육학과 지역학, 복합학 분야에 해당하였으며, 대부분이 어문학 분야로서 지원을 받은 연구사업에서 산출된 논문이었다. 논문의 제목과 초록을 통해 확인한 결과, 어문학 분야의 세부 분야인 언어 교육 관련 연구와 작품 분석 과정에서 지역학적 내용이 포함된 연구 때문에 지역학과 교육학 분야 학술지에 게재된 논문이 존재하는 것으로 확인되었다. 또한 인문학 분야에서도

최근 다학제적 연구를 진행하고 있으므로 복합학 분야에도 논문이 게재된 것으로 파악되었다. 따라서 비인문학에 속한 논문 232건도 결국 인문학 분야 연구와 무관하지 않다고 판단하고 이를 분석의 대상에서 제외하지 않았다.

대부분의 키워드는 KCI와 RISS에서 제공하는 저자키워드 정보를 통해 수집 가능했으며, KCI와 RISS에 저자키워드 정보가 제공되지 않은 경우 논문 본문을 모두 확인하는 방법을 통해 최종적으로 16,989개의 키워드를 수집하였다. 데이터 수집 후 키워드에 병기된 특수기호 제거, 같은 키워드에 대한 다양한 외국어 표현 확인, 띄어쓰기 제거, 인명 표기 방식 통일 등의 데이터 정제 작업을 진행하였다. 고유 키워드는 최소 1회부터 최대 25회까지의 출현빈도를 보였으며, 이 중 출현빈도가 1회뿐인 키워드는 인문학 분야의 특성을 반영하는 것으로 간주하기 어려우므로 분석 대상에서 제외하였다. 최종적으로 중복 포함 6,568개, 중복 제외 2,216종의 키워드를 유형 분석의 대상으로 선정하였다.

<표 1> 분석 대상 논문 학문분야별 구분

학문분야	대상 논문 수(비율)	학문분야	대상 논문 수(비율)
사전학	2 (0.1)	일본어와 문학	231 (8.0)
역사학	265 (9.2)	기타동양어문학	11 (0.4)
철학	183 (6.4)	영어와 문학	213 (7.4)
종교학	24 (0.8)	프랑스어와 문학	98 (3.4)
기독교 신학	84 (2.9)	독일어와 문학	144 (5.0)
가톨릭 신학	8 (0.3)	스페인어와 문학	25 (0.9)
유교학	4 (0.1)	러시아어와 문학	38 (1.3)
불교학	25 (0.9)	서양고전어와 문학	1 (0.0)
언어학	145 (5.0)	기타서양어문학	10 (0.3)
문학	122 (4.2)	통역번역학	25 (0.9)
한국어와 문학	449 (15.6)	기타인문학	222 (7.7)
중국어와 문학	312 (10.9)	비 인문학	232 (8.1)
		합계	2,873 (100)

키워드 유형 분석의 기준으로 3단계 계층구조와 172개 범주로 구성되어 있는 구조적학술용어사전(Structural Terminology Net, 이하 STNet) 개념범주 텍사노미를 활용하였다(고영만 외, 2016). 키워드 유형 분류의 기준으로 사용한 STNet의 1단계와 2단계 개념범주는 <표

2>와 같다.

본 연구에서는 1단계 개념범주를 활용하되 1단계 개념범주가 광범위하여 세부 범주화가 요구되는 경우 2단계 개념범주를 선택하였다. 대표적으로 ‘언어/각국어’ 개념범주의 경우 어문학 분야 연구가 다수를 차지하는 인문학 분야의

<표 2> 유형 분류를 위해 적용한 STNet의 개념범주 1~2단계(2021년 11월 기준)

1단계 개념범주	2단계 개념범주
인간	인간(생물학적특성), 인간(관계), 인간(사회집단), 인간(능력/성향), 인간(직업/지위/역할), 준인간, 신체
기관/조직	행정/공공기관, 교육기관, 기업/회사, 사회(종교)조직/단체
자연물	동물, 식물, 자연(광물)
인공물	물품/제품/생산물, 재료/부품, 의류, 식료품, 도구/수단, 건축물(시설물), 교통수단, 저작(창작)물/정보
행위/행동/역할	행위/활동, 교육활동, 경제/산업활동, 위법행위, 신체운동/행동, 기능/역할
변화(변동)	완화(감소/축소/쇠퇴), 강화(증가/확장/신장), 개혁(개편/재편/혁신), 변천/과정, 분해/합체
특성(성질)	경향/동향, 기질/품질/형질/성향, 수준/정도, 능력/힘/에너지, 분포, 환경, 감각
심리	감정, 인식/의식
현상/이슈	상태/상황, 격차/차이, 문화/생활, 경제/경영/무역, 정치/국제
이론(사상, 이념, 주의, 법칙)	이론/사상, 원칙/법칙, 분과학문, 개념(정의)
제도(체제)	사회제도, 정치/법률제도, 경제경영제도
방법	연구/조사방법, 분석방법, 측정/척도, 지수/지표
기법/전략	기법/방식, 평가/분석, 교수/학습법, 전략/전술
유형/양식/장르	문학장르, 음악장르, 미술(디자인)양식, 스포츠/놀이유형, 공연예술
모형/기준	모델/모형, 패턴, 기준/규칙, 표준/규격, 인프라/구조/범위, 기호/신호
언어/각국어	언어/문자, 각국어
공간	인위적공간, 관념적공간, 자연적공간
지명	대륙(반도)명, 국가명, 도시/구/동명, 산/산맥명, 바다/강/호수명, 천문(체)명
시대/시간	시대, 기간, 시점
관계/상호작용	기원/유래/파생, 비교/차이, 계층/등급/계통, 인과, 상호작용
인명	실존인물, 가상인물
창작물명	문헌명, 예술작품명, 신문/잡지명, 방송(프로그램)명, 지도명, 문서(보고서)명
사건명	사건명/회담명, 국경일/기념일명, 행사/축제명, 수상명
기념물명(문화재)	
법률/제도명	법률/법령, 조약/협약, 정책(제도)명
기관/단체명	조직/단체명, 정부(왕조)명, 학과/종파명, 회의명
상품명	도구/수단명, 제품/브랜드명, 건축물/시설물명
기타주제어	

특성을 반영하여 하위 개념범주인 '각국어'와 '언어/문자'로 구분하고 1단계 개념범주는 활용하지 않았다. '인간' 개념범주에서도 마찬가지로 1단계 개념범주는 광범위한 반면 2단계 개념범주는 인간을 둘러싼 특성에 따라 '인간(관계)', '인간(직업/지위/역할)', '준인간' 등으로 세부구분이 되어 있으므로 2단계 개념범주를 활용하였다. '특성(성질)' 1단계 개념범주의 경우 하위 개념범주인 '수준/정도'와 '감각'에 해당하는 키워드는 두드러지는 반면 '경향/동향'이나 '분포', '기질/품질/형질/성향' 등 다른 하위요소는 비중도 크지 않고 키워드간 구분도 분명하지 않았다. 따라서 '특성(성질)' 1단계 개념범주를 사용하되, 2단계 하위범주 중 구분이 필요한 '수준/정도'와 '감각'만을 별도의 범주로 제시하는 방식을 택하였다([부록 1] 참조).

STNet에 이미 구축되어 있는 키워드는 기존의 키워드 유형 분류를 따랐으며, 그 외 검색을 통해 확인이 불가능한 키워드는 연구자가 분류하되, 제3자의 검토를 병행하였다. 이와 같은 방식을 통해 인문학 분야 연구논문에 수록된 저자 키워드를 <표 3>과 같이 46개 유형으로 구분하였다.

키워드 유형 분석 결과 가장 많은 키워드가 속한 유형은 '인명'이었으며, 인문학 분야 전반에서 확인되었다. '인명' 유형은 원래 실존인물과 작품 속 가상 인물을 포괄하는 유형이지만, 실질적으로 인명 유형에 포함된 키워드 대부분은 역사적 인물, 철학자, 작가이며 가상 인물에 해당하는 키워드는 5종뿐이다. '인명' 유형에 속한 키워드는 주로 연구에서 분석의 대상이 되거나 분석의 대상이 되는 작품, 사건, 이론 등의 창시자로서 연구된 경우이다. 어학과 언어

학이 다수 포함되어 있는 인문학 분야의 특성에 따라 '언어/문자' 유형에 포함되는 키워드도 다수 확인되었으며, 연구에서 주요하게 다뤄지거나 주된 분석 대상이 된 언어학적 현상이나 법칙, 개념, 문법 현상에 대한 키워드가 대부분이었다. '언어/문자' 유형과 달리 '각국어'에 해당하는 키워드는 많지 않았는데, 이는 일반적으로 연구의 대상으로 삼은 언어 자체는 키워드로 포함하지 않기 때문이다. '행위/행동/역할' 유형은 '글쓰기'나 '농업경영'과 같이 연구에서 분석의 대상으로 삼고 있는 행위나 활동을 지칭하는 키워드가 다수를 차지하고 있었으나, 연구 방법이나 분석 방법에 대한 키워드가 포함되기도 하였다. 인문학 분야에는 문학 작품 혹은 문헌을 기반으로 하거나 이를 분석하는 연구가 많기 때문에 다양한 문학 장르가 포함된 '유형/양식/장르' 유형과 작품명 및 문헌명을 포괄하는 '창작물명' 유형에 많은 키워드가 해당되었다. 두 유형은 대부분 연구의 분석 대상에 대한 키워드로 구성되었다는 공통점이 있으나 '유형/양식/장르'의 경우 해당 유형에 속하는 작품을 분석하기 위한 배경 정보로서 다뤄지기도 한다. '지명' 유형이나 '시대/시간' 유형의 경우 연구의 분석 대상에 해당하는 키워드도 존재하지만 분석 대상의 명확한 시공간적 범위를 설정하는 역할을 하는 키워드도 다수 포함되어 있다는 것이 특징이다.

인문학 분야 연구논문에서 추출한 키워드 유형을 분석함으로써 인문학 분야 연구자의 정보요구를 확인하고자 하였다. 인문학 분야의 경우 연구에서 분석의 대상(인명, 언어/문자, 창작물명, 현상/이슈, 사건명 등)이 되거나 주된 분석 대상의 배경 정보(유형/양식/장르, 지명, 시대/

〈표 3〉 STNet 개념범주에 따른 키워드 유형 분석

No	키워드 유형 분류	중복 제외 키워드 수(비율)	중복 포함 키워드 전체 출현 수(비율)
1	인명	257 (11.60)	752 (11.45)
2	행위/행동/역할	197 (8.89)	560 (8.53)
3	언어/문자	189 (8.53)	541 (8.24)
4	유형/양식/장르	143 (6.45)	418 (6.36)
5	이론(사상, 이념, 주의, 법칙)	131 (5.91)	401 (6.11)
6	개념(정의)	107 (4.83)	301 (4.58)
7	특성(성질)	103 (4.65)	287 (4.37)
8	기법/전략	87 (3.93)	264 (4.02)
9	장작물명	108 (4.87)	260 (3.96)
10	인식/의식	73 (3.29)	244 (3.71)
11	지명	69 (3.11)	204 (3.11)
12	제도(체제)	56 (2.53)	185 (2.82)
13	인간(사회집단)	61 (2.75)	175 (2.66)
14	공간	46 (2.08)	161 (2.45)
15	시대/시간	37 (1.67)	151 (2.30)
16	저작(창작)물/정보	51 (2.30)	150 (2.28)
17	감정	36 (1.62)	146 (2.22)
18	상태/상황	48 (2.17)	145 (2.21)
19	분과학문	45 (2.03)	126 (1.92)
20	문화/생활	35 (1.58)	107 (1.63)
21	관계/상호작용	32 (1.44)	104 (1.58)
22	인간(직업/지위/역할)	32 (1.44)	86 (1.31)
23	모형/기준	33 (1.49)	86 (1.31)
24	기관/단체명	32 (1.44)	80 (1.22)
25	사건명	24 (1.08)	66 (1.00)
26	인간(능력/성향)	11 (0.50)	63 (0.96)
27	각국어	18 (0.81)	57 (0.87)
28	현상/이슈	19 (0.86)	50 (0.76)
29	신체	16 (0.72)	49 (0.75)
30	변화(변동)	14 (0.63)	49 (0.75)
31	방법	14 (0.63)	42 (0.64)
32	인간(생물학적특성)	7 (0.32)	35 (0.53)
33	수준/정도	12 (0.54)	29 (0.44)
34	자연물	12 (0.54)	26 (0.40)
35	감각	4 (0.18)	26 (0.40)
36	물품/제품/생산물	10 (0.45)	24 (0.37)
37	기관/조직	8 (0.36)	23 (0.35)
38	도구/수단	8 (0.36)	21 (0.32)
39	준인간	8 (0.36)	18 (0.27)
40	인간(관계)	5 (0.23)	14 (0.21)
41	건축물(시설물)	6 (0.27)	13 (0.20)
42	인공물	3 (0.14)	8 (0.12)
43	법률/제도명	2 (0.09)	4 (0.06)
44	상품명	2 (0.09)	4 (0.06)
45	기념물명(문화재)	1 (0.05)	2 (0.03)
46	기타주제어	4 (0.18)	11 (0.17)
	합계	2,216 (100)	6,568 (100)

시간)를 중요한 정보로 여긴다는 것을 확인할 수 있었다. 또한 연구논문에서 활용된 중요 개념이나 법칙, 이론(이론(사상, 이념, 주의, 법칙), 언어/문자, 개념(정의))과 연구의 특수한 연구방법(행위/행동/역할, 기법/전략, 방법)에 대한 정보 요구도 존재한다는 사실을 파악하였다.

3.2 연구논문의 의미구조

인문학 분야 연구논문의 키워드 유형 분석을

기반으로 인문학 분야 연구자가 주로 사용하는 키워드는 무엇이며 어떤 유형에 해당하는지를 파악하였다. 연구자의 정보 요구를 반영하는 키워드를 다시 활용하여 각 유형에 속하는 키워드를 의미구조에 따라 분석함으로써 연구논문의 의미구조를 반영하는 메타데이터 항목을 도출하였다. 이 과정에서 송민선, 고영만(2015)이 제시한 기존 연구문헌의 의미구조 분석을 통한 공통 항목 개념화를 기준으로 사용하였으며, 이는 <표 4>와 같다.

<표 4> 기존 연구문헌의 의미구조 분석을 통한 공통 항목 개념화(송민선, 고영만, 2015)

참고한 기존 연구들				공통 의미 개념화
Kando(1999)	유사라, 이혜원, 송인석(2009)	고영만, 송인석(2011)		
		학술적 개념요소	논증적 의미관계	
- Background - Purpose	연구목적 조사목적	문제/이슈 개념/정의	연구목적	연구목적
- Rationale	연구의의	관점 이해 구성 특성	연구관점	연구의의
	연구범위	개념/정의 이해 관점 방법 데이터/자료		연구 범위 제한
- Scope of the Study - Framework of the study - Data Collection - Data Analysis	연구유형	연구유형	연구문제 연구방법	연구유형
- Research Questions	연구문제	문제/이슈	연구문제	연구문제
- Hypothesis	가설	문제/이슈		가설
- Validity of the evidence or methods - Examination of the Evidence	연구방법	방법	연구방법	연구방법
- Data Collection	자료수집	데이터/자료 방법	연구활동	연구 대상 데이터
	표본설계	방법		데이터 수집 규모 데이터 수집 방법
- Data Analysis	자료분석	수단/도구		분석방법
- Conclusions	연구결과	이해 구성 특성	연구결과	연구결과
- Future Research	후속연구 제언	방법		후속연구 제언

송민선, 고영만(2015)이 연구문헌의 의미구조 도출을 위해 활용한 기존 연구는 Kando(1999), 유사라, 이혜원, 송인석(2009), 고영만, 송인석(2011)이다. Kando(1999)는 4개 학문분야에 대한 100편가량의 논문 분석과 연구자 인터뷰를 기반으로 3계층까지 세분화된 연구논문의 의미구조를 도출하였다. Kando(1999)의 지식구조 모형은 1단계는 Problems, Validity of the evidence or methods, Examination of the Evidence, Answers로 비교적 간단하게 나뉘지만 2단계는 25개로 구분되고 그중 8개 구분은 3단계 계층까지 존재한다. 송민선, 고영만(2015)에서 Kando(1999)의 의미구조를 기반으로 공통 의미를 개념화할 때는 지나치게 단순한 1단계 요소 대신 2단계와 3단계를 그 대상으로 하였다. 유사라, 이혜원, 송인석(2009)은 연구성과물의 의미검색을 지원하는 온톨로지 구축을 목적으로 KRM의 용어 데이터 레지스트리(이하 DR)와 연구성과물 DB를 분석하고 KRM 업무담당자와의 인터뷰를 진행하였다. 이 과정에서 Kando(1999)의 3계층 의미구조를 5계층까지 세분화하였으며, 특히 연구의 측정분석과 관련된 계층을 구체적으로 다루었다. 유사라, 이혜원, 송인석(2009)의 의미구조를 정리하기 위하여 송민선, 고영만(2015)이 주된 대상으로 삼은 것은 1단계와 2단계이다. 고영만, 송인석(2011)도 KRM의 용어 DR과 연구성과물 DB를 분석의 대상으로 하되, 연구문헌의 지식구조를 파악하는 과정에서 용어DR의 개념클래스와 인스턴스를 주로 분석하였다. 이를 통해 연구문헌의 학술적 개념요소 10개와 개념요소 사이의 논증적인 의미관계 6가지를 제시하였다. 유사라, 이혜원, 송인석(2009)의 연구에 비해 고

영만, 송인석(2011)에서 제시하는 연구문헌 지식구조는 훨씬 단순화되어 있으므로 송민선, 고영만(2015)에서는 학술적 개념요소와 논증적 의미관계 모두를 분석하여 공통의 의미구조를 파악하였다.

송민선, 고영만(2015)은 세 개의 선행연구에서 제시된 연구문헌의 의미구조를 분석하여 13개의 공통 의미를 도출하였다. 송민선, 고영만(2015)이 기존의 연구와 달리 단일 계층 구조를 제시한 것은 의미 검색 시스템을 위한 메타데이터 구성 단계와 데이터 구축 단계를 고려한 결과로, 지나친 계층화는 효율성과 표준화 측면에서 문제가 될 수 있기 때문이다. 인문학 분야 연구문헌의 의미구조 분석에서 그치지 않고 이를 기반으로 연구데이터를 위한 메타데이터 구성까지 고려하는 본 연구에서는 송민선, 고영만(2015)의 의미구조가 분석의 기준으로 적합하다고 판단하였다. 송민선, 고영만(2015)의 의미구조를 기준으로 선행 연구의 의미구조가 어떻게 활용되었는지를 종합적으로 조사하여 각 의미구조를 <표 5>와 같이 정리하였다.

3.3 인문학 분야 연구논문의 의미구조 기반 메타데이터 항목 도출

송민선, 고영만(2015)의 연구논문 의미구조 기준을 활용하여 유형 분석이 완료된 키워드가 어떤 의미구조에 해당하는지 분류하고 각 유형의 경향을 파악했다. 구체적으로, 논문의 제목과 초록을 바탕으로 키워드가 어떤 의미구조에 해당하는지를 파악하였다. 이후 한 유형에서 하나의 의미구조가 10% 이상 나타나는 경

〈표 5〉 의미구조별 정리 및 포함 사항

의미구조	의미구조별 정리 및 포함 사항
연구목적	- 연구 주제를 결정한 동기와 연구를 통해 달성하고자 하는 목적 관련 내용 - 연구 배경, 연구의 필요성, 연구에서 사용하는 개념의 개념화, 이론 정립을 포함함
연구의의	- 연구를 통해 얻고자 하는 학문적·사회적 공헌 및 영향 관련 내용 - 연구 주제의 특징과 이를 기반으로 한 연구의 이점, 중요성, 공헌도 및 활용도, 영향력 포함
연구범위 제한	- 연구의 대상, 기간, 시공간적 범위 등 연구의 범위 설정 관련 내용 - 관찰 대상, 문제영역, 주제범위, 연구대상 특성 및 분석단위 포함
연구유형	- 연구 목적, 내용, 접근 방법, 데이터 수집 방법, 연구 방법 등에 따른 연구 유형 구분 관련 내용
연구문제	- 연구하고자 하는 주제 혹은 질문 관련 내용. '연구목적'보다 구체적인 내용으로 구성 - 연구에서 지적하는 문제점과 목표, 가정을 포함함
가설	- 연구에서 설정한 변수 및 변인 간 관계에 대한 예측 관련 내용 - 가설에 사용된 변인, 가설을 통해 밝히고자 하는 관계를 포함함
연구방법	- 연구 수행을 위하여 사용한 방법 관련 내용. '연구유형'보다 구체적인 내용으로 구성 - 연구 설계, 조사범위, 도구 설정, 조사 방법을 포함함
연구대상	- 연구에서 분석의 대상이 된 객체 - 주된 분석의 대상이 된 문헌, 인물, 개념 등 객체를 포함함
데이터수집방법	- 연구에서 자료 혹은 데이터 수집을 위해 사용한 방법 관련 내용 - 표본 설계, 데이터 수집 절차·도구·재료와 데이터 수집 절차·조건을 포함함
데이터수집규모	- 연구에서 자료 혹은 데이터 수집을 위해 사용한 방법에 따른 규모 및 범위 관련 내용 - 연구의 데이터 수집 범위, 단위 및 기준, 크기를 포함함
분석방법	- 연구에서 수집한 데이터 및 자료에 대한 분석 방법 관련 내용 - '연구방법'보다 분석에 초점을 둔 세부적인 내용 포함 - 분석 절차, 기술, 모형, 방법 선택 이유, 검증 유형 및 방법을 포함함
연구결과	- 연구를 통해 도출된 결과에 대한 의미 및 해석 관련 내용 - 가설의 검증 여부, 연구를 통해 도출된 새로운 개념 정의 및 객체간 관계 정리를 포함함
후속연구제언	- 연구에 포함되지는 않았으나 연구를 기반으로 분석 및 연구할 수 있는 문제 관련 내용 - 연구의 제한점, 연구성과 활용방안을 포함함

우 그 유형이 해당 의미구조에 포함되는 것으로 간주했다. 키워드 전체 출현 수가 300개 이상으로 지나치게 많거나 30회 미만으로 아주 적은 경우 10%라는 기준을 엄격하게 적용하는 대신, 2~5개의 정도의 여유를 두어 유형에 속한 키워드 중 충분히 다수가 특정 의미구조에 해당하는지를 확인했다. 하나의 키워드가 두 개 이상의 의미구조에 속하는 경우에는 임의로 하나의 의미구조를 선택하지 않고 모든 의미구조를 사용하는 방식을 택하였다. 현재 기준으로 활용된 송민선, 고영만(2015)의 연구논문 의미구조에

따라 분류하기 어려운 키워드는 그 특징을 별도로 기술해 놓은 후 취합하여 유형을 확인하였으며, 이를 토대로 추가되어야 할 의미구조를 파악하였다. 이와 같은 절차와 기준을 기반으로 키워드 유형 범주 별 연구논문의 의미구조를 분석한 결과는 〈표 6〉과 같다.

키워드 유형을 의미구조에 따라 분석한 결과, 가장 많은 유형이 해당하는 의미구조는 '연구대상'이다. 연구 대상의 선정은 다른 연구와의 구분점이 될 수 있고, 또 가장 분명하게 드러나기 때문에 대부분의 논문에서 연구 대상을 키워

〈표 6〉 연구논문의 의미구조와 키워드 유형 범주 분석 적용 결과

의미구조	키워드 유형 분석 적용		
연구목적	이론(사상, 이념, 주의, 법칙) 행위/행동/역할 인간(사회집단) 상태/상황 감각 변화(변동) 인간(관계)	제도(체제) 특성(성질) 공간 분과학문 문화/생활 신체	개념(정의) 인식/의식 감정 관계/상호작용 모형/기준 인간
연구의의	이론(사상, 이념, 주의, 법칙) 행위/행동/역할	제도(체제)	개념(정의)
연구 범위 제한	방법 공간 인간(능력/성향)	시대/시간 상태/상황 신체	지명 기관/단체명 수준/정도
연구유형	분과학문		
연구문제	개념(정의) 공간	특성(성질) 감정	인식/의식
가설	개념(정의)	특성(성질)	
연구방법	방법 기법/전략	분과학문 모형/기준	행위/행동/역할
연구 대상	인명 사건명 건축물(시설물) 각국어 지명 저작(창작)물/정보 문화/생활 모형/기준 변화(변동) 인간 물품/제품/생산품	창작물명 도구/수단 인공물 유형/양식/장르 공간 분과학문 인간(관계) 기관/단체명 현상/이슈 감각 준인간	인간(사회집단) 기관/조직 언어/문자 인식/의식 시대/시간 관계/상호작용 인간(직업/지위/역할) 인간(능력/성향) 신체 자연물 건축물(시설물)
데이터 수집 방법	방법	수준/정도	
데이터 수집 규모	방법		
분석방법	방법	기법/전략	
연구결과	이론(사상, 이념, 주의, 법칙) 인간	행위/행동/역할	상태/상황
후속연구 제언	이론(사상, 이념, 주의, 법칙)		

드로 포함하고 있었다. 특정 작품이나 인물, 기관이 분석의 대상이 되기도 하고, 작품의 분류나 대상 사이의 관계, 변화 양상과 같이 추상적 개념을 대상으로 하는 연구도 다수 존재하였다. ‘연구 목적’ 의미구조는 연구의 배경과 필요성,

연구에 필요한 개념 정의를 모두 포함하고 있으므로 이에 해당하는 유형도 상당수 존재하였다. 키워드를 통해서 연구 목적을 밝히는 경우는 대부분 연구를 통해 사회 문제에 대한 해결책을 모색하거나 아직 합의되지 않은 개념을 정리하

고자 하는 경우였다. 모든 의미구조에서 하나 이상의 키워드 유형 범주가 확인되었고, 따라서 기준으로 제시된 의미구조 중 어떤 항목도 제외되지 않았다.

여러 유형에 속하는 각 키워드를 연구논문의 의미구조에 적용하는 과정에서 현재의 의미구조에 포함시키기 어려운 키워드들이 발견되었다. 연구 대상 자체는 아니지만 연구 대상에 대한 이해를 높이고 심도 깊은 분석을 진행하기 위하여 연구 대상의 배경 정보를 함께 연구하는 경우가 다수 존재하였다. 예를 들어 작품이나 문헌에 대한 분석 연구를 진행하는 과정에서 저술자의 생애와 사상, 생활사를 다루거나 역사적 단체나 기관을 분석하는 것을 목적으로 하는 연구에서 관련된 특정 시대나 사건을 더불어 조사하는 경우가 이에 해당한다. 주로 '지명', '인명', '기관/단체명', '시대/시간', '사건명' 유형에 포함된 키워드들이 이에 해당하였다. 이러한 키워드를 의미구조에 포함시키기 위하여 연구 대상의 배경에 대한 의미구조를 추가할 필요가 있다는 점이 확인되었으며, 구체적으로 '연구 대상의 배경이 되는 인물 및 단체', '연구 대상의 배경이 되는 공간적 범위', '연구 대상의 배경이 되는 시간적 범위(사건)'로 세분화하였다. 또한 연구 대상의 분석 기준 수립이나 분석 방법 선택 과정에서 선행 연구의 이론이나 개념, 모델이 활용되는 경우가 다수 존재한다. 단순한 참고 및 활용을 넘어서 연구에서 해당 이론·개념·모델에 대해 탐구하고 이를 연구의 핵심적인 내용 중 하나로 판단하여 키워드로도 사용한다면 별도의 의미구조로 포함시킬 필요가 있으며, 이를 '연구에 중요하게 관련되거나 차용한 이론 및 개념'으로 명명하였

다. 주로 '이론(사상, 이념, 주의, 법칙)' 유형이나 '개념(정의)', '모형(기준)' 유형에 해당하는 키워드들이 연구와 중요하게 관련되거나 연구에서 핵심적으로 차용한 이론이나 개념에 해당한다.

연구데이터가 생산된 바 있는 인문학 분야 연구논문에서 저자 키워드를 추출하고 이를 유형화함으로써 연구자들의 실제 요구를 조사하였으며, 각 키워드가 해당하는 논문의 의미구조를 파악하는 단계에서 추가되어야 할 의미구조를 확인하였다. <표 7>은 이와 같은 분석을 기반으로 도출한 인문학 분야 연구논문의 의미구조를 기반으로 하는 메타데이터이다.

4. 인문학 분야 연구데이터를 위한 연구논문의 의미구조 기반 메타데이터 항목 평가

분석을 통해 도출한 인문학 분야 연구논문의 의미구조 기반 메타데이터를 연구데이터를 설명하는 요소로 활용하기에 적합한지 확인하기 위해 인문학 분야 연구자를 대상으로 설문조사를 실시하였다. 인문학 분야 연구자는 연구데이터의 생산자이면서 동시에 다른 연구자의 연구데이터를 활용하는 이용자이기도 하다. 따라서 인문학 분야 연구자에게 연구문헌의 의미구조 기반 메타데이터가 연구데이터를 이해하는 과정에서 활용하기에도 적합한지를 질문함으로써 연구데이터 생산자와 이용자의 입장을 모두 반영할 수 있다고 판단하였다. 인문학 분야 연구자 중에서도 연구데이터에 대한 이해가 있는 연구자를 선정하기 위하여 연구데이터의 생

〈표 7〉 인문학 분야 연구논문 의미구조 기반 메타데이터 요소

No	메타데이터 요소	메타데이터 설명
1	Research Objectives(연구 목적)	- 논문에 제시된 연구 배경과 목적
2	Research Significance(연구 의의)	- 논문에 제시된 연구 의의 및 연구의 영향
3	Research Limitation(연구범위 제한)	- 논문에 제시된 연구 범위 설정 및 제한사항
4	Research Types(연구 유형)	- 연구에서 사용한 연구 방법, 분석 방법, 자료 수집 방법 등에 따른 연구 유형 구분
5	Research Problem(연구 문제)	- 논문에 제시된 연구 문제 혹은 질문
6	Hypothesis(연구 가설)	- 논문에 제시된 연구 가설
7	Research Subject(연구 대상)	- 연구에서 분석 및 탐구의 대상으로 하는 객체
8	Personage/Affiliation Background (연구 대상의 배경이 되는 인물·단체)	- 연구 대상의 배경이 되는 인물 및 단체 - 연구 대상의 배경 정보로서 연구된 인물 혹은 단체 정보
9	Temporal Background (연구 대상의 배경이 되는 시간적 범위)	- 연구 대상의 배경이 되는 시간적 범위 - 연구 대상의 배경 정보로서 연구된 시대, 시기, 사건 정보
10	Spatial Background (연구 대상의 배경이 되는 지리적 범위)	- 연구 대상의 배경이 되는 지리적 범위 - 연구 대상의 배경 정보로서 연구된 구체적인 지역 정보
11	Research Method(연구 방법)	- 연구에서 사용한 구체적인 연구 방법
12	Related Models(연구 모델)	- 연구에 중요하게 관련되거나 차용한 이론 및 개념, 모형
13	Method of Data Collection (데이터 수집 방법)	- 연구에서 자료 혹은 데이터 수집을 위해 사용한 방법
14	Data Scale(데이터수집규모)	- 자료 혹은 데이터 수집을 위해 사용한 방법에 따른 규모 및 범위
15	Analysis Methods(분석 방법)	- 연구에서 수집한 데이터 및 자료에 대한 분석 방법(통계 방법, 데이터 분석 방법 등)
16	Research Outcome(연구 결과)	- 논문에 제시된 연구 결과
17	Further Research(후속연구 제언)	- 논문에 제시된 후속연구 관련 내용

산 및 관리 경험이 있는 연구자로 대상을 제한하였다. 구체적으로 한국 연구재단에서 인문학 분야 연구 사업을 지원받아 2015년부터 2020년 사이에 KRM에 연구데이터를 제출하고 데이터 구축이 완료된 연구자를 설문 대상자로 하였으며, 이 중 임의 표본 추출로 145명의 연구자에게 메일을 통해 설문 참여 의사를 확인하였다. 설문조사는 2022년 1월 19일부터 2월 8일까지 21일간 진행되었으며, 설문 참여에 동의한 후 이메일을 통해 발송한 설문지에 응답을 회신한 연구자는 40명이다.

연구를 통해 도출한 17개 메타데이터 요소 각각에 요소명과 요소 설명, 사용 예시를 제시

하였고, 각 요소를 인문학 분야 연구데이터의 메타데이터로 사용하였을 때의 적합도를 5점 척도(매우 적합, 적합, 보통, 부적합, 매우 부적합)로 평가할 수 있도록 구성하였다. 설문 분석 단계에서는 각 요소별 응답에 점수를 부여하고(매우 적합 5점, 적합 4점, 보통 3점, 부적합 2점, 매우 부적합 1점) 점수의 평균과 표준편차를 바탕으로 요소의 적합도 여부를 판단하였다. 인문학 분야 연구데이터의 메타데이터로 적합하다고 판단하는 기준은 평균 점수 4점(적합) 이상이면서 표준편차가 1 미만인 요소이다. 평균적으로 3점(보통)을 초과하는 응답을 받은 경우 필요에 따라 사용 가능한 요소로 간주하

였다.

〈표 8〉은 연구자의 설문 응답을 정리하고 이를 바탕으로 점수를 부여하여 평균 점수가 높은 순서대로 나열한 것이다.

가장 높은 점수를 얻은 메타데이터 요소는 'Research Objectives'와 'Data Scale'이며, 7개의 요소('Research Objectives', 'Data Scale', 'Research Outcome', 'Research Subject', 'Research Significance', 'Research Methods', 'Personage/Affiliation Background')가 평균적으로 4점 이상의 평가를 받아 인문학 분야 연구데이터의 메타데이터로 사용하기에 적합한 요소로 판단되었다. 적합으로 판단되지 않은

10개 요소 중 '부적합' 응답을 받은 요소는 존재하였으나, 전체 응답자의 의견을 취합하였을 때 10개 요소 모두의 적합도 평균이 3.5점 이상이었다. 이에 따라 평균 점수 3.5점 이상을 획득한 10개 요소에 대해서는 필요한 경우 인문학 분야 연구데이터의 메타데이터로 사용할 수 있는 요소로 판단하였다. 모든 요소에서 1미만의 표준편차가 나타났으므로 응답 간 편차는 크지 않으며, 따라서 표준편차로 인하여 제외되는 요소는 없었다. 설문 결과를 바탕으로 인문학 분야 연구데이터를 위한 연구논문 의미구조기반 메타데이터를 구성하였다([부록 2] 참조).

〈표 8〉 설문 응답자의 메타데이터 요소 점수 부여 결과(5점 척도)

메타데이터 요소	연구자 응답					점수 평균	표준 편차
	매우 부적합	부적합	보통	적합	매우 적합		
Research Objectives(연구 목적)	0	0	6	17	17	4.275	0.707
Data Scale(데이터수집규모)	0	2	5	13	20	4.275	0.866
Research Outcome(연구 결과)	0	0	8	16	16	4.200	0.748
Research Subject(연구 대상)	0	0	10	15	15	4.125	0.781
Research Significance(연구 의의)	0	1	9	15	15	4.100	0.831
Research Method(연구 방법)	0	0	9	18	13	4.100	0.735
Personage/Affiliation Background (연구 대상의 배경이 되는 인물·단체)	0	0	11	15	14	4.075	0.787
Method of Data Collection(데이터수집방법)	0	2	7	21	10	3.975	0.790
Analysis Methods(분석 방법)	0	2	9	17	12	3.975	0.851
Research Types(연구 유형)	0	1	12	16	11	3.925	0.818
Related Models(연구 모델)	0	2	10	17	11	3.925	0.848
Spatial Background (연구 대상의 배경이 되는 지리적 범위)	0	1	12	18	9	3.875	0.781
Temporal Background (연구 대상의 배경이 되는 시간적 범위)	0	1	13	16	10	3.875	0.812
Hypothesis(연구 가설)	0	1	14	15	10	3.850	0.823
Further Research(후속연구 제안) * 미응답 1건 존재	0	2	17	10	10	3.718	0.904
Research Limitation(연구범위 제한)	0	4	13	16	7	3.650	0.882
Research Problem(연구 문제)	0	2	21	9	8	3.575	0.863

응답자가 서술한 추가 의견을 분석한 결과 'Method of Data Collection', 'Analysis Methods', 'Research Types', 'Related Models'가 4점(적합) 이상을 획득하지 못한 이유는 해당 항목이 'Research Methods' 요소와 내용적으로 상당히 중복될 수 있다고 판단했기 때문으로 확인되었다. 두드러지는 특징은 연구 대상의 배경이 되는 시간적, 지리적 범위(Spatial Background, Temporal Background)는 적합도 점수 평균이 동일하나, 연구 대상의 배경이 되는 인물 및 단체(Personage/Affiliation Background)의 점수 평균만 높게 나타난 것이다. 이 같은 결과를 키워드 유형 분류에서 '인물' 유형에 속하는 키워드가 가장 많았던 것과 종합적으로 고려했을 때, 인문학 분야 연구에서 인물 및 단체의 중요도가 높다는 사실을 추정할 수 있었다.

5. 결론

본 연구는 아직 연구데이터에 대한 논의가 활발하게 이루어지지 않은 인문학 분야에 주목하였다. 인문학 분야 연구의 특성을 고려하였을 때 연구자가 원활하게 연구데이터를 활용하기 위해서는 연구데이터와 연구논문을 연계하는 학술적 이해가 요구된다는 견해를 바탕으로, 연구데이터와 연구논문을 연계하는 학술적 이해 형성을 지원하기 위하여 연구논문의 의미구조를 기반으로 한 메타데이터를 연구데이터의 메타데이터로서 활용하는 방안을 구상하였다.

인문학 분야 연구자들이 필요로 하는 연구논문의 의미 정보를 도출하기 위해 연구데이터가 생산된 인문학 분야 연구논문에서 추출한 키워

드의 유형을 분석한 결과 전체 키워드를 46개 유형으로 구분할 수 있었다. 송민선, 고영만(2015)에서 제시한 연구논문의 일반적인 의미구조를 기준으로 인문학 분야 연구논문의 의미구조를 도출하였으며, 기준으로 제시된 13개 의미구조 외에 연구 대상의 배경이 되는 인물·단체, 지리적 범위, 시간적 범위에 대한 의미구조와 연구에서 핵심적으로 차용한 이론·개념·모델을 포괄할 수 있는 의미구조가 추가되었다. 인문학 분야 연구논문의 의미구조를 기반으로 Research Objectives, Research Significance, Research Limitation, Research Types, Research Problem, Hypothesis, Research Method, Related Models, Research Subject, Personage/Affiliation Background, SpatialBackground, Temporal Background, Data Scale, Method Of Data Collection, Analysis Methods, Research Outcome, Further Research 17개의 메타데이터 요소를 도출하였다. 연구를 통해 도출한 메타데이터 각 요소를 인문학 분야 연구데이터 메타데이터로 사용하기에 적합한지를 확인하고자 연구데이터를 관리한 경험이 있는 인문학 분야 연구자를 대상으로 평가를 실시하였으며, 그 결과 Research Objectives, Data Scale, Research Outcome, Research Subject, Research Significance, Research Methods, Personage/Affiliation Background 7개 요소는 적합한 요소로, 그 외 10개 요소는 필요한 경우 사용 가능한 요소로 확인되었다.

본 연구는 연구문헌의 의미구조와 연구데이터 메타데이터의 연계를 시도하였다는 점에서 의의가 있으며, 인문학 분야 연구 문헌의 키워드 유형화와 연구자들의 메타데이터 적합도 평

가를 통해 인문학 분야 연구자의 정보 요구를 실질적으로 반영하고자 했다는 점에서도 의미를 확인할 수 있다. 그러나 본 연구를 통해 개발한 ‘인문학 분야 연구데이터를 위한 연구논문의 의미구조 기반 메타데이터’를 실제 검색 시

시스템으로 구현하여 메타데이터 항목의 유용성을 평가하기 위해서는 보다 체계화된 메타데이터가 필요하므로 이를 위한 심화 연구가 요구된다.

참 고 문 헌

- 고영만, 김비연, 민혜령, 송민선, 이승준 (2016). 구조적학술용어사전 기반 온톨로지 구축론. 서울: 한국도서관협회.
- 고영만, 송인석 (2011). 연구문헌의 지식구조를 반영하는 의미기반의 지식조직체계에 관한 연구. 정보관리학회지, 28(1), 145-170. <http://dx.doi.org/10.3743/kosim.2011.28.1.145>
- 송민선 (2015). 한국학 분야의 지식구조 분석 연구. 한국문헌정보학회지, 49(4), 125-157. <http://dx.doi.org/10.4275/KSLIS.2015.49.4.125>
- 송민선, 고영만 (2015). 한국학 연구 논문의 의미 구조 기반 메타데이터 연구. 한국도서관·정보학회지, 46(3), 277-299. <http://dx.doi.org/10.16981/kliss.46.3.201509.277>
- 송민선, 고영만, 이승준 (2016). 한국학 연구 논문의 텍스트 구조 기반 메타데이터 검색 시스템 개발 연구. 정보관리학회지, 33(3), 155-176. <http://dx.doi.org/10.3743/KOSIM.2016.33.3.155>
- 신영란, 정연경 (2012). 국내 인문사회 연구데이터 아카이브의 개선방안에 관한 연구. 한국기록관리학회지, 12(3), 93-115.
- 심원식, 안혜연, 변제연 (2015). 인문학 분야 연구데이터의 수집 및 활용성 증진을 위한 전략 연구. 한국문헌정보학회지, 49(3), 155-183. <http://dx.doi.org/10.4275/KSLIS.2015.49.3.155>
- 유사라, 이혜원, 송인석 (2009). 연구자 중심 연구성과를 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>
- 조재인, 박종도 (2019). re3data를 기반으로 한 인문사회 RDR 연구. 한국비블리아학회지, 30(2), 69-87. <http://dx.doi.org/10.14699/kbiblia.2019.30.2.069>
- Burrows, T. (2011, December 12-14). Sharing humanities data for e-research: conceptual and technical issues. Sustainable data from digital research: Humanities perspectives on digital scholarship. Proceedings of the conference held at the University of Melbourne, 177-192.
- de Waard, A., Breure, L., Kircz, J., & Oostendorp, H. (2006). Modeling Rhetoric in Scientific

- Publications. International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006.
- de Waarde, A. (2007). A Pragmatic Structure for Research Articles. In: 2nd International Conference on the Pragmatic Web, October, 2007. Tilburg, Netherlands, 22-23.
- 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, 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.
- Stigler, J. H. & Steiner, E. (2018). GAMS-eine infrastruktur zur langzeitarchivierung und publikation geisteswissenschaftlicher forschungsdaten. Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen und Bibliothekare, 71(1), 207-216.
- Superceanu, R. (1998) The rhetoric of scientific articles: Agenere study. Timisoara, Rumania: Orizonturi Universitre.
- van Dijk, T. A. (1980). Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction, and Cognition. Hillsdale, NJ: Lawrence Erlbaum.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Cho, J. & Park, J. D. (2019). A study on analysis of research data repository in humanities and social sciences. Journal of the Korean BIBLIA Society for library and Information Science, 30(2), 69-87. <http://dx.doi.org/10.14699/kbiblia.2019.30.2.069>
- Ko, Y. M. & Song, I. S. (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>
- Ko, Y. M., Kim, B. Y., Min, H. R., Song, M. S., & Lee, S. J. (2016). Ontology Construction Based on Structural Academic Terminology. Seoul: KLA.
- Shim, W. S., Ahn, H. Y., & Byun, J. Y. (2015). Strategies for improving the collection and use of research data in the humanities. Journal of the Korean Society for Library and Information Science, 49(3), 155-183. <http://dx.doi.org/10.4275/KSLIS.2015.49.3.155>

- Shin, Y. R. & Chung, Y. K. (2012). A study on the improvement plans of the humanities and social sciences research data archives in Korea. *Journal of Korean Society of Archives and Records Management*, 12(3), 93-115.
- Song, M. S. & Ko, Y. M. (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>
- Song, M. S. (2015). A study on the intellectual structure analysis in Korean studies. *Journal of the Korean Society for Library and Information Science*, 49(4), 125-157. <http://dx.doi.org/10.4275/KSLIS.2015.49.4.125>
- Song, M. S., Ko, Y. M., & Lee, S. J. (2016). A study on developing a metadata search system based on the text structure of Korean studies research articles. *Journal of the Korean Society for Information Management*, 33(3), 155-176. <http://dx.doi.org/10.3743/KOSIM.2016.33.3.155>
- Yoo, S. R., Lee, H. W., & Song, I. S. (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>

[부록 1] STNet 개념범주를 기반으로 한 인문학 분야 연구논문 키워드 유형 분류 기준

No	클래스 분류	클래스 설명
1	각국어	- STNet 'E03 언어/각국어' 개념범주의 하위 범주 - 각 국가에서 사용하는 언어(영어, 일본어 등) 및 과거에 사용된 언어(조선어, 중고한어 등), 지역어(중국어방언 등) 포함
2	감각	- STNet 'C01 특성(성질)' 개념범주의 하위 범주 - 오감을 통한 자극과 관련된 주제어(감각, 웃음 등) 포함
3	감정	- STNet 'C02 심리' 개념범주의 하위 범주 - 어떤 현상이나 일에 대하여 일어나는 마음(관용, 교감 등)이나 느끼는 기분(분노, 불안 등)에 대한 주제어 포함
4	개념(정의)	- STNet 'D01 이론(사상, 이념, 주의, 법칙)' 개념범주의 하위 범주 - '이론/사상'에 미치지 못하는 이론·사상적 하위 개념(상호주관성, 이상 등) 포함
5	건축물(시설물)	- STNet 'A04 인공물' 개념범주의 하위 범주 - 건축물이나 시설물의 일반 명칭(아파트, 도성 등) 포함
6	공간	- STNet 1단계 개념범주 - 인위적 공간(염전, 호마목장 등), 자연적 공간(토지, 서해안 등), 관념적 공간(식민도시, 유토피아 등)을 지칭하는 주제어 포함
7	관계/상호작용	- STNet 1단계 개념범주, 기원/유래, 비교/차이, 계층/등급, 인과, 상호작용을 하위 개념범주로 포함 - 인간의 활동 중, 관계적 의미가 중심인 주제어 포함
8	기관/단체명	- STNet 1단계 개념범주, 조직/단체명, 정부(왕조)명, 학파/종파명, 회의명을 하위 개념범주로 포함 - 국내외의 모든 기관·단체명(천주교, 신감각파 등), 모든 역사적 기관(조선총독부 등) 및 단체명(흥사단, 조선의용군 등) 포함
9	기관/조직	- STNet 1단계 개념범주, 행정/공공기관, 교육기관, 기업/회사, 사회(종교)조직/단체를 하위 개념범주로 포함 - 기관단체명의 고유한 명칭이 아니라, 조직과 집단의 유형이나 성질을 나타내는 주제어(교회, 학교, 의회 등) 포함
10	기념물명(문화재)	- STNet 1단계 개념범주 - 기념물(문화재)에 대한 주제어 포함
11	기법/전략	- STNet 1단계 개념범주, 기법/방식, 평가/분석, 교수/학습법, 전략/전술을 하위 개념범주로 포함 - 사회현상과 활동을 분석하는 방법/기법(절대평가, 대조분석 등), 전략/전술(번역전략, 서사 전략 등) 및 교수/학습방법(협동학습, 학습자중심교육 등)관련 주제어 포함
12	도구/수단	- STNet 'A04 인공물' 개념범주의 하위 범주 - 연장, 매체, 미디어, 인터넷, 커뮤니케이션 수단, 무기, 약기류 등 주제어 포함 - 구체적인 도구 및 수단의 명칭은 포함하지 않음
13	모형/기준	- STNet 1단계 개념범주, 모델/모형, 패턴, 기준/규칙, 표준/규격, 인프라/구조/범위, 기호/신호를 하위 개념범주로 포함 - 사회현상 및 활동을 분석하기 위한 모델, 패턴, 기준, 원칙, 표준, 규격, 인프라, 구조 등의 주제어 포함
14	문화/생활	- STNet 'C03 현상/이슈' 개념범주의 하위 범주 - 사회, 문화, 교육, 생활 전반에서 나타나는 현상과 이슈(다문화, 장애의식, 전통 등)에 관한 주제어 포함

No	클래스 분류	클래스 설명
15	물품/제품/생산품	- STNet 'A04 인공물' 개념범주의 하위 범주 - 구체적인 도구 및 수단의 명칭 - 정보통신과 관련된 인공적 생산물(기초재료시스템 등)을 비롯하여 광범위한 생산물(철기, 로봇 등)에 관한 주제어 포함
16	방법	- STNet 1단계 개념범주 - 연구방법(질적연구 등), 조사방법(의식조사 등), 분석기법(통계분석 등) 등과 관련된 주제어 포함
17	법률/제도명	- STNet 1단계 개념범주 - 국내의 과거와 현재의 헌법, 법률, 법령, 시행령, 조례, 조약/협정, 규칙명 등 포함 - 대일평화조약, 수능 등
18	변화(변동)	- STNet 1단계 개념범주. 완화, 강화, 개혁, 변천/과정, 분해/합체를 하위 개념범주로 포함 - 인간과 사물의 행위와 활동 중 변화 및 변동에 중심적 의미를 내포하는 주제어(대체, 발전, 변용, 상태변화 등) 포함
19	분과학문	- STNet 'D01 이론(사상, 이념, 주의, 법칙)' 개념범주의 하위 범주 - 분과학문, 학문의 하위 및 연구 분야, 주제영역 분야에 관한 주제어 포함
20	사건명	- STNet 1단계 개념범주. 사건명/회담명, 국경일/기념일명, 행사/축제명, 수상명 포함 - 전쟁을 포함한 역사적·사회적·문화적 사건명, 회담명 및 집회명, 국경일, 기념일, 상(賞), 행사, 축제명 등 포함
21	상태/상황	- STNet 'C03 현상/이슈' 개념범주의 하위 범주 - 사회 및 현상의 상태 및 상황에 대한 광범위한 주제어 포함 - 소외, 상실, 단절, 인식양태, 근대화, 내면화 등
22	상품명	- STNet 1단계 개념범주. 도구/수단명, 제품/브랜드명, 건축물/시설물명 포함 - 인간 활동에 의해 제작된 모든 인공물, 건축물, 시설물의 고유 명칭 포함
23	수준/정도	- STNet 'C01 특성(성질)' 개념범주의 하위 범주 - 인간 및 사회적 현상에 적용 가능한 표준, 정도, 분량 등에 관한 주제어 포함 - 가독성, 강도, 길이, 깊이 등
24	시대/시간	- STNet 1단계 개념범주. 시대, 시간, 시점 포함 - 일정한 시점, 기간, 역사적 시대 등을 나타내는 주제어 포함
25	신체	- STNet 1단계 개념범주. 신체의 부분, 신체 물질, 신체의 장애와 질병 포함 - 인간 및 동물의 신체와 관련된 주제어 및 장애, 질병에 관련된 주제어 포함
26	언어/문자	- STNet 'E03 언어/각국어' 개념범주의 하위 범주 - 언어, 언어학 및 문자와 관련된 주제어, 동사, 방언, 문자, 한자 관련 주제어 포함
27	유형/양식/장르	- STNet 1단계 개념범주. 문학장르, 음악장르, 미술과 디자인 양식, 스포츠와 놀이유형, 공연예술 포함 - 예술, 체육 분야의 특정 유형이나 장르와 관련된 주제어 포함
28	이론(사상/이념/주의/법칙)	- STNet 1단계 개념범주. 이론과 사상, 원칙 및 법칙 포함 - 사회현상이나 활동을 분석하고 이해하는 관점을 제공하는 이론, 사상, 이념, 법칙, 주의 등을 지칭하는 주제어에 적용
29	인간(생물학적특성)	- STNet 1단계 개념범주. 인간의 생물학적 특성 포함 - 인간을 지칭하는 일반 용어 및 인간의 생물학적 특성을 의미하는 주제어 포함
30	인간(관계)	- STNet 'A01 인간' 개념범주의 하위 범주. 친족관계 및 대인관계 포함 - 의미상 친족이나 대인 관계를 나타내는 주제어 포함
31	인간(능력/성향)	- STNet 'A01 인간' 개념범주의 하위 범주. 인간의 재능, 장애, 질병, 성향 포함 - 능력의 유무와 정도나 성향으로 인간의 유형을 구분하는 주제어 포함
32	인간(사회집단)	- STNet 'A01 인간' 개념범주의 하위 범주. 민족, 인종, 종족, 세대, 국민, 공동체, 성씨, 거주상황, 사회계층 포함 - 국내의 현재와 과거의 인간의 집합체(동성마을, 탈북고려인, 민족공동체 등) 나타내는 주제어 포함

No	클래스 분류	클래스 설명
33	인간(직업/지위/역할)	- STNet 'A01 인간' 개념범주의 하위 범주. 직업, 지위, 벼슬, 역할 포함 - 국내외 현재와 과거 인간의 직업, 지위, 역할 등을 나타내는 주제어(무대작가, 사진, 예비교사 등) 포함
34	인공물	- STNet 1단계 개념범주. 재료 및 부품, 의류, 식료품, 교통수단 포함 - 특정 제품명이나 도구명이 아닌, 이를 통칭하는 유형이나 성질을 의미하는 주제어 포함
35	인명	- STNet 1단계 개념범주. 실존인물과 가상인물 포함 - 국내외의 역사적 인물, 현존 인물, 예술작품 속의 인물, 가상적 인물 포함
36	인식/의식	- STNet 'C02 심리' 개념범주의 하위 범주 - 자의식, 초인지, 역사 인식, 작가 의식, 죽음 의식, 잠재 의식, 공동체 의식, 이해와 관련된 주제어 포함
37	자연물	- STNet 1단계 개념범주. 동물, 식물, 자연물, 광물 포함 - 인간을 제외한 동물, 식물 등 생물과 자연물, 광물 등 무생물에 대한 주제어
38	저작(창작)물/정보	- STNet 'A04 인공물' 개념범주의 하위 범주 - 가상의 인공물로서 영화, 스포츠, 음악, 미술, 레크레이션과 같이 인간의 창작물 및 정보를 지칭하는 일반적인 주제어 포함 - 구체적인 저작물과 창작물의 명칭 제외
39	제도(체제)	- STNet 1단계 개념범주. 사회제도, 정치·법률제도, 경제경영제도 포함 - 경제제도, 정치제도 등을 포함한 사회제도나 양식을 나타내는 주제어 포함 - 법, 시행령, 규칙 등의 구체적인 명칭은 제외함
40	준인간	- STNet 'A01 인간' 개념범주의 하위 범주 - 초인간적 가상 존재에 관한 주제어(괴물, 마녀, 요괴 등) 포함 - 구체적인 인물 명칭 제외
41	지명	- STNet 1단계 개념범주. 대륙명, 반도명, 국가명, 행정구역명, 지형지물명, 천체명 포함 - 국내외 과거와 현재의 구체적인 지역 및 지형지물 명칭 포함
42	창작물명	- STNet 1단계 개념범주. 문헌명, 예술작품명, 신문 및 잡지명, 방송 프로그램명, 지도명, 문서명 포함 - 국내외 과거와 현재 제작된 인간의 정신적 창작물(저작물)과 관련된 주제어 포함
43	특성(성질)	- STNet 1단계 개념범주. 경향/동향, 기질/품질/형질/성향, 능력/힘/에너지, 분포, 환경 포함 - 사회적 현상 및 대상의 분포, 성향, 이미지, 경향, 동향, 기질, 품질, 능력(힘, 에너지) 등의 주제어 포함
44	행위/행동/역할	- STNet 1단계 개념범주. 행위/활동, 교육활동, 경제/산업활동, 위법행위, 신체운동/행동, 기능/역할 포함 - 인간, 기관, 조직 등이 어떠한 목적을 달성하기 위해 취하는 모든 행위와 활동과 관련한 주제어 포함
45	현상/이슈	- STNet 1단계 개념범주. 격차/차이, 경영/경제/무역, 정치/국제 포함 - 인간활동에서 발생하는 모든 사회현상과 상태를 나타내는 주제어 포함 - 상황, 행태, 격차, 차이, 언어현상, 사회적 행위 등의 주제어 포함
46	기타주제어	- STNet 1단계 개념범주 - 다른 클래스에 적용되지 않는 용어

[부록 2] 인문학 분야 연구데이터를 위한 연구논문 의미구조기반 메타데이터

인문학 분야 연구데이터를 위한 연구논문 의미구조기반 메타데이터 적합 요소				
1	요소명	ResearchObjectives	표시상수	연구 목적
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 기술되어 있는 연구 배경과 기획 계기, 목적 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		
2	요소명	ResearchSubject	표시상수	연구 대상
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문에서 분석 및 탐구의 대상으로 하는 객체 • 'STNet'과 'KRM 용어사전'에 구축된 학술용어로 명시 • 요소 반복 가능 		
3	요소명	Personage/AffiliationBackground	표시상수	연구 대상의 배경이 되는 인물 및 단체
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문에서 연구 대상을 분석하기 위하여 함께 조사 및 연구한 인물 및 단체 정보 • 하위속성에 부합하는 방식으로 요소 정보 제시 • 요소 반복 가능 		
	하위속성	<ul style="list-style-type: none"> • BackgroundKeyword: 연구 대상의 배경을 'STNet'과 'KRM 용어사전'에 구축된 학술용어로 명시, 속성 반복 가능 • BackgroundRelation: 연구 대상과의 관계를 자유로운 형식으로 기술, 속성 반복 불가 		
4	요소명	ResearchMethod	표시상수	연구 방법
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 연구자가 구체적으로 명시한 연구 방법 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		
5	요소명	ResearchOutcome	표시상수	연구 결과
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 기술되어 있는 연구 결과 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		
6	요소명	ResearchSignificance	표시상수	연구 의의
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 기술되어 있는 연구 의의와 영향 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		
7	요소명	DataScale	표시상수	데이터 수집 규모
	요소설명	<ul style="list-style-type: none"> • 연구논문에 기술된 데이터셋의 생산·수집 규모 및 범위 • 연구를 위해 분석한 문헌의 건수, 설문 대상자 수, 참여관찰 대상 기관 수 등 • 수치와 단위를 연구논문에 작성된 내용을 기반으로 기술 • 요소 반복 가능 		

인문학 분야 연구데이터를 위한 연구논문 의미구조기반 메타데이터 사용 가능 요소				
1	요소명	ResearchTypes	표시상수	연구 유형
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구에서 사용한 연구 방법, 분석 방법, 자료 수집 방법 등에 따른 연구 유형 구분 • 제시된 유형 중 1개 이상 선택: 질적연구, 양적연구, 문헌연구, 조사연구, 사례연구, 내용분석, 메타분석, 실험연구 • 요소 반복 가능 		
2	요소명	ResearchProblem	표시상수	연구 문제
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 기술되어 있는 연구 문제 혹은 질문 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		

3	요소명	Hypothesis	표시상수	연구 가설
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 제시된 연구 가설, 전제 사항 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		
4	요소명	MethodOfDataCollection	표시상수	데이터 수집 방법
	요소설명	<ul style="list-style-type: none"> • 연구에서 해당 데이터셋의 생산·수집을 위해 사용한 방법을 자유로운 형식으로 기술 • 요소 반복 가능 		
5	요소명	SpatialBackground	표시상수	연구 대상의 배경이 되는 지리적 범위
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문에서 연구 대상을 분석하기 위하여 함께 조사 및 연구한 지리적 범위 정보 • 하위속성에 부합하는 방식으로 요소 정보 제시 • 요소 반복 가능 		
	하위속성	<ul style="list-style-type: none"> • BackgroundKeyword: 연구 대상의 배경을 'STNet'과 'KRM 용어사전'에 구축된 학술용어로 명시, 속성 반복 가능 • BackgroundRelation: 연구 대상과의 관계를 자유로운 형식으로 기술, 속성 반복 불가 		
6	요소명	TemporalBackground	표시상수	연구 대상의 배경이 되는 시간적 범위 및 사건
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문에서 연구 대상을 분석하기 위하여 함께 조사 및 연구한 시간적 범위, 사건 정보 • 하위속성에 부합하는 방식으로 요소 정보 제시 • 요소 반복 가능 		
	하위속성	<ul style="list-style-type: none"> • BackgroundKeyword: 연구 대상의 배경을 'STNet'과 'KRM 용어사전'에 구축된 학술용어로 명시, 속성 반복 가능 • BackgroundRelation: 연구 대상과의 관계를 자유로운 형식으로 기술, 속성 반복 불가 		
7	요소명	AnalysisMethods	표시상수	데이터 분석 방법
	요소설명	<ul style="list-style-type: none"> • 연구에서 해당 데이터셋의 분석을 위해 사용한 방법을 자유로운 형식으로 기술(통계 방법, 데이터 분석 방법 등) • 분석방법론을 비롯하여 구체적인 통계 분석 방법, 데이터 분석 방법 기술 • 요소 반복 가능 		
8	요소명	RelatedModels	표시상수	데이터셋 관련 연구모델
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구를 진행하기 위하여 다른 연구에서 차용한 이론이나 개념, 모형 • 하위속성에 부합하는 방식으로 요소 정보 제시 • 요소 반복 가능 		
	하위속성	<ul style="list-style-type: none"> • ModelsRelation: 연구에서 이론(개념, 모형)이 데이터셋에서 차용된 방식을 자유로운 형식으로 기술 • ModelsKeyword: 연구에서 차용한 이론(개념, 모형)을 'STNet'과 'KRM 용어사전'에 구축된 학술용어로 명시 		
9	요소명	ResearchLimitation	표시상수	연구 제한점
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문의 연구 범위 설정 및 이로 인한 연구의 제한사항 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		
10	요소명	FurtherResearch	표시상수	후속연구
	요소설명	<ul style="list-style-type: none"> • 데이터셋이 생산된 연구 논문 내에 기술되어 있는 후속연구 관련 내용 • 논문에 사용된 문장을 기반으로 기재 • 요소 반복 불가 		