

토픽 모델링을 활용한 한국불교 연구 동향 분석 - 2008~2022년 불교 관련 학술지 게재 논문을 중심으로

박종향(제1주저자)

동국대학교 K학술혁신연구소 연구초빙교수

jonghyangpark8@gmail.com

김은영(제2주저자)

동국대학교 K학술혁신연구소 연구초빙교수

tara@dgu.ac.kr

김용태(교신저자)

동국대학교 불교학술원 HK교수

yotai@dongguk.edu

I. 서론

II. 이론적 배경: 선행연구의 토픽 모델링

III. 연구 방법: 절차, 대상, 분석법

IV. 분석 결과: 논문 현황과 키워드,

시기별 토픽

V. 결론

요약문

본 연구는 최근 15년 동안(2008-2022) ‘한국불교’를 주제로 나온 국내 연구논문을 대상으로 토픽 모델링 기법을 활용하여 빅데이터를 분석하고, 한국불교 연구의

- 이 논문은 2021년 대한민국 교육부와 한국학중앙연구원(한국학진흥사업단)을 통해 K학술혁신연구소사업의 지원을 받아 수행된 연구임(AKS-2021-KDA-1250007).

동향을 파악하여 그 시의성을 도출하였다. 한국 불교학의 대표적 전문 학술지인 『한국불교학』, 『불교학보』, 『불교학연구』를 중심으로 한국불교와 관련된 논문 483편의 초록을 텍스트 마이닝(text-mining) 방식을 사용해 빅데이터로 정리하고 검토하였다. LDA(잠재 디리클레 할당) 기반 토픽 모델링을 활용한 결과, ‘불교 사회화와 대중화’, ‘불교 수행과 신앙 전통’, ‘불교학 연구와 방법론’, ‘원효와 한국 불교철학’, ‘선의 역사와 근대의 선승’, ‘동아시아 속의 한국불교’, ‘신라와 고려의 불교 사상’, ‘사회 복지와 불교적 실천’, ‘불교 교단과 종파의 전개’의 9개 토픽을 추출하였다. 이 토픽들은 한국 불교학의 주요 분야와 주제들을 시대별로 포괄하고 있으며 2000년대부터 꾸준히 재제되는 모습이 확인된다. 시기별로는 특별 기념 학술대회의 개최에 따라 ‘불교학 연구와 방법론’, ‘불교 사회화와 대중화’ 토픽의 논문 수가 급증하기도 했다. 한편 ‘사회 복지와 불교적 실천’과 같은 불교의 사회적 역할과 관련된 주제를 다룬 연구는 코로나 유행 시기인 2020년부터 축소되는 양상을 보인다. 본 연구는 이러한 연구 결과를 토대로 한국불교의 주제별 연구 동향을 고찰하였고, 향후 이를 바탕으로 다양한 관점에서 한국 불교학의 미래를 논의할 수 있기를 기대한다.

주제어

한국불교, 토픽 모델링, 빅데이터, 텍스트 마이닝, 잠재 디리클레 할당, 『한국불교학』, 『불교학보』, 『불교학연구』

I. 서론

한국의 불교학은 역사, 철학, 종교를 넘어서 사회·경제, 문화·예술, 문학 등 다양한 분야에서 연구가 이루어져 왔고 학제 간 연구에서도 성과가 도출되고 있다. 불교학을 연구하는 방법도 문헌학적(philology) 방법인 텍스트 중심의 해석학적 연구나 제도, 인물, 사상을 다루는 역사학·철학적 접근에만 머물지 않고 미시적인 문화학(culture studies) 연구, 그리고 사회과학 분야에서 주로 활용해온 수량적 연구에 이르기까지 매우 다양해졌다. 여러 분야의 다각적 접근을 통해 한국에서의 불교학 연구는 질적·양적 도약을 이루어왔다. 이러한 연구의 결과물은 불교학 전문 학술지에 발표되며 한국불교에 대한 체계적이고 폭넓은 이해를 가능케 했다.¹⁾ 국내 최초의 불교학 관련 전문 학술지는 『불교학보』

로서 1963년 발간된 이래 현재까지 100여 종이 출간되었다. 2022년만 해도 『불교학보』, 『한국불교학』, 『불교학연구』를 비롯하여 30여 개의 불교학 관련 학술지가 발행되었고, 여기에 400편 내외의 논문들이 게재되었다.

그동안 국내 불교학 연구의 동향에 대해서는 여러 학술지에서 지역, 시대, 분야, 주제 등으로 나누어 ‘회고와 전망’ 형태의 연구사 정리가 이루어졌다. 예를 들어, 『한국불교학』을 발간하는 한국불교학회는 2013년에 창립 40주년 기념으로 개최한 ‘결집, 한국불교학 40년: 그 연구성과와 과제’라는 학술대회에서 중국·티베트, 한국·일본, 인도·남방불교, 응용불교, 실천불교의 5개 분야로 나누어 기존의 연구성과를 검토하고 향후 연구의 방향성을 논의하였다. 그리고 같은 해에 『불교학보』에는 동국대 불교문화연구원 설립 50주년을 맞이하여 ‘한국불교의 연구 동향’을 비롯해 불교 교학, 불교 사학, 응용 불교 분야에서 그간의 성과를 정리한 논문들이 수록되었다. 하지만 주제별 연구 동향의 검토와 정리를 넘어서는, 기존 연구 결과에 대한 새로운 방법론적 접근과 구조적 분석은 이루어지지 않았다. ‘한국불교’를 주제로 기왕의 연구 동향을 검토한 논문도 ‘최근의 한국불교 연구 동향과 통불교 논의’,²⁾ ‘현대 한국불교의 동향과 과제’,³⁾ ‘불교학보를 통해 본 한국불교 연구 동향’⁴⁾ 등을 찾아볼 수 있을 뿐이다.

최근에는 여러 학문 분야의 연구 동향을 분석하는 방법으로 텍스트 마이닝(text-mining)을 활용하고 접목하는 것이 대세가 되고 있다. 그도 그럴 것이 세상에 존재하는 데이터의 80% 이상이 비정형 데이터로 추산되고, 비정형 데이터를 특정 목적에 맞게 유의미한 정보를 추출하는 분석과정이 바로 텍스트 마이닝이다. 이 방식은 1970년대부터 활성화되어 반세기 이상의 역사를 갖고 있다.⁵⁾ 텍스트 마이닝에는 단어 사이의 거리와 동시 출현 빈도 등을 바탕으로 단어기반의 네트워크를 표현하는 분석이나, 텍스트 데이터 내 몇 개의 잠복된 주

1) 민순의 2021 참조.

2) 최유진 2005 참조.

3) 강동균 2007 참조.

4) 이기운·조기룡·윤기엽 2013 참조.

5) 통계교육원 2017, https://sti.kostat.go.kr/window/2017b/html/2017_win_3.html (2024.01.08. 검색) 참조.

제를 찾는 토픽 모델링 기법 등이 포함된다. 또 논문, 도서, 학술자료 등 다양한 자료 유형을 종합적으로 분석하는 데 활용된다. 지금까지는 연구 동향을 분석하기 위하여 연구자가 수많은 연구 논저를 일일이 조사하고 검토하였는데, 이 방법은 오랜 시간이 소요될 뿐만 아니라 연구자의 주관적인 해석이 들어갈 가능성이 크다는 점에서 한계를 가진다. 반면 텍스트 마이닝은 방대한 분량의 데이터를 통계적·확률적 기법을 활용하여 객관적 문헌분석이 가능하다는 점에서 이 기법을 적용한 연구들이 활발히 이루어지고 있다. 그럼에도 그간 불교학 연구의 동향을 분석할 때는 한자, 고어 등에 대한 접근 문제 때문에 빅데이터를 활용한 이러한 사례는 거의 확인되지 않는다.

본 연구에서는 한국불교를 주제로 연구 동향 및 추세를 살펴보고 관련 연구에서 등장했던 주제(토픽)를 살펴본다. 다만 한국불교를 사상 및 특징, 맥락, 시기 등으로 구분하고 분류 유형에 따라 선행연구를 추려 그 경향을 살펴보는 일반적 연구와는 방법을 달리한다. 국내의 대표 학술지인 『한국불교학』, 『불교학보』, 『불교학연구』에서 한국불교와 관련된 논문들을 대상으로 빅데이터를 활용하여 연구 동향을 분석하고 이를 토픽 모델링 방식으로 유형화하여 살펴보려 한다. 특히 시기별로 연구의 주제별 경향이 어떻게 변화해 왔는지 그 특징을 검토하여 ‘연구의 시의성’을 고려해 볼 것이다. 본 연구의 문제의식은 다음과 같은 질문에서 시작한다. 첫째, 2008년부터 2022년까지 최근 15년 동안 한국불교 관련 연구에서 드러난 주제는 무엇이 있었는가? 둘째, 시기별 연구 주제의 변화 추이는 어떠한가? 이 질문들에 대한 답을 모색함으로써 한국불교 연구에서 다루었던 연구주제들과 그 시기별 변화 양상을 검토하고 이를 토대로 향후 연구의 전망을 논의하고자 한다.

II. 이론적 배경: 선행연구의 토픽 모델링

1. 한국불교 연구동향 선행연구

한국불교의 연구 동향을 검토하기 위해서는 무엇보다도 그 대상에 대한 정

의가 선행될 필요가 있다. 먼저 나름의 정의를 내리려는 시도가 있었는데, 그 중에 하나로 한국인들에 의한, 한국인들을 위한, 한국인들의 불교학을 한국불교학이라 명명하기도 한다.⁶⁾ 혹은 근대 이후 한국에서 근대적인 학문 방법론을 적용하여 불교를 대상으로 이루어진 연구를 한국불교학이라 정의하기도 한다.⁷⁾ 또 통불교 등과 같은 특성을 미리 전제하기보다 한국불교에 대한 연구들과 그에 따른 다양한 특징을 검토하다 보면 새로운 윤곽이 드러날 것이라고 보는 입장도 있다.⁸⁾

그런데 ‘한국불교학이란 무엇인가’라는 질문에 대한 정답은 없기에, 학문적 견해에 따른 다양한 의견들이 표출될 수 있다. 예를 들어 한국불교학을 ‘한국 출신의 연구자가 하는 불교학’으로 정의하면, 같은 시기 해외의 문헌학 동향과 한국의 문헌학 경향을 비교해보았을 때 유의미한 차이를 찾을 수 있을 것이다. 한국불교학을 ‘한국불교만을 다룬 연구’로 제한하더라도, 미국 내의 한국불교 연구도 포함될 수 있는 동시에 현대 한국에서의 남방불교 수행문화에 대한 연구도 배제할 수 없다. 이처럼 한국에서의 불교학 연구의 방향과 내용이 끊임없이 다양하고 풍부해지고 있어 규정하기가 점점 쉽지 않을 것이다.

앞서 언급한 바와 같이, 근대학문으로서 불교 연구가 시작된 이래 여러 잡지와 학술지를 통해 그 연구성과가 소개되어왔다. 학술지 논문의 주제들을 통해 한국불교학의 연구 동향을 검토한 글에서는 시기별로 다음과 같이 구분하고 있다. 근대 한국 불교학의 본격적인 시작과 체계화(1950~1970년대), 불교학 연구주제의 전문화와 양적 확대(1980~1990년대), 불교학 연구의 전문성 심화와 다변화(2000년대 이후)가 그것이다. 특히, 불교 연구의 발표의 장이 잡지에서 학술지로 중심이 옮겨가고, 더욱이 최근에 학술지는 한국연구재단의 평가와 지원 기준에 따라 운영되면서 학술지의 주제별 성격이 전문화되었다고 해석

6) 김원명 2021, 13.

7) 민순의 2021, 50.

8) 조은수 2004, 50-51.

한다. 다만 학술지의 질적·양적 확대는 발행 초기의 목적이나 특화된 성격을 유지하기 어렵게 만들었고 점차 종합학술지로 변모하는 양상을 보인다는 지적도 있다.⁹⁾ 그렇기에 현재로서는 학술지의 표제에 따른 성격보다는 개별 논문들의 연구 동향을 종합하여 분석하는 것이 한국의 불교학을 유형화하여 이해하는데 유용한 접근 방식이 될 것이다.

한국불교를 연구의 대상으로 하여 연구사, 방법론, 경향 등을 정리한 영어 논문도 발표되고 있다. 가장 최근의 주목할 만한 연구는 다음과 같다. 한국불교에 대한 연구주제가 삼국 등 고대에 치우쳐 있다가 조선으로 관심의 시기가 바뀌었고, 불교 전통과 관련한 현상까지도 다루어지며 새로운 연구와 접근 방식이 도입되고 있다는 평가가 나왔다.¹⁰⁾ 또한 2000년대 이후 북미 지역의 한국 불교 연구 중 전근대 시기에 해당하는 50여 편의 성과물 분석을 통해 비록 연구 편수가 증가하고 있지만 다른 분야에 비해 아직 확고한 영역을 형성하지 못했음을 지적하는 논문도 있다.¹¹⁾ 1868년 이후 150여 년간 프랑스에서 이루어진 한국불교 연구를 검토하여, 한국학의 성장, 연구논문과 학위논문의 현황, 관련 연구자들의 활동 등을 정리하면서 향후 한국불교학의 전망을 긍정적으로 바라보는 시각도 있다.¹²⁾

이상 살펴본 바와 같이 한국불교 연구는 역사, 철학, 종교, 문화, 문학 등 다양한 분야와 주제를 포괄하고 있다. 또 명상 수행과 같은 실천이나 응용 불교 분야도 최근 괄목할 만한 성장을 하고 있다. 그 결과 한국 불교학의 연구범위와 대상은 그 폭이 매우 넓으며 분석의 대상이 되는 연구 성과물의 양 또한 적지 않다. 그렇기에 토픽 모델링과 같은 새로운 방법론을 도입하여 한국불교에 대한 이해를 체계화하고 연구 경향을 구조적으로 분석할 필요가 있다.

9) 민순의 2021, 53-74.

10) Cho Eun-su 2022 참조.

11) Kim Seong Uk 2022 참조.

12) Bruneton 2022 참조.

2. 토픽 모델링(Topic Modeling)

토픽 모델링은 단어의 활용 및 맥락적 사용 패턴을 찾아내고, 유사한 패턴을 보이는 문서를 연결함으로써 이들 사이에 유용한 구조를 찾아내는 텍스트 마이닝 분석 방법 가운데 하나이다.¹³⁾ 토픽 모델링 가운데 흔히 사용되는 분석은 Blei, Ng, & Nordan(2003)의 잠재 디리클레 할당(Latent Dirichlet allocation: LDA)으로 다수의 문서에서 잠재적으로 의미 있는 토픽을 발견하는 절차적 확률분포 모델이다. LDA의 기본 개념은 <그림 1>에서와 같이 문서(document)들이 여러 개의 토픽(topics)으로 구성되고, 토픽은 복수의 단어들과 연관되며, 문서별 토픽 분포와 각 토픽별 단어들의 분포를 동시에 추정하여 숨겨진(latent) 변수인 토픽을 찾아내는 것이다.¹⁴⁾ 즉, 전체 문서 집합에서 확인되는 토픽 구조를 같은 문서에서 빈번하게 등장하는 단어들의 집합으로 간주함으로써 특정 토픽의 출현 정도에 따라 문서를 분류한다.¹⁵⁾ 이 원리에 따르면 토픽 모델링을 통해 문서들을 아우르는 주제의 내용을 파악하여 주요 주제별로 문서를 구분하여 분석할 수도 있지만, 연구자에 따라 문서를 아우르는 프레임으로 분류하기보다는 ‘이슈’ 또는 ‘주제’로 토픽을 주제 묶음화(군집화)하여 해설할 수도 있다.¹⁶⁾

토픽 모델링에서 하나의 문서는 여러 토픽을 포함하고 있고 여러 토픽이 동시에 대응되어 대량의 데이터에서 내재된 주제를 찾아내기 때문에 하나의 문서가 하나의 주제로 구분되는 기존 군집화 기법과는 차이가 있다.¹⁷⁾ 또한 통계적 추론을 통해 더 객관적이고 명확하게 연구주제를 분석할 수 있다는 점에서 연구자의 주관성이 최대한 배제되며, 기존 내용 분석보다 시간과 인력 투여에 대한 부담이 적다는 장점이 있다.¹⁸⁾ 그렇기에 토픽 모델링은 여러 문

13) 유주영 2022, 212.

14) Blei 2012, 79.

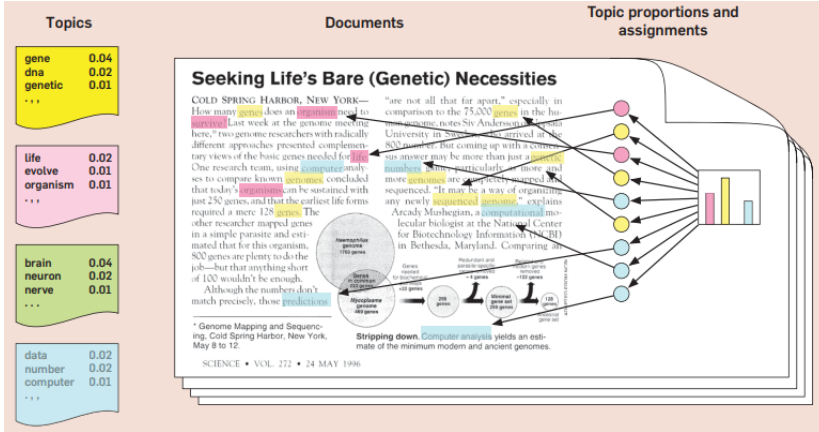
15) Blei 2012, 79.

16) 문안나 2020, 107.

17) 박준형·오효정 2017, 239.

18) 정영주·김혜진 2020, 105.

서에서 공통의 주제를 도출하고 이들을 추적하는 연구에 활용되어 여론 분석, 사회적 이슈 동향 파악, 특정 분야의 연구 동향 및 기술 동향을 파악하는 데 널리 사용된다.¹⁹⁾



〈그림 1〉 잠재 디리클레 할당에 따른 토픽 모델링²⁰⁾

토픽 모델링이 처음 언급된 국내 논문은 2003년으로, 2000년 이후 공학 계열에서는 70편, 사회과학 계열에서는 96편의 KCI 등재 논문에서 토픽 모델링이 활용되었다.²¹⁾ 반면 인문학에서는 50여 편 정도로 다른 분야에 비하여 아직 활성화되지는 않았다. 토픽 모델링을 활용한 인문학 연구주제로는 무형문화유산 분야,²²⁾ 문헌정보학,²³⁾ 한국어교육,²⁴⁾ 학교도서관, 독후감 텍스트 등이 대표적이다. 인문, 과학 및 기술, 사회과학, 학제간 연구 등 연구자들의 관심사를 기반으로 다양한 학문 분야에서 토픽 모델링이 사용되고 있음에도 한국불교학 분야에서는 토픽 모델링을 적용한 연구를 찾기 어려웠다.

19) 김남규 외 3인 2017, 477; 백평규·김남희 2021, 170.

20) Blei 2012, 78.

21) 최성철·박한우 2020, 819.

22) 오정심, 2020 참조.

23) 박자현·송민 2013 참조.

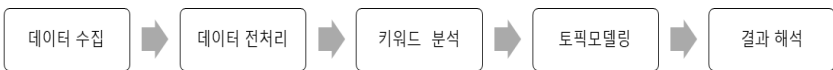
24) 김보영 2022 참조.

III. 연구 방법: 절차, 대상, 분석법

1. 연구 절차

본 연구는 학술연구 정보 서비스(RISS)를 활용하여 2023년 2월부터 자료 수집을 시작하였고 6월 이후 본격적인 분석을 시작하여 <그림 2>와 같이 한국불교를 주제로 한 연구 동향 분석을 진행하였다. 연구대상 논문의 주제어를 중심으로 수집하고, 데이터분석을 위해 데이터 전처리 작업을 시행하였다. 데이터 전처리 작업은 비정형 데이터를 분해하고 정형화된 구조로 만드는 작업으로 띄어쓰기, 오타자, 한자의 한글화 등 분석 대상, 데이터를 정제하는 작업과 유의어, 지정어, 제외어 등의 단어 정리를 거쳐 이루어졌다.

다음으로는 키워드 분석과 토픽 모델링을 한 후 이를 해석하였다. 데이터 분석은 넷마이너 프로그램(NetMiner 4.4)을 사용하였다. 텍스트 마이닝 및 토픽 모델링 분석에 사용되는 프로그램으로는 R, Python, UCINET, TEXTOM, NetMiner 등 다양하나 넷마이너 4.4는 단어 추출 시에 숫자나 대명사 같은 불용어(stop-word)가 자동 제거되기에 본 연구에서는 이를 활용하였다. 또한 주요 개념 파악을 위해 추출 단어의 품사는 ‘명사(noun)’로 지정하였고 연구자가 유의어(thesaurus), 제외어(exception list) 및 지정어(defined words) 사전들을 적용하여 분석에 사용할 의미형태소를 추출하였다.



<그림 2> 연구 절차

2. 분석 대상

분석 대상은 한국교육학술정보원의 RISS에서 ‘한국불교’를 키워드로 검색하여 ‘한국’, ‘불교’ 또는 ‘한국불교’가 제목, 목차, 초록, 키워드에 포함되는 모든 자료 가운데 2008년 1월 1일부터 2022년 12월 31일까지 발행된 연구논문을

수집하였다. 2008년 이전의 연구논문은 전자화된 초록(abstract)이 없어서 데이터분석 자체가 어려우므로 2008년 이후의 자료를 사용하였다. 또 ‘한국불교’를 키워드로 다양한 분야의 연구논문들을 조사 수집했기 때문에 한국불교 연구의 유형 및 특성을 파악할 수 있는 대표 학술지에 게재된 논문으로 한정하였다. 대표 학술지로는 『한국불교학』, 『불교학보』, 『불교학연구』를 선정하였다.

『불교학보』는 1963년, 『한국불교학』은 1975년, 『불교학연구』는 2000년에 창간하여 최소 20년 이상의 발행 역사가 축적되어 있고, KCI 등재지로서 학술지의 질적 수준을 유지해왔다. 또한 역사, 철학, 문학 등과 같은 전공 분야를 망라한 종합적 학술지로서 한국 불교학 영역에서 나름의 대표성을 가진다. 이처럼 이들 학술지는 오랜 기간 한국불교 연구의 성과를 고스란히 담아낸 학술적 보고로서 한국불교의 다양한 분야와 주제를 망라하고 있다. 따라서 한국불교를 키워드로 연구 동향을 분석하는데 적합한 조건과 기준을 갖추었다고 할 수 있다.

RISS의 서지정보 내보내기(export) 기능을 활용하여 엑셀 데이터로 추출한 뒤 저자, 논문명, 초록, 학회지명, 발행 연도를 중심으로 정리하였다. 수집된 데이터에서 중복된 논문 또는 서평, 한글 이외의 논문, 초록이 없는 논문을 걸러낸 결과 총 809건이 추출되었다. 이 자료에서 연구대상이 ‘한국불교’와 정확히 부합하지 않는 주제의 논문은 배제하였다. 즉 유교, 도교, 민속, 또는 중국불교, 일본불교의 경우처럼 한국불교 관련 연구라고 볼 수 없는 논문을 제외한 결과 483건이 최종 선정되었다.

3. 분석 방법

1) 텍스트 전처리

수집된 데이터는 정규화 과정을 거치지 않은 비정형 데이터로서 전처리 작업을 하였다. 넷마이너의 한국어 형태소분석기를 이용하여 비정형 텍스트 데이터를 분해하고 명사를 추출하였다. 첫째, 추출된 형태소들의 정확도를 높이기 위해 먼저 특수문자, 숫자, 영어 등은 삭제하였고 한문은 한글로 바꾸어 초록 데이터를 교정하였다.

둘째 ‘연구’, ‘보고’, ‘논고’, ‘논문’, ‘의의’, ‘의미’, ‘해석’, ‘결과’, ‘결론’, ‘전개’, ‘목적’, ‘때’, ‘다음’과 같이 논문초록에 자주 사용되는 단어들과 단어 중요도(Term Frequency-Inverse Document Frequency: TF-IDF)를 측정하여 값이 0.3 미만인 단어인 ‘한국’, ‘불교’, ‘부처’, ‘중심’은 제외하였다. TF-IDF 값은 모든 문서에서 사용될수록 0에 가깝고 소수의 문서에서 사용될수록 높은 수치로 나타나기 때문에 낮은 수치에 해당하는 단어를 제외하면 대부분 문서에서 사용하는 단어들을 배제하는 것이다.

또한 ‘뇌’, ‘각’, ‘계’, ‘결’, ‘끝’, ‘날’ 등처럼 한 글자로 큰 의미를 찾기 어려운 단어도 제외하였다. Query 기능을 이용하여 한 음절 글자를 제외할 수 있었지만, 한자가 자주 등장하는 본 연구에서는 한 글자만으로도 의미가 있는 단어들 이 존재하기 때문에 행(行), 체(體), 용(用), 심(心), 공(公), 선(禪), 의(義) 등과 같이 초록에 자주 등장하는 단어들의 경우 한 글자임에도 지정어로 등록하고, 그 외 한 음절의 단어들은 제외하였다.

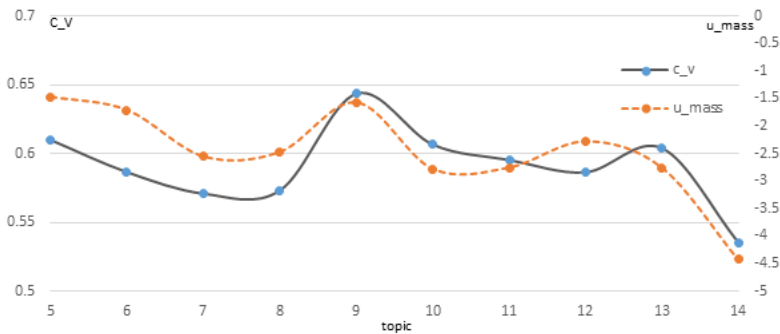
셋째, 뜻이 비슷한 말이나 축약어, 승려의 법명은 법호와 법명을 붙여 하나로 통일하였다. 예를 들면 ‘경허’, ‘경허선사’, ‘경허성우’, ‘성우’ 는 ‘경허 성우’ 로 하였다. 또 ‘부처’, ‘석가모니’, ‘석존’, ‘붓다’, ‘석가’, ‘불’, ‘佛’, ‘세존’은 ‘부처’로 대표어를 지정하였다. 넷째, 인명, 지명, 서명 등은 고유명사를 형태소 분석 없이 그대로 추출하도록 지정하였다. ‘팔도선교도총섭’, ‘대관정광진언’, ‘관불삼매해경’, ‘집일금광명경소’, ‘금동아미타여래좌상’, ‘가흥대장경’, ‘간화결의론’, ‘계초심학인문’, ‘관무량수경’ 등 단어들을 분리하지 않고 지정어로 등록하였다.

2) 데이터 분석

LDA(Latent Dirichlet allocation: 잠재 디리클레 할당) 기반 토픽 모델링은 토픽 수에 따라 결과가 상이하게 나타나기 때문에 연구자는 분석 목적에 적절한 토픽 수를 선정해야 한다. 연구자가 토픽 수를 결정하고 토픽 모델링 분석을 하면 추출된 주제에 해당하는 문서와 단어들의 분포, 이들의 관계를 통해 토픽

명을 도출한다. 만약 토픽 수를 너무 높게 설정하면 의미 없는 주제들이 추출될 수 있으며 너무 적게 설정하면 토픽에 상당히 많은 단어들이 포함되어 토픽 구분이 어려워질 수 있다. 본 연구에서는 매개변수 α 는 0.1, β 는 0.01,²⁵⁾ 반복 수행 횟수는 100회로 설정하고, 최적의 토픽 수는 타당성을 검증하는 방법을 사용하였다.

타당성을 검증하는 방법으로는 통계적인 방법과 해석적인 방법을 활용하였다. 통계적인 방법으로는 토픽 일관성(coherence)과 실루엣(silhouette) 계수를 참고하였다. 토픽 일관성은 각 주제에서 상위 단어를 뽑아 상위 단어 간의 공출현 유사도를 계산함으로써, 같은 주제에 의미적으로 유사한 단어들이 분류되는지 확인하는 방법으로 C_V 값은 1에 가까울수록, u_{mass}는 0에 가까울수록 응집성이 높은 것으로 판단하였다. 또한 K-means clustering을 이용한 실루엣 계수는 각 데이터가 얼마나 조밀하게 모여 있는지를 나타내는 값으로 1에 가까울수록 군집들이 적절히 분리되었다고 볼 수 있다(silhouette=0.83).²⁶⁾ 해석적인 방법으로는 토픽 일관성과 실루엣 계수 값을 참고하여 토픽 수를 변경해가며 주제 범주화가 적절하게 이루어졌는지 판단하였다. 이 과정에서 불교학 전공 연구자 2인이 참여하여 토픽 수를 확인한 후 정리하였고 논의를 거쳐 최종 9개로 결정하였다.



〈그림 3〉 토픽 일관성

25) 손태권·이광호 2020, 2006; Zhao, Chen and Zou 2015, 13.

26) 광희연·김숙향·이병인 2022, 124.

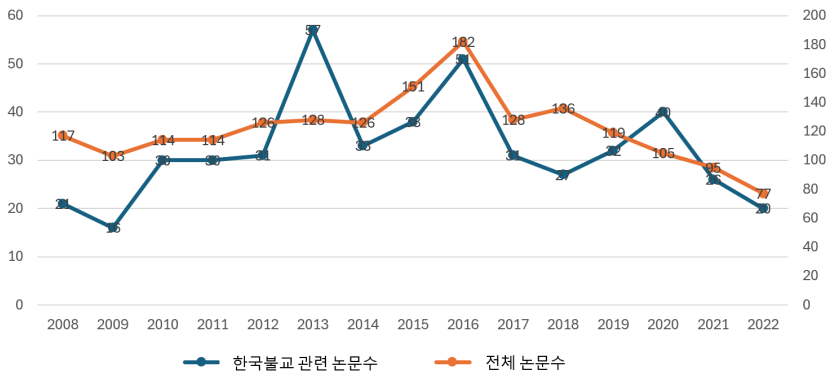
IV. 분석 결과: 논문 현황과 키워드, 시기별 토픽

1. 수집 논문 현황

2008년부터 2022년까지 ‘한국불교’를 주제로 『한국불교학』, 『불교학보』, 『불교학연구』 학술지에 게재된 논문 수는 총 483개(각각 384개, 90개, 11개)였다. 지난 15년 동안 이들 학술지에 게재된 논문들은 약 1,800여 건에 이르는데 이 가운데 한국불교를 주제로 한 연구의 비중은 약 28%였다. 시기별로는 다음 <표 1>과 같으며, 참고로 세 학술지에서 발간한 전체 논문의 수는 <그림 4>와 같다.

<표 1> 연도별 한국불교 관련 논문 게재 수

연도	2008	2009	2010	2011	2012	2013	2014	2015
논문수	21	16	30	30	31	57	33	38
연도	2016	2017	2018	2019	2020	2021	2022	total
논문수	51	31	27	32	40	26	20	483



<그림 4> 전체 논문 수와 한국불교 관련 논문 수

2. 토픽 모델링 결과

1) 주제어 빈도분석

483개의 초록에서 추출된 8,524개 단어 가운데 출현 빈도 분석으로 키워드

의 양상을 확인하였다. 또한 가장 많이 등장한 10개 단어는 <표 2>와 같다. 키워드의 순위에는 다소 차이가 있지만 빈도 수와 문서 등장 수의 상위 키워드는 유사한 것으로 나타나며, 공통적으로 가장 많은 빈도를 차지한 용어는 ‘사상’, ‘수행’, ‘사회’로, 이 키워드들은 한국불교에 대한 저자들의 공통적인 관점으로 확인된다. 한국불교는 철학, 종교, 역사, 문학을 다루는 인문학에 한정되지 않고 사람들의 행위, 일상 그리고 관계 등을 담아낸, 다시 말해 사회학적 관점에서의 유의미한 연구도 이루어졌다. <그림 5>는 주요 단어를 한눈에 파악할 수 있도록 시각화한 워드 클라우드(word cloud)로서 등장 빈도를 기준으로 단어의 크기가 크고 작게 표시된다.

<표 2> 주제어 출현 빈도

순위	빈도수		문서 동시 등장 수	
	단어	횟수	단어	횟수(문서 수)
1	사상	437	사상	162
2	사회	313	수행	140
3	수행	305	사회	121
4	종교	277	전통	117
5	조선	253	역사	109
6	중국	235	관계	100
7	원효	235	종교	97
8	교육	220	조선	96
9	전통	216	인식	96
10	신앙	216	중국	95



<그림 5> 워드 클라우드

2) 토픽별 키워드 추출

토픽 모델링 결과로 추출된 토픽 9개와 각 토픽의 핵심어는 <표 3>과 같으며 토픽별 네트워크는 <그림 4>와 같다. 산출된 토픽에 대한 주제명과 해석은 관련 분야 연구자의 전공 지식과 자료에 대한 이해가 필수적이므로,²⁷⁾ 불교학을 전공한 연구자가 토픽별 핵심어와 관련 문헌을 바탕으로 토픽명을 선정하였다.

토픽 1은 사회, 문화, 종교, 포교, 교육, 대중, 의식, 국가, 정치 등의 키워드로 <불교 사회화와 대중화>라는 토픽명을 정하였다. 대표논문인 「한국사회의 종교시장화와 한국불교의 포교 정향」(조기룡, 『불교학연구』, 2019/ 주제 적합도 72.7%)은 종교가 시장경제의 원리에 따라 선택 가능한 상품처럼 존재하는 현상인 종교 시장화를 비판적인 관점에서 바라보고 올바른 전법교화를 위한 포

<표 3> 토픽 모델링 결과

	주제명	핵심어	논문수(빈도)
1	불교 사회화와 대중화	종교, 교육, 사회, 의식, 문화, 국가, 포교, 한용운, 신앙, 사리, 사찰	54 (11.18%)
2	불교 수행과 신앙 전통	수행, 정토, 염불, 스님, 실천, 신앙, 사상, 현실, 중국, 세계, 대중, 교화	46 (9.52%)
3	불교학 연구와 방법론	불교학, 방법, 성과, 분야, 학문, 철학, 명상, 과제, 비교, 사상	77 (15.94%)
4	원효와 한국 불교철학	원효, 사상, 마음, 일심, 인식, 유식, 주장, 비판, 입장, 회통	63 (13.04%)
5	선의 역사와 근대의 선승	틴허 태성, 한암 중원, 오대산, 사상, 경허 성우, 신앙, 화엄, 신라, 자장, 중국	60 (12.42%)
6	동아시아 속의 한국불교	근대, 일본, 조선, 중국, 전통, 승려, 교사, 후기, 인식, 간행, 발전	46 (9.52%)
7	신라와 고려의 불교 사상	화엄, 신라, 경전, 의상, 고려, 문헌, 화엄경, 대승기신론, 주석서, 보살	39 (8.07%)
8	사회 복지와 불교적 실천	사회, 간화, 복지, 수행, 조사, 실천, 예술, 전승, 노인, 현대	40 (8.28%)
9	불교 교단과 종파의 전개	사찰, 운동, 활동, 조선, 승려, 비구니, 역사, 종단, 교단, 창건, 천태종, 조계종	58 (12.01%)

27) 임중현·오여진·안미리 2021, 1203; 김효선 2021을 참조할 것.

교의 당위성을 강조하는 내용이다. 사회, 종교, 포교, 대중 등 이 토픽의 주요 주제어와 깊은 관련이 있다. 이에 더하여 「국가 교육과정 변천과 불교종립학교 종교 교과서」(김은영, 『한국불교학』, 2014/ 주제 적합도 63.8%)는 학교에서의 종교교육에 대한 문제를 불교계 종립학교의 입장에서 재검토하고 바람직한 방향을 제시한 글이다. 이 또한 사회, 종교, 교육, 국가, 정치 등의 주제어를 다루고 있다. 이 토픽은 3대 종교로서 개신교, 가톨릭과 경쟁하는 불교의 사회적 역할 제고와 대중적 확산이 어느 시기이나 중요한 문제로 인식되었음을 잘 보여준다.

토픽 2를 설명하는 키워드는 수행, 실천, 신앙, 정토, 염불, 계율, 진언, 교화, 의식 등으로 <불교 수행과 신앙 전통>으로 토픽명을 정하였다. 대표논문으로는 「영명연수와 보조지눌의 유심정토와 타방정토」(서왕모, 『한국불교학』, 2018/ 주제 적합도 67.2%)와 「고려시대 법화삼매참법의 두 양상에 대한 비교 고찰: 백련사계와 묘련사계의 법화참법을 중심으로」(오지연, 『한국불교학』, 2022/ 주제 적합도 62.4%)를 들 수 있다. 전자는 중국 송대에 선과 염불 정토를 융합한 영명 연수와 고려의 대표 선승 보조 지눌의 유심정토, 타방정토 수용 문제를 비교 고찰한 논문이다. 이는 수행, 실천, 신앙, 정토, 염불 등 이 토픽의 핵심 키워드와 연계되어 있다. 후자는 고려 후기 천태종 계열 사찰인 백련사와 묘련사를 중심으로 행해진 법화참법의 양상을 비교 검토한 글이다. 이 또한 수행, 실천, 신앙, 계율, 교화 등 이 토픽의 주요 주제어를 포괄하고 있다. 이 토픽은 전근대 시기의 주요 신앙이자 수행 전통이었던 불교의 특성을 드러낸 것으로 연구자들의 관심이 이어지고 있다.

토픽 3은 불교학, 분야, 학문, 성과, 방법, 철학, 사상, 역사, 문헌, 번역 등의 키워드로 <불교학 연구와 방법론>으로 토픽명을 정하였다. 대표논문으로는 「한국 불교학 40년 연구 성과와 과제: 초기불교」(황순일, 『한국불교학』, 2013/ 주제 적합도 71.9%)와 「철학으로서 불교철학의 지형과 방법」(고영섭, 『한국불교학』, 2016/ 주제 적합도 68.2%)을 꼽을 수 있다. 앞의 논문은 현대 한국의 초기불교 연구사를 정리한 것으로서 1980년대 이후와 2000년대 이후로 나누

어 학계의 경향과 과제를 중점적으로 서술하였다. 뒤의 논문은 불교를 철학으로 바라보고 연구 지형과 방법을 검토한 것으로 문헌학적 방법론과 해석학적 방법론의 적용 가능성을 모색하였다. 이들 연구는 불교학, 분야, 학문, 성과, 방법, 철학, 사상, 문헌 등 이 토픽의 키워드들과 밀접히 연계되어 있다. 이 토픽은 한국에서 불교 연구가 수십 년 동안 축적되면서 확고한 기반을 다졌음을 보여주고 있다.

토픽 4는 <원효와 한국 불교철학>을 토픽명으로 하며, 키워드로는 원효, 사상, 마음, 일심, 인식, 유식, 회통, 통합, 논리 등이 있다. 대표논문은 원효 철학과 관련된 「원효 ‘판비량론’에서 ‘알아야식’의 의미」(김현희, 『한국불교학』, 2015/ 주제 적합도 91.7%), 동아시아의 관점에서 원효를 바라본 「동아시아 불교의 통합 담론 전개에서 원효의 역할: 지의의 원융사상과 원효의 화쟁사상의 연속성과 불연속성」(지혜경, 『한국불교학』, 2017/ 주제 적합도 77.3%)을 꼽을 수 있다. 첫 번째 논문은 원효의 『판비량론』에서 알아야식의 의미를 찾는 것으로 결론으로 여래장 일심을 제시하고 있다. 두 번째 논문은 동아시아의 통합 담론 전개라는 문제의식에서 중국의 천태 지의와 신라 원효 간의 사상적 연속성과 불연속성을 드러내 일심에 의한 원효 화쟁 사상의 의미를 조명했다. 이들 논문은 원효, 사상, 마음, 일심, 인식, 유식, 회통, 통합 등 대부분의 주제어와 관련되어 있다. 원효는 한국불교를 대표하는 사상가로서 이 토픽에서 그의 연구 사적 비중과 위상을 확인할 수 있다.

토픽 5의 키워드는 도의, 수행, 신앙, 전통, 경허 성우, 탄허 택성, 한암 중원 등으로 <선의 역사와 근대의 선승>이라고 주제명을 정하였다. 여기에는 근대의 오대산 출신 선승을 다루면서 오대산의 신앙 전통과 그 개조인 자장을 다룬 논문들도 포함되어 있다. 대표논문은 「경허의 오도가와 ‘기허주장자’ 시의 상관성」(홍현지, 『한국불교학』, 2014/ 주제 적합도 65%)으로 한국 근대 선의 중흥조인 경허 성우의 ‘오도가’와 ‘기허주장자’라는 시를 분석하여 선의 전법 계승의 중요성과 현실을 다룬 글로 수행, 전통, 경허 성우 등의 키워드를 담아내고 있다. 또 「신라 오대산의 문수신앙과 오만진신신앙 검토」(염중섭, 『한국불

교학』, 2019/ 주제 적합도 86.3%)처럼 오대산 문수신앙과 그 개창자인 자장을 다룬 논문도 몇 편 포함되어 있다. 이 토픽은 한국불교의 주류 전통인 선종의 역사와 경허 성우 등 근대의 선 계보를 대상으로 하며 최근에도 연구가 이어지고 있다. 흥미로운 점은 오대산 관련 연구가 적지 않게 포함된 사실인데 이는 특정 연구자가 해당 논문을 다수 발표했기 때문이다.

토픽 6의 토픽명은 <동아시아 속의 한국불교>로서 핵심 키워드는 전통, 근대, 중국, 일본, 조선, 민족, 승려, 수용, 발전 등이 있다. 대표논문은 「『선원제전집도서』의 간행과 유통판본 고찰: 중국과 한국 유통본을 중심으로」(서수정, 『한국불교학』, 2011/ 주제 적합도 70.8%)와 「한국근대불교 연구의 과제와 전망」(조성택, 『한국불교학』, 2012/ 주제 적합도 61.2%)을 들 수 있다. 전자는 종필의 『선원제전집도서』가 언제 어디서 어떻게 간행되었는지를 살펴본 후 현존본의 계보와 지형도를 그렸다. 후자는 근대불교의 개념을 정의하기 위한 가설을 제기한 것으로 불교는 한국사의 일부이자 불교사의 일부이지만 동아시아를 넘어 세계불교의 시각에서 바라보아야 한다고 제언했다. 이들 논문에는 전통과 근대, 중국과 일본, 수용과 발전 등의 키워드가 각각 적용된다. 이 토픽은 불교의 보편성과 한국의 특수성을 결합한 한국불교의 특성을 거시적으로 주목한 것으로 최근 15년간의 주목되는 학문적 경향이라 할 수 있다.

토픽 7은 <신라와 고려의 불교 사상>으로 신라, 화엄, 의상, 밀교, 고려, 경전, 문헌, 주석서, 대승기신론 등이 주요 키워드이다. 대표논문은 「신라 및 고려시대 화엄학 문헌의 성격과 내용」(최연식, 『불교학보』, 2011/ 주제 적합도 81.3%)으로 고대와 중세 한국 불교계에서 나온 10여 종 이상의 『화엄경』 관련 문헌들을 종합적으로 검토하였다. 한국불교의 사상적 특성을 보여주는 의상의 『법계도』를 비롯해 원효, 균여, 지눌의 저작 등과 신라 화엄학의 영향 아래 찬술된 책을 대상으로 신라와 고려, 화엄, 의상, 문헌, 주석서 등의 키워드를 다루었다. 이 밖에 「동아시아 불교문헌에 보이는 특이한 다라니의 형태: 석마하연론을 중심으로」(김지연, 『한국불교학』, 2021/ 주제 적합도 75.5%)는 『대승기신론』의 주석서이자 신라에서 찬술된 것이라고 하는 『석마하연론』의 다라니를 분

석하여 책의 성립 시기를 추정하였는데, 신라와 고려, 밀교, 문헌과 주석서, 대승기신론 등의 키워드가 이에 해당한다. 이 토픽은 교학의 전성기였던 신라와 고려의 불교 사상을 다룬 것으로 논문 수는 다른 주제에 비해 많은 편은 아니지만, 불교 교학 연구가 여전히 중요한 분야임을 볼 수 있다.

토픽 8 <사회 복지와 불교적 실천>의 핵심 키워드는 사회, 복지, 노인, 유아, 수행, 실천, 조사, 간화, 예술 등이다. 대표논문으로는 「서울시 소재 불교신자 가구의 특성과 사회복지욕구」(김노은, 『한국불교학』, 2012/ 주제 적합도 80%)와 「조사선과 간화선의 수행체계 비교」(신명희, 『한국불교학』, 2018/ 주제 적합도 62.3%)를 선정하였다. 앞의 논문은 서울의 불교 신자 가구의 특성을 살펴보고 그들의 복지 욕구를 파악하여 불교 사회 복지를 위한 기본 정보를 제공한 것으로 사회, 복지, 노인, 유아 등의 키워드를 포함하고 있다. 뒤의 논문은 당나라 때 성립한 조사선과 송대에 전개된 간화선을 비교하여 양자의 공통점과 차이점을 도출하고 현재적 의미를 찾은 것으로서 수행, 실천, 조사, 간화의 키워드를 공유하고 있다. 이 토픽은 불교의 이타행에 해당하는 사회 복지뿐 아니라 스스로의 깨달음을 위한 실천 수행을 모두 아우르고 있다.

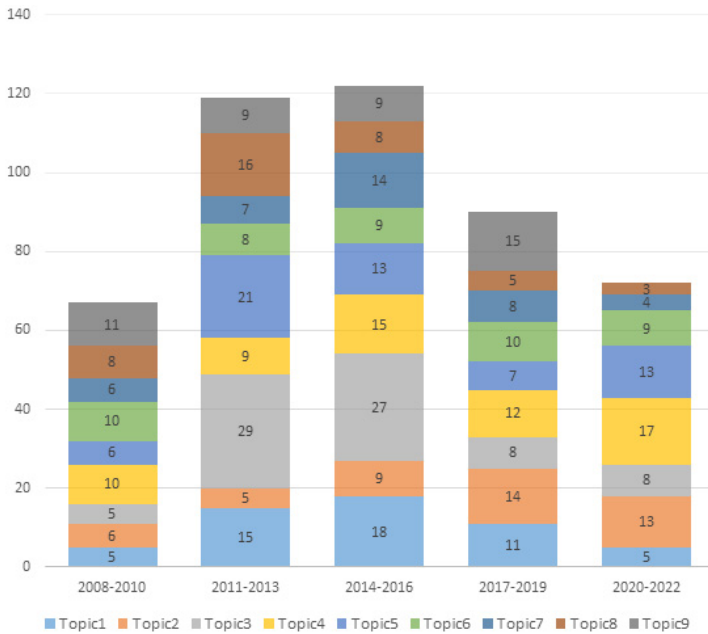
토픽 9는 <불교 교단과 종파의 전개>로서 사찰, 승려, 비구니, 운동, 역사, 종단, 교단, 조계종, 선종 등이 키워드이다. 대표논문으로는 「조선후기 경기 서북부 일대 비구니 사찰과 비구니의 활동」(탁효정, 『한국불교학』, 2022/ 주제 적합도 71.2%)과 「고려후기·조선초 천태종단의 존재양상 추이 및 동향- 주요 고승과 사찰을 중심으로」(황인규, 『한국불교학』, 2015/ 주제 적합도 53.4%)를 택하였다. 전자는 전등사본말사지에 수록된 사료를 토대로 조선 후기 경기 서북부 일대의 비구니 사찰을 조사하고 비구니의 활동을 분석한 논문으로, 사찰, 승려, 비구니, 운동, 역사, 교단 등의 키워드와 직접적 관련이 있다. 후자는 고려시대 천태종의 창설 이후 조선 초에 선종과 교종으로 통합되기까지 천태종단의 존재 양상 추이 및 동향을 살펴본 연구이다. 이에 부합하는 키워드는 사찰, 승려, 역사, 종단, 교단, 선종 등이다. 이 토픽은 불교사의 주요 주제인 교단과 종파의 역사적 전개를 다룬 것으로 연구가 계속 이어지고 있다.

3) 시기별 토픽 추이

한국불교에서 주요하게 다루어진 주제를 시간의 흐름에 따라 파악하기 위해 전체 기간의 LDA 분석 결과를 <그림 6>과 같이 토픽별로 확인하였다. 기간은 2008년부터 2022년까지를 3년 단위로 구분하여 추이를 살펴보았다. 매년 증가 또는 감소하는 양상이 일정한 경향을 보이지는 않지만 2013년, 2016년, 2020년을 기점으로 상승하는 모습이 나타났기에 3년을 주기로 추이를 분석하였다. 분석 결과 2011년부터 2013년에 해당하는 2주기와 2014년부터 2016년까지의 3주기에 한국불교 관련 논문이 다소 빈번하게 나왔고, 그 이후부터 2020년을 제외한 3주기에는 다소 축소되는 경향이 나타났다. 전체적인 논문의 양은 줄었지만 각 주제들은 한국불교 연구에서 일정 비율을 차지하며 비중 있게 다루어진 편이다.

주목할 점은 토픽 1 <불교 사회화와 대중화>와 토픽 3 <불교학 연구와 방법론>이 2011년부터 2016년까지 큰 폭의 상승세를 보였는데, 이러한 모습은 2013년 한국불교학회 창립 40주년을 맞이하여 불교학 연구의 현황과 성과를 논의하는 대규모 학술대회가 열렸기 때문에 이와 관련하여 한국 불교학을 주제로 한 논문들이 다수 게재되며 편수가 늘어난 것이다. 이후 토픽 1과 토픽 3의 비중은 2010년대 후반기부터는 축소되었고 특히 2021년과 2022년에는 토픽 1 주제의 논문이 확인되지 않았다. 또 토픽 9 <불교 교단과 종파의 전개> 관련 논문들도 3년에 10여 편의 수준을 유지하다가 2020년 이후 급격히 감소하는 경향을 보이는데, 2021년에는 1편이었고 2022년에는 확인되지 않는다.

한편 토픽 2 <불교 수행과 신앙 전통>, 토픽 4 <원효와 한국 불교철학>은 2010년대 후반기로 갈수록 점차 늘어나는 추세를 보였다. 불교사의 주요 주제에 해당하는 수행과 신앙, 그리고 원효 관련 연구는 꾸준히 진행되고 있음을 볼 수 있다. 그런데 응용불교학 분야로 분류할 수 있는 토픽 1 <불교 사회화 및 대중화>와 토픽 8 <사회 복지와 불교적 실천>은 특히 코로나 유행 시기인 2020년부터 2022년 사이에 크게 줄어드는 모습을 보여 주목된다.



〈그림 6〉 시기별 토픽 추이

V. 결론

지금까지 한국불교학 연구의 최근 동향을 검토하기 위해 텍스트 마이닝의 토픽 모델링 방법을 활용하여 분석하였고, 연구의 시의성을 확보하기 위해 ‘한국불교 관련 연구에서 드러난 주제’와 ‘시기별 연구주제별 변화’를 중심으로 살펴보았다. 연구는 ‘데이터 수집→ 전처리→ 키워드 분석→ 토픽 모델링→ 결과 해석’의 절차로 진행되었다. 분석 대상은 2008~2022년의 15년 동안 불교학 분야의 전문 학술지로서 대표성과 상징성을 지닌 『한국불교학』, 『불교학보』, 『불교학연구』에 실린 ‘한국불교’를 키워드로 하는 논문들이었다. 주제의 빈도분석 결과, <불교 사회화와 대중화>, <불교 수행과 신앙 전통>, <불교학 연구와 방법론>, <원효와 한국 불교철학>, <선의 역사와 근대의 선승>,

<동아시아 속의 한국불교>, <신라와 고려의 불교 사상>, <사회 복지와 불교적 실천>, <불교 교단과 종파의 전개>의 9개 토픽을 선정할 수 있었다.

해당 토픽들은 완전히 독립적이지 않으며 서로 연관되어 있는데 이 9가지 토픽은 크게 두 가지 관점으로 구분하여 해석할 수 있다. 토픽 1 ‘불교 사회화와 대중화’, 토픽 2 ‘불교 수행과 신앙 전통’, 그리고 토픽 8 ‘사회 복지와 불교적 실천’은 한국 사회 속에 나타난 불교의 다양한 현실적 모습을 보여주며 서로 연동된 주제라고 할 수 있다. 이에 비해 토픽 3 ‘불교학 연구와 방법론’, 토픽 4 ‘원효와 한국 불교철학’, 토픽 5 ‘선의 역사와 근대의 선승’, 토픽 6 ‘동아시아 속의 한국불교’, 토픽 7 ‘신라와 고려의 불교 사상’, 토픽 9 ‘불교 교단과 종파의 전개’는 한국불교의 역사와 사상 등에 관한 학술적 연구 성과를 담은 것이다.

여기에서 전자에 해당하는 사회와 대중 속의 불교, 수행과 실천 신앙으로서의 불교는 학술연구 주제로 구성된 후자에 비해 비중이 작고 시기적으로 점차 축소되는 경향을 보인다. 물론 토픽별로 발표가 집중되는 시기에는 조금씩 차이가 있는데, 이는 특별 기념 학술대회 개최 및 연구방법론과 주제의 변화 추이와 관련이 있다고 추정하였다.

본 연구의 한계 및 후속 연구를 위한 제언은 다음과 같다. 첫째, 한자 기반의 언어로 구성된 단어들 많았기 때문에 텍스트 분석에 일정한 한계가 있다. 한자로 구성된 복합어의 경우 그 의미를 텍스트로 전환하여 분석하는 데 어려움이 있었고, 한자 단어를 한글로 변환했을 때 한자의 원래 뜻을 충분히 살리지 못하는 경우가 발생하였다. 둘째, 토픽 모델링 기법은 기존의 주제별 연구사 정리에 비해 훨씬 많은 대상을 효율적으로 추출하여 분석할 수 있지만 추출한 토픽을 유형화하고 명칭을 붙일 때 전문성이 요구된다.²⁸⁾ 셋째, 본 연구는 디지털화된 초록 데이터를 이용하였기 때문에 최근 15년간의 연구만을 다루었고 그 이전의 논문은 대상에 포함하지 않아서 그에 대한 후속 작업이 필요하다.

28) 임종현 외 3인 2022, 205.

한국 불교학을 주제로 연구 동향을 살펴본 결과 한국 불교학의 특징과 정체성을 명확하게 찾아내기는 쉽지 않았다. 구조적 토픽 모델링만을 활용하여 특정 연구주제의 특징을 뽑아내는 것은 인문학 분야의 경우 일정한 한계를 가지며, 데이터 분석을 보다 전문적이고 체계적으로 행하고 비판적인 검토 과정이 수반되어야 한다. 본 연구는 한국불교라는 카테고리에서 어떠한 주제를 중심으로 연구가 되어왔는지 최근 15년 동안 축적된 한국 불교학 연구의 궤적을 그리며 9개 주요 토픽을 간추리고 그 시기별 경향과 특징을 도출해 본 것은 첫 시도였다. 이 연구를 통해 지금까지 한국 불교학계에서 한 번도 시도하지 않았던 토픽 모델링이라는 빅데이터 분석 기법을 적용한 데서 의의를 찾고자 하며, 향후 이를 바탕으로 다양한 관점에서 한국 불교학의 연구 동향을 전망할 수 있기를 기대한다.

참고 문헌 REFERENCES

- BAEK, Pyounggu 백평구 and KIM, Namhee 김남희. 2021. 「토픽 모델링을 활용한 학술지 ‘여성연구’의 연구동향 탐색」 (“Research Trends in ‘Women’s Studies’ Using Topic Modeling”), 『여성연구』 (*The Women’s Studies*), vol. 110, no. 3, 159-194.
- BLEI, David M. 2012. “Probabilistic Topic Models,” *Communications of the ACM*, vol. 55, no. 4, 77-84.
- BRUNETON, Yannick. 2022. “An Attempt to Construct and Characterize the History of Korean Buddhist Studies in France (1868-2022),” *International Journal of Buddhist Thought & Culture*, vol. 22, no. 2, 41-70.
- CHO, Eun su 조은수. 2004. 「통불교 담론을 중심으로 본 한국 불교사 인식」 (*“Perception of Korean Buddhist History Centered on the Discourse of *T’ongbulgyo*”), 『불교평론』 (*The Buddhist Review*), vol. 6, no. 4, 30-51.
- _____. 2022. “The Rise of New Studies and Approaches in Korean Buddhism,” *International Journal of Buddhist Thought & Culture*, vol. 22, no. 2, 23-39.
- CHOI, Sungchel 최성철 and PARK Han woo 박한우. 2020. 「토픽 모델링 연구동향 분석: 공학파 사회과학 분야 KCI등재지를 중심으로」 (“A Study on the Trend of Topic Modeling in South Korea Using KCI Journal Publications”), 『한국자료분석학회』 (*Journal of the Korean Data Analysis Society*), vol. 22, no. 2, 815-826.
- CHOI, Yu jin 최유진. 2005. 「종교연구동향: 최근의 한국불교 연구동향과 통불교 논의」 (*“Trends in Religious Studies: Recent Trends in Korean Buddhism Research and Discussions on *T’ongbulgyo*”), 『종교문화비평』 (*The Critical Review of Religion and Culture*), vol. 7, 260-290.
- GANG, Dong gyun 강동균. 2007. 「현대 한국불교의 동향과 과제」 (“The Trends and Challenges of Modern Korean Buddhism”), 『동아시아불교문화』 (*Journal of Eastern-Asia Buddhism and Culture*), vol. 1, 101-117.
- JUNG, Young ju 정영주 and KIM Hae jin 김혜진 2020. 「토픽모델링을 활용한 학교도서관 연구동향 분석」 (A Study on the School Library Research Trends Using Topic Modeling), 『한국도서관·정보학회지』 (*Journal of Korean Library and Information*

- Science Society*), vol. 51, no. 3, 103-121.
- KIM, Boyoung 김보영. 2022. 「토픽 모델링과 언어 네트워크 분석을 활용한 재중동포 대상 한국어교육 연구 동향 분석」 (“An Analysis of the Research Trends on Korean Language Education to Korean Chinese through Topic Modeling and Language Network Analysis”), 『외국어로서의 한국어교육』 (*Teaching Korean as a Foreign Language*), vol. 66, 1-23.
- KIM, Hyoseon 김효선. 2021. 「기업 아카이브에 관한 연구 동향 분석: 토픽 모델링 분석을 중심으로」 (“Analysis of the Research Trends on Business Archives: Focusing on the Topic Modeling Analysis”), 『한국기록관리학회』 (*Journal of Records Management & Archives Society of Korea*), vol. 21, no. 3, 163-186.
- KIM, Namgyu, et al. 김남규 외 3인. 2017. 「텍스트 분석 기술 및 활용 동향」 (“Investigations on Techniques and Applications of Text Analytics”), 『한국통신학회논문지』 (*The Journal of Korean Institute of Communications and Information Sciences*), vol. 42, no. 2, 471-492.
- KIM, Seong Uk. 2022. “Trends of Research on Pre-Modern Korean Buddhism in North America (2000-2021),” *International Journal of Buddhist Thought & Culture*, vol. 22, no. 2, 71-94.
- KIM, Wonmyoung 김원명. 2021. 「한국 불교학 어디까지 왔나」 (“What Has Been the Progress of Korean Buddhist Studies”) 『불교평론』 (*The Buddhist Review*), vol. 23, no. 1, 8-30.
- KWAK, Hee yeon 곽희연, KIMSook hyang 김숙향 and LEE Byong in 이병인. 2022. 「텍스트 마이닝을 활용한 특수교육대상영유아 가족 관련 연구 동향분석: 2003~2021년 국내 학술지 논문을 중심으로」 (“Analyzing Trends in Research on Families of Children with Special Needs Using Text Mining: Focusing on Domestic Academic Journal Articles from 2003 to 2021”), 『특수교육논총』 (*Journal of Special Education*), vol. 38, 110-139.
- LEE, Kiwoon 이기운, CHO Kiryong 조기룡 and YOON Kiyeop 윤기엽. 2012. 「佛敎學報(불교학보)를 통해 본 한국불교연구의 동향: 불교문화연구원 50년의 회고와 전망」 (“The Trend of Korean Buddhist Research from the Perspective of the *Bulgyohakbo*: 50 Years Retrospect and Prospect of the Institute for Buddhist Culture”), 『불교학

- 보』(*Journal of Buddhist Culture Research Institute*), vol. 63, 17-48.
- LIM, Jonghyun 임종현, OH Yeon jin 오여진 and AHN Mi lee 안미리. 2021. 「토픽 모델링을 활용한 국내 메이커교육 연구 동향 분석」 (“An Analysis of Research Trends on Maker Education Using LDA-based Topic Modeling”), 『교육정보미디어연구』(*Journal of Korean Association for Education Information and Media*), vol. 27, no. 3, 1189-1219.
- LIM, Jonghyun, et al. 임종현 외 3인. 2022. 「2010~2021년 국내 메타버스와 가상세계를 활용한 교육 연구 동향 분석: LDA 기반 토픽 모델링과 시계열 회귀분석을 적용하여」 (“Education Use of Metaverse and Virtual Worlds from 2010 to 2021: An Analysis of Research Trends Using LDA-based Topic Modeling and Time Series Regression Analysis”), 『교육정보미디어연구』(*Journal of Korean Association for Educational Information and Media*), vol. 28, no. 2, 187-214.
- MIN, Sun euy 민순의. 2021. 「학술지 논문 주제로 살펴본 연구경향」 (“Research Trends Examined Through Academic Journal Paper Topics”), 『불교평론』(*The Buddhist Review*), vol. 23, no. 1, 50-75.
- MOON, Anna 문안나 and LEE Shin Haeng 이신행. 2020. 「사회서비스원 정책 보도의 프레임 분석」 (“News Frames in the Coverage of Social Service Agency Policy”), 『한국광고홍보학회』(*Korean association of AD&PR*), vol. 22, no. 4, 100-134.
- OH, Jung shim 오정심. 2020. 「빅데이터 텍스트 마이닝을 통한 무형문화유산 분야 연구동향 및 지식체계 분석」 (“Research Trends and Knowledge System Analysis of the Academic Field of Intangible Cultural Heritage through Big Data Text-Mining”), 『무형유산』(*Korean Journal of Intangible Heritage*), no. 8, 93-127.
- PARK, Ja hyun 박자현 and SONG Min 송민. 2013. 「토픽 모델링을 활용한 국내 문헌정보학 연구동향 분석」 (“A Study on the Research Trends in Library & Information Science in Korea using Topic Modeling”), 『정보관리학회지』(*Korea Society for Information Management*), vol. 30, no. 1, 7-32.
- PARK, Jun hyeong 박준형 and OH Hyo jung 오효정. 2017. 「국내 기록관리학 연구동향 분석을 위한 토픽모델링 기법 비교 - LDA와 HDP를 중심으로」 (Comparison of Topic Modeling Methods for Analyzing Research Trends of Archive Management in Korea: Focusing on LDA and HDP), 『한국도서관·정보학회지』(*Journal of*

Korean Library and Information Science Society), vol. 48, no. 4, 235-258.

SON, Tae kwon 손태권 and LEE Kwang ho 이광호. 2020. 「토픽 모델링을 활용한 국내 수학과 교육과정 연구동향 분석: 1997년부터 2019년까지 게재된 국내 수학교육 학술지 논문을 중심으로」 (“An Analysis of Domestic Research Trends of Mathematics Curriculum Research through Topic Modeling: Focused on Domestic Journals Published from 1997 to 2019”), 『한국수학교육학회』 (*The Mathematical Education*), vol. 59, no. 3, 201-216.

Statistic Training Institute 통계교육원. 2017. ‘텍스트는 어떻게 분석되는가’ (*How is the Text Analyzed), 통계의 창 (*Window of Statistics), https://sti.kostat.go.kr/window/2017b/html/2017_win_3.html (2024.01.08. 검색).

ZHAO, Weizhong, James CHEN, and Wen ZOU. 2015. ‘Best Practices in Building Topic Models with LDA for Mining Regulatory Textual Documents’, *Dokumen.Tips*, <https://dokumen.tips/documents/best-practices-in-building-topic-models-with-lda-for-mining-regulatory-textual-2015-11-11.html?page=1> (2024.01.08. 검색).

A Topic Modeling Analysis of Trends in Research on Korean Buddhism in Papers Published in Academic Journals (2008-2022)

PARK, Jonghyang

Research Professor

Center for the Expansion of Academics on Korea, Dongguk University

KIM, Eunyoung

Research Professor

Center for the Expansion of Academics on Korea, Dongguk University

KIM, Yongtae

HK Professor

Academy of Buddhist Studies, Dongguk University

This research analyzes academic papers published on Korean Buddhism over the past 15 years (2008-2022). Topic modeling was used to identify trends in research on Korean Buddhism and analyze changes over time. The abstracts of 483 papers related to Korean Buddhism were organized and analyzed through a big-data approach, focusing on the *Journal of Buddhist Culture Research Institute*, the *Korean Journal of Buddhist Studies*, and the *Journal of the Korean Association for Buddhist Studies*, which are representative academic journals in Korea. The following nine topics were extracted through topic modeling based on latent Dirichlet allocation: 1) socialization and popularization of Buddhism, 2) Buddhist practice and religious tradition, 3) Buddhist research and methodology, 4) Wonhyo and Korean Buddhist philosophy, 5) history of Seon and modern Seon masters, 6) Korean Buddhism in East Asia, 7) Buddhist thought in Silla and Goryeo, 8) social welfare and Buddhist practice, and 9) the development of Buddhist orders and sects. These topics encompass the major fields and topics of Korean Buddhist studies

since 200. The number of papers on Buddhist research and methodology and on the socialization and popularization of Buddhism increased rapidly as special commemorative conferences were held. In contrast, research on the social role and popularization of Buddhism, as exemplified by the topic of social welfare and Buddhist practice, appears to be decreasing since 2021, when the COVID-19 pandemic played a dominant role in social trends. This paper describes trends in academic research in Korean Buddhism, and it is hoped that the findings will promote shared innovation and the sustainable exploration of perspectives in the future.

Keywords

Korean Buddhism, Topic Modeling, Big data, Text-mining, Latent Dirichlet Allocation, *Journal of Buddhist Culture Research Institute*, *Korean Journal of Buddhist Studies*, *Journal of the Korean Association for Buddhist Studies*

2023년 11월 08일 투고
2024년 03월 05일 심사완료
2024년 03월 07일 게재확정