

# 텍스트 마이닝과 연관 관계 분석을 이용한 건축역사 용어 분석

김민정\*, 김철주\*\*

숙명여자대학교 소비자경제학과\*, (주)동방문화유산\*\*

## Analyzing Architectural History Terminologies by Text Mining and Association Analysis

Min-Jeong Kim\*, Chul-Joo Kim\*\*

Dept. of Consumer Economics, Sookmyung Women's University\*  
Dongbang Cultural Heritage Corp.\*\*

요 약 건축의 한 분야인 동시에 역사학의 한 분야이기도 한 건축역사는 건축양식의 변천을 다루기는 하나 사회적, 경제적, 문화적, 기술적 상황 등의 시대 배경을 종합적으로 고찰할 필요가 있다. 그러므로 건축역사에서 주로 사용되는 용어는 다양한 분야를 아우를 수밖에 없다. 따라서 본 연구에서는 건축역사 관련 문헌을 대상으로 텍스트 마이닝과 연관 관계 분석을 수행하여 어떤 용어가 건축역사에서 핵심적인 용어인지를 파악해보았다. 우선 국내 건축역사 분야 유일한 학술지인 “건축역사연구”를 선정하여 지금까지 게재된 논문의 제목과 주제어, 초록에 사용된 용어 중 고빈도로 출현하는 핵심 용어들을 도출하였다. 다음으로 연구 분야별 문헌들을 구분하여 핵심 용어의 특징을 분석하였다. 마지막으로, 연관 관계 분석을 통해 핵심 용어들 간에 유기적인 관계를 분석하고 시각화하였다. 이러한 건축역사 핵심 용어의 파악은 건축역사 분야의 지금까지의 논의 내용과 향후 방향성을 이해하는데 유용할 것이다.

주제어 : 건축역사 용어, “건축역사연구” 학술지, 텍스트 마이닝, 연관 관계 분석, 분야별 분석

**Abstract** Architectural history traces the changes in architecture through various traditions, regions, overarching stylistic trends, and dates. This study identified terminologies related to the proximity and frequency in the architectural history areas by text mining and association analysis. This study explored terminologies by investigating articles published in the “Journal of Architectural History”, a sole journal for the architectural history studies. First, key terminologies that appeared frequently were extracted from paper that had titles, keywords, and abstracts. Then, we analyzed some typical and specific key terminologies that appear frequently and partially depending on the research areas. Finally, association analysis was used to find the frequent patterns in the key terminologies. This research can be used as fundamental data for understanding issues and trends in areas on the architectural history.

**Key Words** : Architectural history terminology, “Journal of Architectural History”, Text mining, Association analysis, Sectoral analysis

Received 1 December 2016, Revised 2 January 2017  
Accepted 20 January 2017, Published 28 January 2017  
Corresponding Author: Min-Jeong Kim  
(Sookmyung Women's University)  
Email: min-jeong.kim@sm.ac.kr

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. 서론

건축역사는 건축술, 건축양식, 건축미, 건축문화 등의 변천 과정에 대한 역사를 연구하는 분야로 건축의 한 분야인 동시에 역사학의 한 분야이기도 하다[1]. 그러므로 건축역사를 연구하는 기본전제는 시대 구분이며 건축의 모습이 달라지는 시점으로 중심으로 시대를 구분하고 각각의 다른 모습에 고유의 이름을 붙인다. 서양건축사는 일반적으로 시대와 지역으로 구분하며 우리나라의 건축사는 원시 건축, 삼국시대 건축, 통일신라 건축, 고려 건축, 조선 건축 등 왕조가 시대구분의 기준이 된다. 시대와 지역으로 구분하기는 하지만 사회적, 경제적, 문화적, 기술적 상황 등의 다양한 시대 배경을 종합적으로 고찰하여야 하므로 건축역사는 건축, 역사, 미술사, 전통건축 등 다학제적인 학문이라고 할 수 있다. 그러므로 건축역사 분야에서 주로 사용되는 ‘용어’를 파악하는 것은 우리나라에서 지금까지 이루어진 건축역사 관련 논의주제 및 트렌드를 파악할 수 있는 방법이다. 해외에는 건축역사 관련 여러 종류의 전문적인 학술지가 발간되고 있지만 국내에는 유일하게 “건축역사연구”라는 학술지를 중심으로 건축역사 연구가 진행되어 오고 있다. 학술지는 해당학문의 연구결과를 발표하는 장으로 학문의 역사적 배경과 최신 경향이 결집된 결과물이라고 할 수 있다. 따라서 학술지의 학술연구에서 주로 사용하는 용어를 분석하여 해당 분야의 논의의 흐름을 이해하고 미래 방향을 제시할 수 있다. “건축역사연구” 학술지는 건축역사 관련 다양한 분야가 융합된 학술지로 1992년에 발간되어 현재 25년째 학술지를 발간하고 있다. 따라서 학술지에 게재된 문헌들을 분석하면 우리나라 건축역사에서 어떤 용어가 핵심적이고 쟁점이 되는지를 파악할 수 있다.

텍스트 마이닝은 정형 또는 비정형 텍스트 데이터에서 가치와 의미가 있는 정보를 찾아내는 기술이다. 정보통신 기술의 발전과 다양한 디지털 기기 등의 등장으로 텍스트 데이터의 양이 증가함에 따라 텍스트 마이닝의 중요성도 부각되고 있다. 최근 들어 빅데이터 처리 기술이 발달됨에 따라 이전에는 처리하는데 많은 시간과 노력이 필요했던 방대한 자료의 분석이 가능해졌고 이러한 시대적 변화에 따라 특허정보, 학술지 논문정보, 신문기사, 온라인 담론 등 다양한 분야에서 연구가 활발히 진행되고 있다[2,3,4,5,6,7].

본 연구는 건축역사 용어의 분석을 위해서 텍스트 마이닝과 연관 관계 분석을 이용한다. 우선 텍스트 마이닝을 사용하여 주로 사용되는 용어를 추출하고 건축역사 연구 분야별 핵심 용어의 특징을 분석할 것이며, 연관 관계 분석을 통해서 핵심 용어간 인과성과 함께 연결의 맥락을 파악할 것이다. 분석대상 텍스트는 건축역사학 및 이론 연구 분야에서 국내 유일한 한국건축역사학회의 학술지 “건축역사연구”에 창간호(1992년 6월)부터 2016년 8월까지 게재된 682편 논문의 제목, 키워드, 초록이다. 이를 통해 과거부터 현재까지 국내 건축역사 분야에서 주요한 주제는 무엇이고 트렌드는 어떠한지 점검하였다.

## 2. 기존 연구 및 본 연구의 평가

건축역사 관련 분야에서 용어에 대한 연구는 “건축역사연구” 학술지에서 주로 다루어져 왔는데 2003년 이희봉은 한국건축사에서 동일 현상에 대해 다르게 붙인 용어의 호칭에 대한 제검토 필요성을 제기하였고[8] 그 이후 한국건축용어에 대한 정립을 위해 연구가 지속적으로 시도되었다[9,10,11,12]. 이와 같은 연구는 특히 한국건축사에서 이루어지는데 그 이유는 건축용어의 표기가 차자(借字)로 이루어져 있어서 그 대표어형을 찾기가 쉽지 않으며[13] 일제 강점기를 거친 이후 일본식 용어를 사용하는 경우가 많기 때문이다[14]. 이와 같이 건축역사 분야에서 용어에 대한 연구는 같은 의미가 다른 용어를 통해 전달되는 경우를 바로잡고자 대부분 진행되었으며 어떤 용어가 주로 사용되는지 과거부터 최근까지 트렌드는 어떠한지에 대한 연구는 거의 없다.

이러한 연구는 최근 들어 텍스트 마이닝과 연관 관계 분석 등 언어통계학적 관점에서 키워드의 특성과 시사점을 도출함으로써 해당 분야의 핵심용어와 트렌드를 분석한 연구 형태로 많이 진행되고 있다. 건축분야에서 텍스트 마이닝 기법을 적용한 연구로는 최정민과 강순주가 아파트단지 취재기사로부터 모범적인 아파트 관리상에서의 논점에 대한 특성과 시사점을 고찰하였으며[15] 이상훈은 현대 일본의 소규모 도시형 집합주택 계획에 대한 논문들로부터 집합주택 동수별로 키워드 특성이 차이가 있다는 것을 발견하였다[16]. 또한 건축분야 중 전통 건축 분야에서는 박은수와 김지은이 학위논문과 국내학

술지 논문 정보를 바탕으로 텍스트 마이닝 분석을 활용하여 종묘의 유·무형 콘텐츠 영향요인을 분석하였으며 [17], 수원 화성의 문화·기술적 영향요인을 분석한 연구가 있다[18].

본 연구에서는 건축역사 분야에서 텍스트 마이닝을 통해 문서에 잠재되어 있는 전달 메시지와 핵심 용어를 파악하고 연구 분야별로 핵심 용어들이 차이가 있는지 분석하고자 한다. 또한 연관 관계 분석을 통해 핵심 용어 간의 유기적인 관계를 분석하고자 한다. 이와 같은 연구는 국내 건축역사 관련 문헌을 분석함으로써 용어의 특징을 도출한 최초의 연구로서 의의를 가진다.

<Table 1> Analysis Target: Articles Published in Journal of Architectural History

Year	Published Articles	Number of Issues	Year	Published Articles	Number of Issues
1992	29	2	2005	39	4
1993	17	2	2006	29	5
1994	17	2	2007	39	6
1995	14	2	2008	27	6
1996	11	2	2009	31	6
1997	20	3	2010	44	6
1998	37	4	2011	39	6
1999	29	4	2012	41	6
2000	14	3	2013	27	6
2001	18	4	2014	33	6
2002	25	4	2015	33	6
2003	22	4	2016	22	4
2004	25	4	계	682	107

### 3. 자료의 수집

본 논문에서는 “건축역사연구”에 게재된 1992년 창간호부터 2016년 8월호까지 모든 논문 682편을 대상으로 하였으며 분석대상 논문들은 <Table 1>에 요약하였다. 우리는 682편 논문의 제목과 키워드, 초록을 분석 텍스트로 하였다. 그러나 682편 중 19편은 제목이 국문으로만 작성되었기 때문에 키워드와 초록만 포함시켜 분석하였고 682편 중 68편은 키워드가 국문으로만 작성되었으며 제목과 초록만 포함시켜 분석하였고 682편 중 1편은 영문초록이 없기 때문에 제목과 키워드만 포함시켜 분석하였다. 국문이 아닌 영문으로 분석한 이유는 “건축역사연구” 학술지의 초록이 기본적으로 영문으로 작성되어 텍스트의 대부분이 영문이기 때문이다.

## 4. 분석 방법

### 4.1 분석 도구

텍스트 마이닝과 연관 관계 분석을 적용하기 위해 사용한 프로그램은 오픈 소스 프로그램인 R 프로그램[19] 3.3.1 버전이다. 우선 수집된 논문의 제목과 키워드, 초록에 대하여 tm(text mining) 패키지를 사용하였으며 고빈도 용어를 추출하였으며 이를 바탕으로 핵심 용어-문헌 행렬을 구성하였다. 또한 연구 분야별(동/서양, 시대별) 논문의 텍스트에 대하여 tm 패키지를 사용하여 고빈도어를 추출하였다. 마지막으로 최초로 구성한 핵심 용어-문헌 행렬을 가지고 핵심 용어간 연관 관계 분석을 위해 arules(association rules) 패키지와 arulesViz(Visualizing association rules) 패키지를 사용하였다.

### 4.2 분석 과정

첫째, 1992년부터 2016년 8월호까지 “건축역사연구” 학술지에서 수집한 논문 682편을 대상으로 영문 제목, 키워드, 초록을 수집하였다.

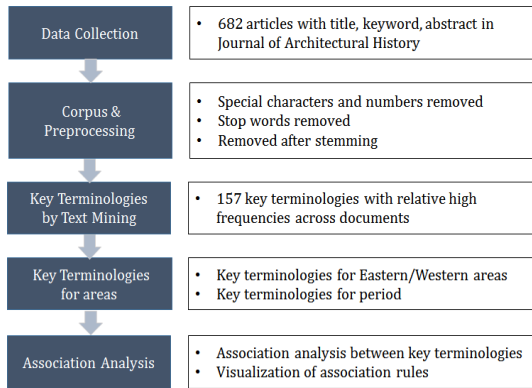
둘째, 수집한 텍스트에 대하여 말뭉치(corpus)를 작성한 후 텍스트 처리를 수행(기호, 숫자, 공백, stopwords 제거 등)하여 분석대상 용어들을 1차로 추출하였다. 1차 추출된 분석대상 용어로부터 추가적인 수정 작업을 수행하였다. 예를 들면 복수로 표현된 단어는 단수로 변환하고 과거형은 현재로 변환하는 작업을 수행하는 것이다.

셋째, 최종적으로 수정된 분석대상 용어 중에서 빈도 분석을 수행한 후 일정 횟수(30회) 이상 출현한 고빈도 용어들을 추출하는데 그 중에서 ‘study’, ‘research’ 등 일반적인 논문에서 추출되는 단어는 제외하고 최종적으로 분석대상 핵심 용어들을 선정하였다. 이를 통해 건축역사 관련 주로 사용되는 용어의 특징을 살펴보았다.

넷째, 수집한 논문에 대해서 연구 분야를 구분하고(동/서양, 시대별) 연구 분야별 각 그룹의 핵심 용어를 추출하여 연구 분야가 달라짐에 따라 사용하는 핵심 용어의 차이점이 존재하는지 살펴보았다.

다섯째, 한편의 문헌에서 존재하는 핵심 용어간의 연관규칙을 연관 관계 분석을 통해 연관 관계가 높은 핵심 용어들을 찾아내고 그 관계를 시각적으로 표현하였다.

본 연구의 분석 과정을 요약하면 [Fig. 1]과 같다.



[Fig. 1] Analysis Process of Study

### 4.3 사전처리작업

건축역사 용어에 대한 선행연구에서 제시된 것처럼 한글 용어가 동일 현상에 대해 다르게 붙인 경우가 많아 용어의 호칭에 대한 재검토 필요성이 제기되었다. 우리는 “건축역사연구“ 학술지로부터 영문 제목, 키워드와 초록을 추출하는 과정에서 한글 용어뿐만 아니라 한국건축사 관련 용어의 영문 표기(English Orthography) 역시 저자에 따라 다양하게 표기됨을 발견할 수 있었다. 이와 같이 하나의 단어에 대해 영문 표기가 달라지면 용어 출현빈도가 정확하게 추출되기 어렵기 때문에 하나의 단어에 대해 논문들에서 다르게 표시한 영문 표기를 동일하게 수정한 후 분석하였다. 본 연구에서는 국립국어원의 영어 표기 세칙을 적용하였다[20]. <Table 2>는 특정 단어에 대해 저자에 따라 영어 표기법을 다르게 작성한 일부 예이다.

<Table 2> Example of English Orthography in Journal of Architectural History

English orthography in this study	Example of English orthography
Gyeongbokgung	Gyeongbokgung, Kyeongbok palace, Kyongbok palace, Gyeongbok palace, Gyeongbokgung palace
Gyeongbuk	Kyungbuk, Kyung buk, Kyungbook
Daengjeon	Daewoongjeon, Daeungjeon, Daeong jeon, main temple
Baekje	Baegje, Baekje, Baikje, Backje, Paekche, Baekjae, Baejae, Paekje, Bagjae, Bakjae

## 5. 분석 결과

### 5.1 핵심 용어 분석

본 연구의 분석대상인 682편의 논문 제목과 키워드, 초록에서 사용된 총 단어의 수는 15711개인데 분석에 필요하지 않은 용어들을 제거하여 최종적으로 11157개의 단어들을 얻었다. <Table 3>에서 11157개의 키워드의 빈도분포를 보면 빈도 100 이하의 키워드가 추출결과의 약 98.94%를 차지하고 있으며 특히 빈도 1의 키워드가 전체의 약 42.92%를 차지하여 하나의 사례에서 밖에 사용되고 있지 않음을 알 수 있다. 이는 지역명, 건축명, 건축 관련 어휘 등과 같은 고유명사의 사용에 일부 원인이 있을 것으로 사료되지만 본 연구에서는 이에 관한 추가 분석은 생략하기로 한다.

<Table 3> Summary of Basic Text mining Results

Frequency	No. of word	Percentage	Frequency	No. of word	Percentage
1	4789	42.92%	~600	3	0.03%
~10	4800	43.02%	~700	2	0.02%
~50	1253	11.23%	~800	0	0.00%
~100	197	1.77%	~900	1	0.01%
~200	76	0.68%	~1000	0	0.00%
~300	17	0.15%	~1100	2	0.02%
~400	10	0.09%			
~500	7	0.06%	Total	11157	100%

본 논문에서 관찰대상이 되는 용어는 최종적으로 얻어진 11157개의 단어들 중 총 30회 이상 출현한 571개의 단어를 1차 분석대상으로 선정하였으며 건축역사 관련 용어분석에서 불필요한 무의미한어나 중의적 단어를 제외하고 최종적으로 157개의 단어를 핵심 용어로 선정하였다. <Table 4>에 30회 이상 출현한 건축역사 관련 핵심 용어들이 빈도 순서로 나열되어 있다. 우선 건축역사에서 가장 중요한 것이 건축이기 때문에 ‘architecture’가 압도적으로 많은 것을 알 수 있다. 또한 최상위에 ‘building’, ‘space’, ‘house’, ‘construction’, ‘structure’, ‘site’ 등 건축과 밀접한 관련을 가진 일반적인 용어들이 랭크되어 있음을 볼 수 있다. 그러나 ‘temple’, ‘pagoda’, ‘palace’, ‘stupa’, ‘castle’ 등 전통건축 관련 키워드와 ‘modern’, ‘dynasty’, ‘Joseon’, ‘century’, ‘history’, ‘royal’ 등 시대와 역사 관련 키워드들이 차상위에 랭크되어 있

<Table 4> Most Frequent Key Terminologies in Journal of Architectural History and Their Frequency

Key Terminologies	Freq	Key Terminologies	Freq	Key Terminologies	Freq	Key Terminologies	Freq	Key Terminologies	Freq
architecture	1044	western	136	province	77	confucian	49	appearance	36
building	1009	Japan	135	fire	75	corridor	49	classical	36
space	696	castle	133	society	75	eastern	49	platform	36
house	651	floor	125	heritage	73	landscape	49	archae	35
temple	553	shape	123	conservation	72	cave	47	distance	35
modern	495	China	122	pattern	72	ceremony	47	lime	35
dynasty	480	scale	122	ritual	72	ondol	47	Bukchon	34
Joseon	458	gate	119	wood	72	residence	47	entrance	33
construction	456	technique	117	eupseong	71	ttuljip	47	railway	33
structure	446	culture	116	colonial	67	court	46	craftsman	32
pagoda	431	column	115	pillar	65	excavation	46	exterior	32
Korea	366	drawing	112	restoration	65	Andong	45	ideology	32
site	320	brick	111	formation	64	joint	45	museum	32
stone	301	hanok	110	transformation	64	kan	45	pit	32
palace	265	office	108	arrangement	62	line	45	plane	32
urban	261	art	105	interior	60	reconstruction	45	commercial	31
buddhist	250	tradition	97	Seoul	60	renaissance	45	Dabotap	31
century	245	Goryeo	96	preservation	59	seowon	45	political	31
city	234	location	96	window	59	station	45	reign	31
design	230	region	96	environment	58	stair	43	soil	31
history	228	Baekje	95	relic	58	kitchen	42	sutra	31
wall	222	foundation	92	bridge	56	war	42	territory	31
material	193	Silla	91	fengshui	56	villa	41	bronze	30
hall	191	ancient	89	movement	56	shrine	40	community	30
government	186	contemporary	85	square	55	tower	40	core	30
village	182	nature	85	repair	52	settlement	39	Gyeongbokgung	30
king	181	room	84	aesthetic	51	rectangular	38	monk	30
stupa	180	tomb	84	complex	51	Asia	37	organic	30
dwelling	172	pavilion	83	painting	51	Le Corbusier	37	timber	30
royal	161	mountain	82	beam	50	corner	37		
part	153	living	81	military	50	eupchi	37		
roof	139	tile	80	bracket	49	fortress	37		

는 것으로 볼 때 일반적인 건축 관련 용어와는 차별적인 성격이 강하다고 말할 수 있다. 우리나라의 전통건축에 관련된 용어들도 <Table 4>에 많이 분포하는데 ‘hanok (한옥)’, ‘eupseong(읍성)’, ‘ondol(온돌)’, ‘ttulzip(뜰집)’, ‘kan(칸)’, ‘seowon(서원)’, ‘eupchi(읍치)’, ‘anchae(안채)’, ‘bukchon(북촌)’, ‘Gyeongbokgung(경복궁)’ 등 전통건축 종류, 형태, 구조, 단위, 위치, 실제 전통건축물들이 주로 관심있게 이어져온 주제로 나타나 있다. 또한 ‘Japan’, ‘China’가 많이 출현된 것으로 판단하건대 우리나라의 건축과 일본, 중국의 건축의 비교분석도 건축역사에서 주로 다루어져 온 주제임을 알 수 있다. 또한 ‘Le Corbusier’가 고빈도로 출출되었는데 건축역사에 대한 연구주제가 국내 또는 동양 건축에만 한정되지 않고 서양의 건축까지 논의내용으로 삼고 있다는 것을 의미한다. 이러한 결과는 건축역사 분야에서 핵심 용어는 일반 건축용어와

동서양의 건축용어, 또한 우리나라의 전통건축용어를 모두 포함하고 있음을 알 수 있다.

### 5.2 연구 분야별 핵심 용어 특징 분석

연구 분야별로 핵심 용어의 특징이 어떻게 달라지는가를 각 연구 분야마다 출현빈도가 높은 용어를 그 연구 분야의 특징어로 간주하여 추출, 분석하기로 한다. 우선 동양과 서양 건축 관련 논문의 최빈 핵심 용어 리스트를 살펴보면 다음 <Table 5>와 같다. 단 동/서양건축 연구 분야의 논문 수는 동양건축이 682편 중 528편, 서양건축이 128편으로 논문수가 다르므로 키워드 출현빈도 역시 상대적으로 서양건축 관련 논문들이 낮다. <Table 5>의 동양건축 관련 논문들에서는 ‘building’이 서양건축 관련 논문들에서는 ‘architecture’가 가장 출현빈도가 높는데 같은 건축물을 영문으로 표현할 때 동양건축은 ‘building

‘으로 많이 표현하고 서양건축은 ‘architecture’로 표현하는 것으로 나타났다. 동양건축 관련 논문에서 출현빈도가 높은 키워드는 ‘house’, ‘temple’, ‘pagoda’, ‘palace’, ‘stupa’ 같은 건축물에 관련된 키워드, ‘building’, ‘architecture’, ‘space’, ‘construction’, ‘structure’ 같은 일반 건축관련 키워드, ‘stone’, ‘wall’, ‘hall’, ‘material’, ‘roof’ 같은 건축자재 및 건축구조 관련 키워드가 있다. 또한 ‘dynasty’와 ‘Joseon’, ‘king’, ‘royal’이 상위 키워드로 추출되어 조선시대 왕실에 대한 연구가 많이 이루어졌으며 ‘buddhist’ 역시 상위에 랭크되어 불교관련 건축에 대한 연구가 많음을 유추할 수 있다.

<Table 5> Top 30 Key Terminologies in Eastern/Western Research Articles

Eastern(528)	Freq	Western(128)	Freq
building	884	architecture	339
house	629	modern	151
architecture	574	space	136
temple	542	building	110
space	513	design	98
dynasty	479	art	74
Joseon	456	history	65
pagoda	431	city	58
construction	423	urban	53
structure	403	church	51
Korea	356	century	45
modern	309	renaissance	42
stone	299	spatial	38
site	294	composition	37
buddhist	250	structure	35
palace	242	aesthetic	34
wall	197	language	34
century	193	movement	33
urban	190	villa	33
hall	188	nature	32
government	186	Le Corbusier	30
village	181	construction	29
king	180	tradition	29
stupa	180	interpretation	28
material	173	contemporary	25
city	170	interior	25
loyal	161	site	25
dwelling	159	technology	25
history	144	wall	25
roof	136	visual	24

다음으로 서양건축 관련 논문에서 출현빈도가 높은 키워드는 ‘architecture’, ‘space’, ‘building’, ‘spatial’, ‘composition’, ‘structure’, ‘construction’ 같은 일반 건축 관련 키워드와 함께 ‘design’, ‘art’, ‘aesthetic’, ‘visual’ 같

이 건축의 시각적인 측면에 관련된 키워드가 특징어로 다수 추출되었는데 특히 시각적인 측면에 대한 키워드는 동양건축 관련 논문에서는 상위에 랭크되지 않은 키워드이다. 또한 ‘modern’, ‘contemporary’ 등이 추출된 것으로 보아 현대 건축관련 연구가 많으며 시대를 나타내는 키워드는 ‘renaissance’, 건축물로는 ‘church’가 고빈도어로 추출된 것으로 보아 르네상스 시대와 교회건축물에 대한 연구가 많이 이루어졌음을 알 수 있다. 이와 같이 동/서양건축 분야에서 사용된 용어는 일반건축 관련 용어는 유사하나 동양건축 분야에서는 전통건축 및 건축자재와 구조, 서양건축 분야에서는 시각적인 측면의 용어가 많이 사용된 것으로 나타났다.

다음은 우리나라 시대별 연구 분야에 따른 핵심 용어를 추출해보면 다음 <Table 6>과 같다. 시대별 구분은 시대에 해당되는 키워드가 제목, 키워드, 초록에 포함된 논문으로 구분하였다. 각 시대별 문헌 수는 삼국시대가 39편, 고려시대가 9편, 조선시대가 190편으로 구분되며 이는 조선시대의 건축물이 가장 많이 남아있으며 의례 등 조선시대 건축물에 대한 참고자료 역시 건축역사 연구에 중요한 역할을 하기 때문인 것으로 보인다. 각 시대별 최빈 키워드의 경향을 비교하면 다음과 같다. 삼국시대에는 ‘pagoda’, ‘temple’, ‘buddhist’ 등 불교건축에 관련된 키워드와 ‘wooden’, ‘stone’, ‘brick’ 등 건축 재료들도 다수 추출되었다. 또한 통일신라시대까지를 삼국시대로 포함시켰음에도 불구하고 ‘Baekje’가 ‘Silla’보다 더 많이 추출되어 “건축역사연구” 학술지에서 백제시대의 건축에 대한 연구가 상대적으로 많이 이루어졌음을 유추할 수 있다. 고려시대에 관련된 논문은 상대적으로 적은 편이기 때문에 일의적으로 설명하기에는 근거가 부족하지만 ‘temple’, ‘sutra’, ‘Jeonryunjang(전륜장)’, ‘cabinet’과 같은 대장경과 연관된 불교 관련 키워드들이 추출되었으며 ‘China’가 추출된 것으로 보아 중국으로부터의 영향을 다룬 주제가 있었음을 유추할 수 있다. 조선시대는 “건축역사연구” 학술지에서 가장 많은 논문이 주제로 삼은 시대로 ‘house’, ‘palace’, ‘temple’, ‘government’, ‘castle’과 같은 전통건축물에 대한 키워드들이 다수 추출되어 이전 시대보다 다양한 건축 분류에 대한 연구가 이루어졌으며 ‘Japanese’가 추출된 걸로 보아 일제강점기 시대를 거치면서 조선시대 건축의 변화 등이 연구되었기 때문이라고 사료된다. <Table 6>에서 보는 것처럼 삼국시대와 고려



타내며 상관관계가 커질수록 색깔이 짙어진다. 이와 같은 규칙 하에 살펴보면 'Joseon'과 'dynasty'는 전체 논문 중에서 동시에 두 개의 용어가 존재하는 비율이 가장 큰 것으로 나타났는데 이는 전체 논문 중에서 조선 왕조에 대한 문헌이 많음을 알 수 있다. 핵심 용어간 상관관계를 살펴보면 'Ttuljip'과 'Andong'간 'Bukchon'과 'hanok'간 강한 양의 상관관계가 존재하는 것으로 나타났는데 'Ttuljip'은 조선시대 안동문화권에서 건물 가운데에 안뜰을 두어서 '口'자 모양을 이루고 있는 집을 말하는 것으로 두 용어간에는 상관관계가 높을 수밖에 없으며 'Bukchon' 역시 한옥마을과 함께 고유명사로 사용하기 때문에 상관관계 높은 용어로 나타났다. 상관관계 분석결과 위의 두 가지 이외에 상관관계가 높은 용어쌍을 살펴보면 'eupseong'과 'castle', 'interior'와 'exterior', 'gyeongbokgung'과 'palace', 'Baekje'와 'Silla', 'hanok'과 'Seoul', 'bukchon'과 'urban', 'king'과 'reign', 'beam'과 'column' 등이 있다.

그림에서 보는 바와 같이 건축역사 관련 용어들은 'architecture'와 'building'을 중심으로 모든 용어가 고르게 연관되어 있는 것으로 나타났는데 <Table 5>에서도 나타난 것처럼 'architecture'와 연결된 용어는 서양건축 관련 용어와 거시적인 관점의 건축관련 용어이며 'building'은 동양건축 관련 용어 및 재료, 구조, 기능 등의 미시적인 관점의 건축관련 용어들인 것으로 나타났다. 또한 'structure'를 중심으로 재료, 배치 등의 용어가 연결되었으며 'space'에는 'room', 'kitchen', 'floor', 'living' 같은 구조 등의 용어가 연결되어 나타났다. 그러나 별도로 고립된 핵심 용어들의 집합은 발견되지 않았는데 이는 건축역사 관련 용어들은 다양한 분야의 용어가 함께 사용되지만 'architecture', 'building', 'structure', 'space' 같은 용어들과 항상 연결되어 있다는 것을 의미한다.

## 6. 맺음말

본 연구의 목적은 건축역사 분야에서 주로 사용하는 용어를 파악하여 우리나라에서 지금까지 이루어진 건축역사 관련 논의주제 및 트렌드를 이해하는 것이다. 이를 위해 국내 건축역사 관련 유일한 학술지인 "건축역사연구"의 1992년 창간호부터 2016년 8월호까지 25년 동안 출간된 682편의 논문을 수집하고 해당 논문의 제목, 키워

드, 초록에 대한 텍스트 마이닝과 연관 관계 분석을 수행하였다. 우선 682편의 논문 제목 및 키워드, 초록에서 가장 많은 출현빈도를 가지는 157개의 핵심 용어들을 도출하였으며 도출된 용어를 분석하여 건축역사 관련된 용어들의 특징을 살펴보았다. 또한 연구분야별로 고빈도를 가지는 핵심용어들을 분석하여 연구분야별로 주로 사용하는 용어에 차이가 있는지를 분석하였다. 마지막으로 도출된 핵심 용어간 연관 관계 분석을 이용하여 시각화하였다. 주요 연구결과는 다음과 같다.

첫째, 건축역사 분야에서 주로 사용하는 용어는 일반 건축용어와 동서양의 건축용어, 우리나라의 전통건축 용어를 모두 포함하고 있으며 건축용어 범주에는 건축종류, 형태, 구조, 단위, 위치, 실제 건축물 등이 있다. 또한 시대를 나타내는 용어와 그 시대의 종교와 연관되는 건축물 등도 함께 사용되고 있었다.

둘째, 동/서양건축 분야에서 주로 사용하는 용어는 차이가 나타났는데 동양건축 분야에서는 전통건축 및 건축자재와 구조에 관련된 용어가 고빈도어로 추출되었으며 서양건축 분야에서는 건축의 시각적인 측면에 관련된 용어가 특징어로 다수 추출되었다.

셋째, 우리나라의 시대별 건축역사 용어 역시 출현경향의 차이가 나타났다. 삼국시대와 고려시대는 불교와 관련된 건축물에 대한 연구로 인해 주로 사용하는 단어들인 불교와 관련된 건축물인데 반해 조선시대 관련 연구에서는 건축 종류가 다양해지면서 사찰 이외에 집, 궁궐, 성 등 사용된 용어가 확대된 것을 볼 수 있다.

마지막으로 핵심 용어간 연관 관계 분석을 이용하여 시각화하였는데 분석 결과 'architecture'와 'building' 용어를 허브로 하는 두 개의 커다란 그룹과 'structure', 'space'를 중심으로 하는 소규모 그룹이 형성되었다. 그러나 위의 그룹들과 별도로 고립된 키워드 그룹은 존재하지 않았는데 이는 건축역사 분야는 다양한 주제를 다루지만 건축관련 일반적인 키워드들과 항상 함께 출현하고 있음을 보여준다.

본 연구에서는 텍스트 마이닝과 연관 관계 분석을 통해 건축역사 분야에서 어떤 용어가 가장 핵심적이고 서로가 어떤 밀접한 연관을 보이는지를 파악하였다. 본 연구의 결과물은 국내의 건축역사에 대한 논의가 어떻게 이어지고 있으며 기존 연구들에 유사한 핵심 용어를 가진 문헌들을 찾아내는데 활용할 수 있다. 향후 학술지 외



에 건축역사 분야에서 다량의 텍스트를 분석함으로써 매체간의 공통점과 차이점을 도출하고 매체별 정보 생산자들이 건축역사를 풀어나가는데 있어 관점의 다양성을 제공할 수 있을 것이다.

그러나 건축역사 분야의 핵심 용어 분석 시 국문으로 분석하는 것이 중요하지만 아직 국내 건축역사 분야의 텍스트가 많지 않아서 “건축역사연구” 학술지의 영문 제목과 키워드, 초록을 중심으로 텍스트 마이닝을 하였다는 점은 본 연구의 한계이며 그에 관한 후속연구가 필요하다고 본다. 또한 본 연구를 수행하는 과정에서 도출된 이슈인 관련 어휘에 대한 영어표기법의 통일은 향후 반드시 추진되어야 할 것으로 사료된다. 한국학을 연구하는 외국인들이 점점 늘어나고 있는 상황에서 건축역사 분야가 선도적으로 관련 어휘에 대한 표준화된 영어표기법을 제시하는 것은 다른 분야에서도 참고할 수 있는 유익한 기준으로써 기대할 수 있다.

## REFERENCES

- [1] H. Khang, C. S. Yoon, “A Study on the Approach and Interpretation in the Architectural History.” Proceedings of the Architectural Institute of Korea, Vol. 2, No. 2, pp. 27-30, 1982.
- [2] S. J. Lee, S. H. Lee, H. J. Seol, Y. T. Park, “Using Patent Information for Designing New Product and Technology: Keyword based Technology Roadmapping.” R and R Management, Vol. 38, No. 2, pp. 169-188, 2008.
- [3] S. G. Cho, S. B. Kim, “Finding Meaningful Pattern of Key Words in IIE Transactions using Text Mining.” Journal of Korean Institute of Industrial Engineers, Vol. 38, No. 1, pp. 67-73, 2012.
- [4] A. Balahur, R. Steinberger, “Rethinking Sentiment Analysis in the News: from Theory to Practice and Back.” Proceedings of the 1st Workshop on Opinion Mining and Sentiment Analysis, pp. 1-12, 2009.
- [5] M. H. Hwang, J. Park, “Lexical and Phrasal Analysis of Online Discourse of Type 2 Diabetes Patients based on Text-Mining.” Journal of Digital Convergence, Vol. 12, No. 6, pp. 655-667, 2014.
- [6] Y. S. Jeong, “Business Process Model for Efficient SMB using Big Data.” Journal of IT Convergence Society for SMB, Vol. 5, No. 4, pp. 11-16, 2015.
- [7] S. H. Hong, “New Authentication Methods based on User’s Behavior Big Data Analysis on Cloud.” Journal of IT Convergence Society for SMB, Vol. 6, No. 4, pp. 31-36, 2016.
- [8] H. B. Lee, “For Terms of Korean Traditional Architecture-Toenmaru or Jjokmaru ?” Journal of Architectural History, Vol. 12, No. 4, pp. 226-228, 2003.
- [9] H. B. Lee, “Effort for Establishment of Terms of Korean Traditional Architecture-1. Toenmaru, Jjokmaru.” Journal of Architectural History, Vol. 20, No. 4, pp. 197-204, 2011.
- [10] H. B. Lee, “Effort for Establishment of Terms of Korean Traditional Architecture-2. Dori: Napdori, Doridori, Jjakjjakkung ?” Journal of Architectural History, Vol. 20, No. 5, pp. 144-147, 2011.
- [11] H. B. Lee, “Effort for Establishment of Terms of Korean Traditional Architecture-3. ‘Soeseo?’ Cow tongue ?” Journal of Architectural History, Vol. 21, No. 1, pp. 217-222, 2012.
- [12] H. B. Lee, “Gyoduhyeong ?-Establishment of Terms of Korean Traditional Architecture 5.” Journal of Architectural History, Vol. 23, No. 1, pp. 84-92, 2014.
- [13] W. J. Kim, “Can Architectural Terminology in Uigwe be a Standard for Terminology Unification ?” Journal of Architectural History, Vol. 14, No. 1, pp. 227-234, 2005.
- [14] K. I. Jang, “The Course of Establishment of Architecture Terms.” Review of Architecture and Building Science, Vol. 33, No. 5, pp. 66-68, 1989.
- [15] J. M. Choi, S. J. Kang, “Issues of Articles Covering Outstanding Management of Apartment Complexes - Content Analysis of Newspaper Reports with Lexical Statistics -.” Journal of the Korean Housing Association, Vol. 17, No. 4, pp. 131-143m 2006.
- [16] S. H. Lee, “A Study on the Characteristics of Small-scale Urban Housing Planning in Modern

Japan-Feature Analysis of Architect's Discourse in Group Housing by Text mining-." Journal of the architectural institute of Korea : Planning & design, Vol. 30, No. 9, pp.129-136, 2014.

- [17] E. S. Park, J. E. Kim, "A Study on the Tangibility and Intangibility Value Contents Influence Factor of Kongmyo Shrine Using Text Mining Analysis." Korea Science & Art Forum, Vol. 22, pp. 169-183, 2015.
- [18] E. S. Park, J. E. Kim, "A Study on the Cultural and Technical Influence Factor Using Unstructured Data Analysis." Korea Science & Art Forum, Vol. 20, pp. 227-241, 2015.
- [19] K. H. Choi, J. A. Yoo, "A Reviews on the Social Network Analysis using R." Journal of the Korea Convergence Society, Vol. 6, No. 1, pp. 77-83, 2015.
- [20] National Institute of Korean Language, "Rules for English Orthography." [http://www.korean.go.kr/front/page/pageView.do?page\\_id=P000124&mn\\_id=97](http://www.korean.go.kr/front/page/pageView.do?page_id=P000124&mn_id=97).
- [21] Y. B. Jung, E. S. Park, "Keyword Analysis of Two SCI Journals on Rock Engineering by using Text Mining." Tunnel & Underground Space, Vol. 25, No. 4, pp. 303-319, 2015.
- [22] Y. J. Shin, M. S. Yim, "A Study of the Relationship Analysis between Mobile Application by using an Association Rules." Journal of the Korea Convergence Society, Vol. 3, No. 2, pp.19-26, 2012.

김 철 주(Kim, Chul Joo)



- 1994년 2월 : 서울대학교 경영학과 (경영학사)
- 2015년 1월 ~ 현재 : (주)동방문화유산 대표
- 관심분야 : 고건축 시공, 건축역사
- E-Mail : eckehomo@hanmail.net

김 민 정(Kim, Min Jeong)



- 2012년 2월 : 서울대학교 산업공학과(공학박사)
- 2015년 3월 ~ 현재 : 숙명여자대학교 소비자경제학과 교수
- 관심분야 : User & Data Analysis
- E-Mail : min-jeong.kim@sm.ac.kr