

다차원 분석방법을 활용한 중소규모 공동주택 건축심의 의견의 경향과 비정형 데이터로서의 특성분석

김진희¹ · 황태언² · 김재식³ · 허영기^{4*}

¹(주)해안건축 상무 · ²부산대학교 건축공학과 학사과정 · ³한국수력원자력 안전경영단 차장 · ⁴부산대학교 건축공학과 교수

Multidimensional Analysis of Unstructured Data and Trends in Architectural Review Opinions of Small and Medium-Sized Apartment Projects

Kim, Jinhee¹, Hwang, Taeon², Kim, Jae-Sik³, Huh, Youngki^{4*}

¹Busan Office Director, Haeahn Architecture

²Undergraduate Student, Department of Architectural Engineering, Pusan National University

³Senior Manager, Safety Management Unit, Korea Hydro & Nuclear Power Co. Ltd.

⁴Professor, Department of Architectural Engineering, Pusan National University

Abstract : This study examines the characteristics of architectural review opinions as unstructured data, focusing on the most challenging risk for developers of small and medium-sized apartment projects in response to the increasing number of single-person households in Korea. Using multidimensional analysis methods, the study analyzes the review opinions of 25 projects in B City. Correspondence analysis and MDS (Multidimensional Scale) analysis show that, consistent with prior research, the keywords related to 'structure' and 'planning' dominate architectural review opinions in B City. While the MDS model's stress is very poor at 34.4%, correspondence analysis reveals that this is due to the characteristics of unstructured data in architectural reviews. In addition, the non-structured data analyzed in this study, such as architectural review opinions, exhibited a probability distribution with low kurtosis and high skewness, as they involved various combinations and occurrences of data depending on the discretion of the review committee members and the specific formats of different local governments. This often led to the emergence of keywords that differed significantly from commonly mentioned terms. Although the study has some limitations, it provides a foundation for future detailed analysis by identifying the characteristics of architectural review opinions as unstructured data.

Keywords : Unstructured Data, Architectural Review, Small and Medium-sized Apartment, Multidimensional Scale

1. 서론

1.1 연구의 배경 및 목적

중소규모 공동주택에 대한 관심과 수요는 지속적으로 증가하는 추세이다. Lee (2012)에 따르면 국내 1인 가구 수는 2035년까지 전체 가구 수의 34.3%까지 증가할 것으로 예상된다. 또한, 2020년에는 소규모 공동주택 관리지원 법제화가 이루어졌으며, 이러한 거주유형에 대한 정책적 보완도 이루어지고 있어 중소규모 공동주택의 심의 건수도 증가하고

있다. 그러나, 전체 건축프로세스 중 인허가 절차는 전문가들에게도 매우 큰 불편을 느끼게 하는 과정이며(Ministry of Land, Infrastructure, and Transport, 2015), 특히 비교적 경험이 많은 시행자가 진행하는 대규모 공동주택사업에 비하여 개인 혹은 소규모 사업주체가 시행하는 중소규모 공동주택 사업에서는 큰 리스크로 작용하고 있다.

건축심의의 절차와 결과는 사업비에 직접적인 영향을 미침으로, 건축심의 과정에서 발생하는 문제점과 개선방안에 대한 연구는 지속적으로 이루어지고 있다. 그러나, 건축심의는 건축심의 위원들이 사용하는 용어나 지자체별로 사용하는 표준 양식의 차이 등으로 인해 데이터의 비정형성이 발생하여 컴퓨터 연산을 통한 분석이 매우 어렵다. 따라서, 본 연구에서는 Ryu et al. (2022)의 선행 연구에서 직접 분류한 건축심의 의견의 상위 분류 키워드를 근간으로, 대응분석과

* **Corresponding author:** Huh, Youngki, Department of Architectural Engineering Pusan National University, Pusan 46241, Korea
E-mail: ykhuh@pusan.ac.kr
Received September 14, 2023; **revised** -
accepted November 27, 2023

다차원 척도법(Multidimensional scaling, 이하 MDS)을 활용하여 중소규모 공동주택의 건축심의 과정에서 도출된 심의 의견의 경향을 분석하고, 비정형 데이터로서의 심의 의견의 특성을 파악해 사업경험이 적은 중소규모 공동주택 사업주의 건축심의에서의 리스크를 경감시키는데 도움을 주는 것을 목적으로 한다.

1.2 연구의 범위 및 방법

본 연구는 Ryu et al. (2022)가 직접 분류한 B도시에서 이루어진 25개 중소규모 공동주택사업에 대한 403건의 건축심의 의견을 대상으로 하였다. 건축심의 의견은 다양한 주제와 의견을 다루며, 심의위원의 주관과 동일 개념에 대한 다른 용어들의 사용 등으로 인해 비정형 데이터로 인식된다. 비정형 데이터의 특성은 구조화된 데이터와 달리 직접적인 분석이 어렵고, 수작업으로 데이터를 분류해야 하는 번거로움이 존재해 일관성 있는 분석이 어려운 한계점이 있다.

이러한 문제를 해결하기 위해서 본 연구는 한국어 분석 모델인 Python의 KoNLP (Korean Natural Language Processing)를 사용하여 데이터를 전처리 후, Romesburg (2004)의 저서를 참고하여 대응분석을 진행해, 선형적 관계를 가지는 군집과 데이터의 특징을 개략적으로 확인하였다. 이후, 앞선 분석을 토대로 MDS (Multidimensional Scaling)을 활용해 군집분석을 진행하고, 결과를 선행연구의 건축심의 분석 및 분류한 전문가 집단과 동일한 인물로 구성된 집단 인터뷰를 통해 건축심의 자료의 경향성과 건축심의의 비정형 데이터로서의 자료특성을 파악하였다(Fig. 1).

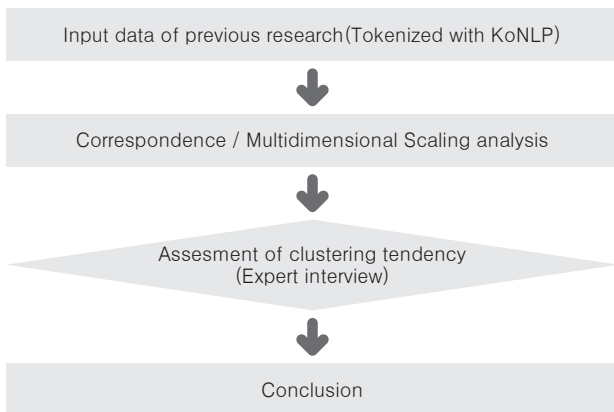


Fig. 1. Analyzing process for the architectural review opinions as unstructured data

1.3 문헌연구

건축심의회에 대한 연구는 주로 설문조사 및 직접분류를 통하여 진행되어왔다. Kang (2012)는 직접분류를 통해 광역지

자체별 심의기준을 분석해 지자체별 심의기준의 중점요소에 따라 “지침을 폐지하여 자율재량기능을 최대화한 유형”, “성능기준의 성격을 강화하여 자율재량을 우선하되 자동행정의 기능을 가미한 유형”, “자동행정과 자율재량을 복합한 유형”, 총 3가지로 구분하였다. Choi and Kim (2022)는 지방건축위원회의 심의결과를 분석하고 소형집합주택 사업에서의 주요 대응 검토항목을 체크리스트화 하는 연구를 진행하였다. 해당 연구에서는 심의 의견의 내부 부위별 요소 빈도 분석을 통하여 의견빈도가 5회 이상인 7개분야 대부분류와 39개 항목 중분류를 구분하여 서울시 소형집합주택의 심의결과에 대한 내부 대응 설계항목들의 체계화된 기초자료를 제안하였다.

본 연구의 분석대상인 Ryu et al. (2022)의 연구는 B광역시시의 대표적인 25개 중소규모 공동주택 사업들에 대하여 항목별 심의 의견을 분석하였다. 전문가 집단의 인터뷰를 통해 선행연구의 분류체계를 발전시켜 30가지 주요 키워드를 도출하고 상위 및 하위 개념으로 분류해 광범위한 자료 분석을 위한 기초를 제시하였다. Lee et al. (2016)은 통계분석 프로그램인 R을 사용하여 비정형 건설 데이터에 해당하는 영국의 중재 판례들에 대한 자료들을 추출하여, 건설분쟁판례에 대하여 컴퓨팅 기술을 활용하는 문서 분석방법론을 제시하고 다양한 분쟁사례들에 대한 분석 결과를 도출하였다.

1.4 다차원 데이터의 분석방법

본 연구에서는 비정형 데이터의 2차원 그래프 분석을 위해 MDS와 대응분석을 사용하였다. MDS는 다차원상에 존재하는 개체 간 유사성 또는 비유사성을 저차원 공간에 기하적으로 나타내어 개체 간 관계를 파악하는 다변량 기법이다(Choi, 2014). Jang (2001)에 따르면, MDS 분석은 다차원 데이터를 저차원으로 축소시키는 과정에서 군집화(Clustering)의 탐사능력이 급격히 저하되는 차원의 저주(Curse of dimensionality)를 극복하기 위한 유용한 방법론이다. 따라서, MDS 분석 결과를 바탕으로 미래의 예측 모델링에 적용할 수 있으며, 이를 통해 보다 정확하고 신뢰성 높은 결과를 도출할 수 있다. 본 연구에서 적용한 MDS의 작성방법은 가장 좋다고 생각되는 것을 해답으로 나타내는 탐욕적 알고리즘(Greedy Algorithm) 중 가장 최적의 해답을 주는 Kruskal 알고리즘(Chun et al., 2019)을 적용하였으며, 분포데이터의 특징이 계층적 군집을 이루며 범주형(Categorical)인 성질을 고려하여 자카드 지수 계산법(Jaccard index)를 통해 산개도를 도출하였다. 해당 계산법은 두 범주의 합집합이 아닌 원소는 제외하고 합집합, 교집합의 원소만 활용하여 1 이하의 지수값을 계산하는 방법이다(H. Charles, 2004). Park (2016)에 의하면, 대응분석은 분

할표(Contingency table)상의 행(row)과 열(column)을 저차원의 공간에 나타내어 행과 열의 관계를 알아보는 탐색적 자료분석(explanatory data analysis)기법 중 하나로, 분할표상의 행과 열 변수간의 변동정도를 나타내는 변량(Variate)을 수리적으로 나타내는데 활용된다. 이후 Chikio Hayashi에 의하여 수량화 제3방법(quantification method III)으로 발전해, 다차원 분석방법의 발전에 기여하였다.

2. 건축심의 의견 분석 및 검증

2.1 텍스트 마이닝

본 연구는 앞선 중소규모 공동주택 사업 및 건축심의의 연구 동향을 바탕으로, 전문가 집단과의 인터뷰를 통해 직접 분석한 25건의 대표 사업들(Table 1)의 건축심의 의견에 대한 키워드를 (Fig. 2)와 같은 절차로 분석하였다.

Table 1. A summary of 25 representative building projects

Province	Building Size (Basement floors/ Ground floors/ Gross floor area)	Building use	Review date	Reviewed opinions
	Unit (floor/floor/m ²)			
D	2 / 20 / 17,462	Apartment	2020.07.17	16
	3 / 17 / 15,253	Apartment	2020.10.27	6
	1 / 15 / 4,032	Officetel	2021.03.17	9
	1 / 17 / 5,785	Apartment	2021.03.17	9
	1 / 18 / 5,413	Officetel	2022.06.10	3
I	2 / 20 / 8,025	Apartment	2018.03.19	12
	1 / 17 / 3,817	Apartment Complex	2021.05.14	19
	1 / 19 / 7,968	Apartment Complex	2021.05.14	42
	1 / 15 / 7,040	Apartment	2021.08.19	28
	1 / 19 / 3,988	Officetel	2022.08.03	37
K	2 / 15 / 4,400	Officetel	2022.01.14	11
	4 / 19 / 27,112	Apartment	2022.03.08	16
	2 / 20 / 19,128	Apartment Complex	2022.04.13	11
	5 / 20 / 14,672	Officetel	2022.05.12	9
	1 / 17 / 3,460	Apartment Complex	2022.05.12	8
N	1 / 20 / 7,103	Apartment Complex	2019.02.13	15
	0 / 9 / 1,361	Apartment Complex	2019.08.26	9
	0 / 19 / 5,591	Apartment Complex	2019.08.26	10
	1 / 8 / 1,283	Apartment Complex	2020.01.15	4
	1 / 20 / 3,647	Apartment Complex	2022.07.29	8
P	1 / 18 / 4,007	Apartment Complex	2020.06.05	17
	1 / 6 / 2964	Apartment	2021.12.10	22
	2 / 12 / 6694	Apartment	2022.01.21	30
	2 / 6 / 4371	Apartment	2022.05.17	31
	1 / 6 / 2996	Apartment	2022.06.24	21

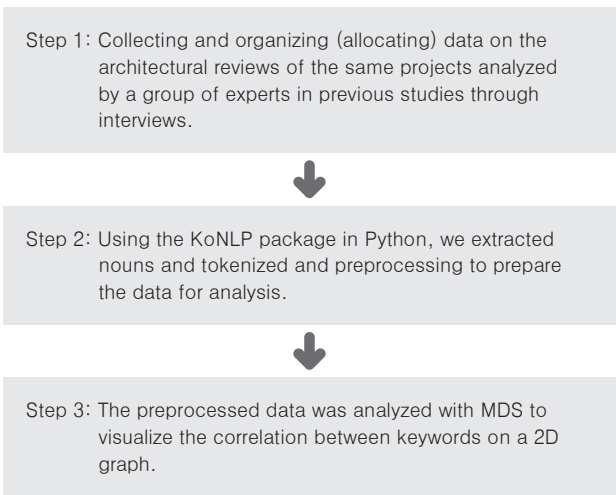


Fig. 2. Natural language processing procedure

본 연구에서는 상관성 분석을 위하여 KoNLP 패키지의 명사 추출 기능을 활용하고, 남은 형태소는 도출된 키워드의 의미 해석에 영향을 주지않는 단어들에 대하여 불용어(stopwords)를 다음 예시와 같이 설정하여 제거하였다.

불용어 삭제예시 1) ['바','나','모','리','의','가','이','은','들','는','좀','잘','강','과','도','니','받','서','면서','적용권고','내','를','으로','자','에','와','한','하다','할','준용','것','계','시','고']

불용어 삭제예시 2) 형성하시,'추가하시','수립하시','조치바','제시바','재판단하','검토바람','조정바람','확보하시','설치하시','보완하시','설계하시','확인검토하시','해당'

제외된 불용어들은 대부분 심의결과를 해석하는데 의미가 없는 종결형 어미('검토바람', '조정바람' 등), 혹은 명사 추출과정에서 접속사가 남아 의미가 모호한 단어들('면서', '도', '것' 등) 및 단일 의미로서 곡해가 가능한 단어('내', '한', '계', '회' 등)를 추출된 명사 목록을 확인하여 제거하였다.

전처리 완료된 데이터의 KH coder를 통한 출현빈도 상위 20개의 키워드는 (Table 2)와 같다.

관련 키워드 중 빈도수 상위 20개는 '구조', '계획', '기초', '설계', '지하', '적용' 등이며, 전체 프로젝트 심의결과는 403건으로, 가장 많이 언급된 '계획'은 73번이다. 중복된 의견이 여러 번 언급 경우도 있으므로, 단일 키워드로는 유의미한 분석이 어려워, 다차원 분석을 통해 빈도수가 10회 이상인 데이터들 간의 상호 연관성을 분석하고, 해당 군집에서 나타나는 키워드들의 빈도수를 합산하여 해당 군집의 상위분류를 결정하여 경향성을 파악하는 데 활용하였다.

Table 2. The top 20 extracted keywords and frequency

Rank	Word	Frequency
1	Planning	73
2	Designing	54
3	Structure	48
4	Underground	46
5	Application	45
6	Soil	42
7	Deliberation	39
8	Foundation	38
9	Reflection	36
10	Facility	36
11	Location	36
12	Ground	35
13	Construction	32
14	Case	31
15	Design	31
16	Secure	29
17	Architecture	28
18	Installation	28
19	Car	26
20	Building	25

2.2 대응분석

MDS 분석에 선행하여, 비정형 데이터를 다차원 특정 벡터로 전환하여 대응분석을 실시하였다. 2차원으로 차원을 축소하여 대응분석한 결과는 <Fig. 3>과 같다. 해당 대응분석 결과 총 7개의 선형적 관계를 가지는 군집이 분석되었으며, 이 군집들은 Ryu et al. (2022)의 선행연구에서 직접 분류한 상위 키워드와 비교하기 위해 '계획', '구조', '설비', '안전', '기타' 항목으로 구분하여 해석하였다.

대응분석 결과, 많은 키워드 간의 상관관계에서 데이터 패턴이 두드러지게 유사하여 선형적 상관성을 보이는 군집은 구조 및 계획 관련 키워드의 군집으로 분석되었다. X축의 차원의 0값을 기준으로 음수영역엔 상위분류 '구조'와 관련된 키워드가 주를 이루었고, 양수영역엔 상위분류 '계획'과 관계된 키워드가 주를 이루었다. 이 외의 군집은 앞서 언급한 '구조' 및 '계획' 관련 군집과는 패턴이 다른 집단으로 분류하였다. '순환', '골재', '콘크리트', '사용'과 같은 데이터는 각 차원의 0값과 멀게 나타나므로, 이 데이터로 인해 저차원해석의 적합도가 낮아질 수 있다는 것을 인지하였다.

2.3 MDS 분석

이전에 수행한 대응분석 결과에서는 구조 및 계획 관련 키워드들이 상호 연관성이 높고 서로 유사한 거리를 가지는 것으로 나타났다. 이에 따라 Kruskal Algorithm을 이용하여 2차원 MDS를 작성한 결과는 <Fig. 4>와 같다.

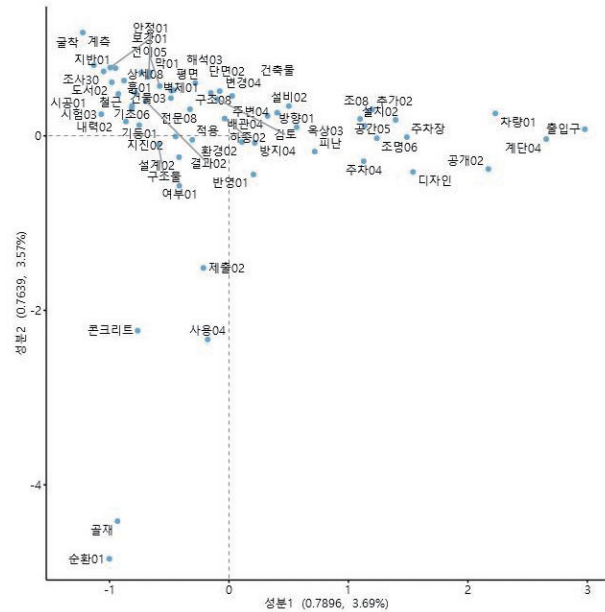


Fig. 3. Correspondence analysis for tendency identification

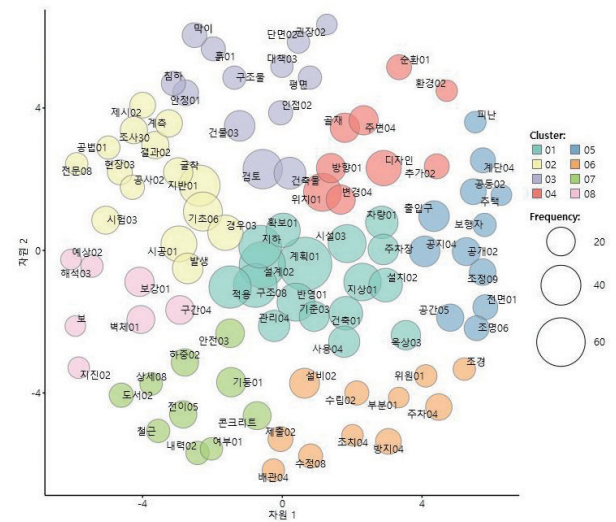


Fig. 4. The result of a two-dimensional MDS analysis

MDS 그래프는 총 8개의 군집으로 구성되었으며, 각 원의 크기는 해당 데이터가 나타난 빈도수를 나타낸다. 적합성을 나타내는 Stress 값은 34.4%로 나타났으며 각 차원의 0값과 거리가 먼 일부 데이터로 인해 적합성이 손상된 것이 주된 원인으로 추론되며, 이러한 결과는 건축심의 의견의 비정형 데이터의 특성으로 인한 것이다. 해당 MDS 작성결과에 대한 분석 및 분류는 선행연구와 동일한 전문가 집단과의 브레인스토밍을 통해 분류하였다<Table 3>.

Table 3. Clusters and words in two-dimensional MDS

Cluster	Class	Words
01	Planning	Basement, Designing, Structure, Management, Application, Secure, Facility, Ground, Installation, Parking lot, Car, Usage, Standard, Reflect, Rooftop
02	Structure	Construction, Experiment, Occur, Case, Foundation, Soil, Excavation, Result, Measurement, Expertise, Shoring system, Site, Surveying, Suggest
03	Structure	Building, Deliberation, Adjacent, Floor plan, Dirt, Stability, Submerge, Section plan, Recommend, Structure
04	The others	Environment, Surroundings, Design, Addition, Change, Location, Direction, Aggregate, Circulation
05	Planning	Entrance, Stair, Evacuation, Joint, Apartment, Space Pedestrian, Public, Vacant land, Adjustment, Frontal, Lighting,
06	The others	Commissioner, Parking, Part, Establishment, Action, Submit, Prevention, Utility, Modify, Plumbing, Landscaping
07	Structure	Safety, Load, Column, Concrete, Inertia, Detail, Documents, Rebar, Congestion, Whether
08	Structure	Reinforcement, Segment, Wall, Beam, Prediction, Interpretation, Earthquake

위 표의 01번 군집은 상관성이 높은 최다빈도 키워드로 구성되어 있으며, 본 연구에 적용된 대표 프로젝트의 심의 의견에서는 '계획', '설계', '적용', '구조' 등 각 분야의 심의 의견에서 보편적으로 나타날 수 있는 공통어가 나타났다. 해당 군집은 '구조' 관련 키워드가 존재하지만, 대부분의 키워드가 '계획'과 관련되어 상위분류 '계획'으로 분류하였다. 02번 군집은 '지반', '기초', '굴착', '계측' 등으로, 구조의 토질 및 기초에 해당하는 내용이 주를 이루어 상위분류 '구조'로 분류하였으며, 03번 군집도 동일하게 '흙막이', '인접', '침하' 등의 키워드가 도출되어, 상위분류 '구조'로 분류하였다.

04번 군집은 '디자인', '순환', '골재', '환경' 등으로, 동일한 경향을 가진 상위분류에 포함되기 어려운 키워드를 포함하여 상위분류 '기타'로, 05번 군집은 '출입구', '계단', '보행자', '공개', '공지', '공간' 등 평면 및 배치에 관련한 키워드가 주를 이루어 상위분류 '계획'으로 분류하였다.

06번 군집은 '설비', '주차', '방지', '수정' 등 4번 군집과 동일하게 동일한 경향을 가지지 못하므로 상위분류 '기타'로 분류하고, 07번 군집은 '기둥', '전이', '하중', '철근' 등을 포함하여, 부재의 상세설계 등에 관한 내용으로 해석하여 상위분류 '구조'로 분류하였다. 끝으로 08번 군집은 '보강', '벽체', '해석', '보', '지진' 등 구조부재 해석 및 보강에 관련한 키워드로 상위분류 '구조'로 분류하였다.

2.4 다차원 분석에 대한 상세분석

선행 연구에서 적용한 키워드 상위 분류에서 각 군집에서 적용 가능한 상위 분류가 존재하지만, 출현 빈도수를 10회

로 제한한 결과, 설비, 안전 등과 같은 키워드는 높은 상관성을 보이지 않아 '기타' 분류로 04번과 06번 클러스터에서 혼합되어 나타났다.

'계획' 분류에 해당하는 01번, 05번 군집의 키워드들의 빈도수의 합은 전체 단어의 출현 빈도수 대비 39%이며, '구조' 분류에 해당하는 02번, 03번, 07번, 08번 군집의 키워드의 빈도수의 합은 44%를 차지하였다. '기타' 분류에 해당하는 04번, 06번 군집의 키워드의 빈도수 합은 18%로 나타났다.

단순 빈도수의 합을 통해 분석한 군집별 분석 결과, 본 연구에서는 '구조' 44%, '계획' 39%로 나타나 선행연구의 상위분류별 심의 의견 분류 빈도수의 합인 '구조' 40%, '계획' 27%와 근접하나, 04번 군집과 같이 군집 내에서 동일한 상위분류로 묶이기 어렵거나, 01번 군집과 같이 다른 상위분류에 해당하는 키워드가 혼합되어 나타나는 특징을 보이는 군집으로 인해 '계획' 군집의 빈도가 높게 나타났음을 알 수 있다.

이와 같은 분석을 통한 건축심의 의견의 데이터 특징은 심의위원의 주관 및 지자체별 양식에 따른 각기 다른 요구사항으로 인해, 건축심의를 주로 언급되는 단어와 전혀 다른 키워드('순환', '골재' 등)가 등장 할 수 있으며, 다양한 데이터의 조합 및 출현이 이루어지는 것을 알 수 있었으며, 이로 인해 건축심의 의견은 확률분포적 개념에서 첨도(Kurtosis)가 낮고, 주로 언급되는 단어들의 군집이 대응분석 상에서도 0값을 기준으로 분류되어, 군집을 이루는 집단이 상위분류를 기준으로 여러개 나뉘어 왜도(Skewness)가 크게 나타나는 것으로 파악되었다.

또한, 본 연구결과로 도출된 MDS 군집별 단순 빈도의 합은 전문가 집단 인터뷰를 활용하여 직접 분석 및 분류한 Ryu et al. (2022)의 연구결과와 비교할 때 상위분류 '계획'에서 유사치를 보이지 않은 것으로 나타났다. 이는 B도시 심의 의견의 경향파악을 위해 군집을 임의로 분류하여 전체의 빈도수를 단순 합산한 것, 모델의 적합도가 낮은 것, 차원상의 데이터의 거리를 이해하기 쉽도록 2차원으로 한정된 것 등의 이유로 판단된다.

3. 결론

국내 1인 가구수의 증가 등에 따른 중소규모 공동주택사업의 지속적인 확대에 비교적 경험이 적은 해당 사업유형의 사업주가 대비하기 매우 어려운 리스크인 건축심의를 보다 효율적으로 준비하고 대응하기 위한 자료를 제공하기 위해, 다차원 척도법(Multidimensional scaling, MDS)을 적용해 B도시의 25개 중소규모 공동주택 프로젝트에 대한 건축심의 의견을 분석하였다.

건축심의 의견은 비정형 데이터로, 컴퓨터 연산을 활용한 분석은 매우 어렵다는 한계점이 있다. 이에 선행연구에서 Ryu et al. (2022)는 해당 프로젝트들의 심의 의견을 전문가의 인터뷰 등을 활용해 직접적으로 분석하여, '계획', '구조', '설비', '안전', '기타'의 항목으로 구분하여 제시하였으며, 본 연구에서는 대응분석 및 MDS 기법을 작성 활용하여 동일 자료들을 분석하고, 앞의 직접 분류한 결과값들과의 경향성 차이를 확인하였다.

대응분석 결과, '순환', '골재' 등의 키워드는 다른 키워드와의 상관성이 매우 낮은 데이터로 나타났으며, 건축심의 의견에 대한 분석 시 이와 같은 키워드로 인해 적합도가 낮아지도록 영향을 줄 수 있음을 확인하였다.

MDS 분석결과 B도시 건축심의 의견은 상위분류 '구조'와 '계획'에 해당하는 키워드들이 가장 높은 비중을 차지하였지만, 건축심의 의견 데이터의 비정형성으로 인해 분석모델이 나쁜 적합도를 보였다. 모든 키워드의 출현빈도수 대비 각 상위분류의 빈도수의 합은 선행연구의 직접분류결과 '구조' 40%, '계획' 27%이며, 본 연구의 분석결과는 '구조' 44%, '계획' 39%으로 나타났다.

본 연구에서 분석한 건축심의 의견의 비정형적 데이터로서의 특징은 심의위원의 주관과 지자체별 양식에 따라 다양한 데이터의 조합과 출현이 이루어지며, 주로 언급되는 단어와 전혀 다른 키워드가 등장할 수 있어 첨도가 낮고 왜도가 높은 확률분포적 특성을 파악할 수 있었다.

본 연구는 선행연구와 비교해 상위분류 '계획'에서 유사치를 보이지 못하는 한계를 보이거나, B도시의 건축심의 의견의 경향과 비정형 데이터로서 건축심의 의견의 특징을 파악하여 추후 빅데이터와 기계학습을 통한 건축심의 의견분석에 기초연구로서 활용 될 수 있을 것으로 기대된다.

감사의 글

이 과제는 부산대학교 기본연구지원사업(2년)에 의하여 연구되었음.

References

- Choi, K.O., and Kim, W.P. (2022). "A Checklist of Major Responsive Review Elements of Small-size Collective Housing Based on a Local Architectural Committee's Review Results." *Journal of the Architectural Institute of Korea*, AIK, 38(4), pp. 61-72.
- Choi, Y.S. (2014). *Multidimensional Scling*, 1rd ed, Jayou academy, Paju-si, Chapter 1.
- Chun, I.G., Kong, Y.H., and Ha, S.H. (2019). *Data Structures*

- in C, 2nd ed, Sangneung chulpansa, Paju-si, Chapter 11-2.
- Jang, M.H. (2001). "A Clustering Technique using Two-Dimensional Projection in High-Dimensional Data." MS thesis, Myeongji Univ., Yongin-si, Gyeonggi-do.
- Kang, I.H. (2012). "A Study on the Characteristics of Design Council Criteria of Housing - Focused on Metropolitan Cities -." *Journal of The Korean Housing Association*, *The Korean Housing Association*, 23(1), pp. 97-105.
- Lee, J.H., Yi, J.S., and Son, J.W. (2016). "Unstructured Construction Data Analytics Using R Programming - Focused on Overseas Construction Adjudication Cases -." *Journal of the Architectural Institute of Korea Structure & Construction*, AIK, 32(5), pp. 37-44.
- Lee, Y.S. (2012). "Trend and Characteristics of Single Person Households." *Korean Social Trends 2012*, *Statistical Research Institute*, pp. 68-76.
- Ministry of Land, Infrastructure, and Transport (2015). *Proposing Policies for Efficient Architectural Administration -A Focus on the Improvement of Architectural Review System*, KRIHS Research Report, 2015-04.
- Romesburg, H.C. (2004). *Cluster analysis for researchers*, 1st ed, Lulu Press, Morrisville, p. 151.
- Ryu, J.W., Kong, I.W., Hwang, B.W., Gu, B.G., and Huh, Y.K. (2022). "Analysis on Main Opinions drawn from Municipal Architectural Review for Small and Medium-Sized Apartments." *Journal of the Regional Association of Architectural Institute of Korea*, AIKRA, 24(6), pp. 105-110.

요약 : 본 연구는 국내 1인 가구수가 증가함에 따라 중소규모 공동주택에 대한 정책적 관심이 높아지는 가운데, 해당 사업 유형의 사업 주들이 대비하기 가장 어려운 리스크로 건축심의가 지적되고 있다. 본 연구는 B도시의 25개 중소규모 공동주택 프로젝트에 대해 다차원 분석방법을 적용하여 건축심의 의견의 경향과 비정형데이터로서 건축심의 의견의 특성을 분석하였다. 대응분석 및 MDS 분석을 실시한 결과, 선행연구와 동일하게 B도시의 건축심의 의견은 주로 상위분류에서 '구조'와 '계획'에 관련된 키워드가 주를 이루었다. 즉, 모든 키워드의 출현빈도수 대비 각 상위분류의 빈도수의 합은, 선행연구의 직접분류결과는 '구조' 40%, '계획' 27%이며, 본 연구의 분석결과는 '구조' 44%, '계획' 39%인 것으로 나타났다. MDS모델의 적합도는 34.4%로 비교적 낮은 편이나, 대응분석을 통해 확인한 결과 건축심의의 비정형적 자료의 특성에 기인한 것으로 확인하였다. 또한, 본 연구에서 분석한 건축심의 의견과 같은 비정형적 데이터는 심의위원의 주관과 지자체별 양식에 따라 다양한 데이터의 조합과 출현이 이루어지며, 주로 언급되는 단어와 전혀 다른 키워드가 등장할 수 있어 첩도가 낮고 왜도가 높은 확률분포적 특성을 파악할 수 있었다. 본 연구는 일부 한계점이 있으나, 비정형 데이터로서 건축심의 의견의 특징을 도출해내었으며 추후 세부분석을 위한 기초 연구로서 활용 될 수 있을 것이다.

키워드 : 비정형 데이터, 건축심의, 중소규모 공동주택, 다차원 분석법
