

건축도시 정책동향

vol. 94

2024.12

건축·도시 AURI POLICY UPDATES

정책 보도자료 분석을 통한
2024년도 건축·도시 정책 동향



건축도시정책동향 Vol. 94

발 행 : 건축공간연구원(auri)
발행인 : 이 영 범
발행일 : 2024년 12월 31일
I S S N : 2635-5140
기 획 : 건축·도시정책정보센터
주 소 : (30116) 세종특별자치시 가름로 143, 8층
연락처 : 044-417-9697
이 메 일 : leejm@auri.re.kr

[안내] 건축도시정책정보센터에서 운영하는 ‘아우름 사이트(www.aurum.re.kr)’에 접속하면
건축·도시 관련 분야 최신 정책정보를 더욱 빠르게 받아볼 수 있습니다.

Contents

01 ● 건축·도시 분야 정책동향 분석 개요

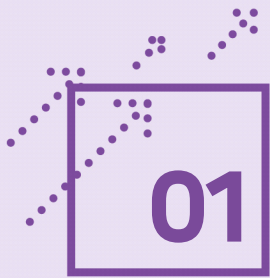
배경 및 목적

분석 방법

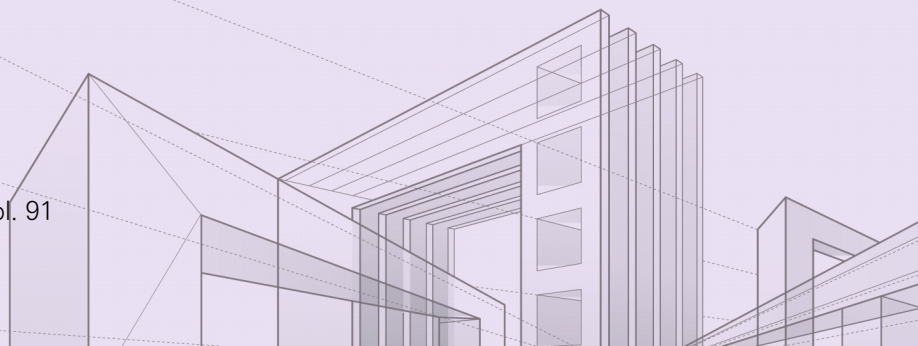
02 ● 2024년도 건축·도시분야 정책동향

키워드 빈도(TF) 및 TF-IDF로 본 건축·도시 정책동향

토픽 모델링으로 본 건축·도시 정책 8대 토픽



건축·도시분야 정책동향 분석 개요

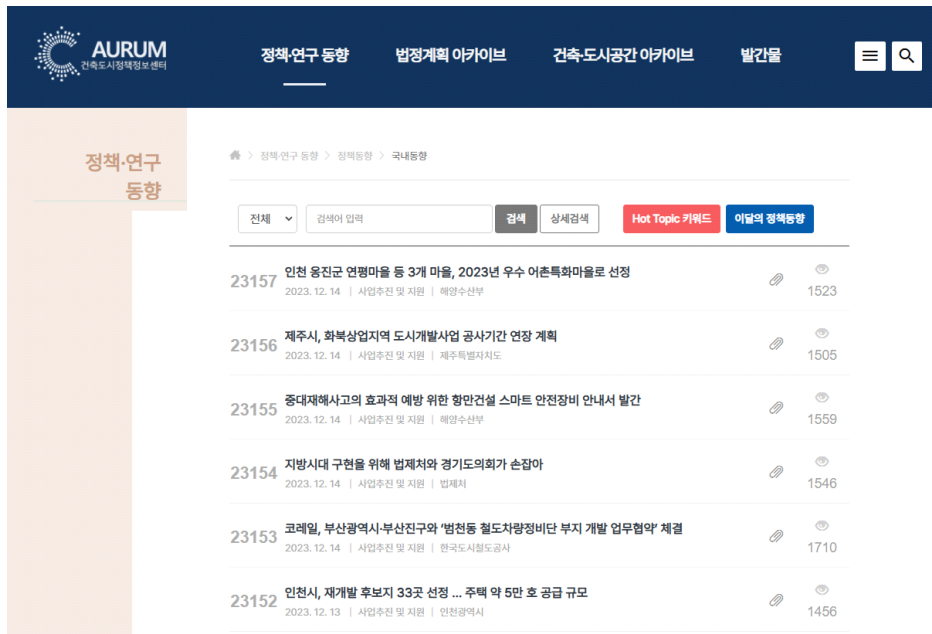


[건축·도시분야 정책동향 분석 개요]

배경 및 목적

■ 배경

- 건축공간연구원 건축·도시정책정보센터에서는 중앙부처(국토교통부, 행정안전부 등)와 지방자치단체에서 제공하는 건축·도시분야 보도자료를 매주 수집하여 AURUM 홈페이지를 통해 대국민 서비스 제공 중



[그림] AURUM 홈페이지 내 정책동향(정책 보도자료) 대국민 서비스 화면

- 중앙부처 및 지방자치단체에서 배포하는 정책 보도자료는 정책·사업·제도에 대한 주요 정보 알림, 홍보 및 확산을 목적으로 하고 있어, 정책동향을 분석할 수 있는 중요한 자료임
- 특히 정책 보도자료는 보도시기, 정책 주체, 정책 세부내용 등에 대한 세부적인 정보가 포함되어 있어, 정책 보도자료가 담고 있는 사회적인 이슈, 정책 방향 등을 가늠할 수 있음

■ 목적

- 정책 보도자료를 분석하여 건축·도시 분야에서 주요하게 다루어지고 있는 키워드 및 토픽을 도출하여 제시함으로써 연관 정책개발 및 정책연구의 방향성을 모색하는 자료를 제공하고자 함

분석 방법

■ 분석대상

- 2024년 1월 1일부터 2024년 12월 31일까지 중앙부처 및 지방자치단체에서 배포된 건축·도시 분야 정책 보도자료

■ 데이터 수집

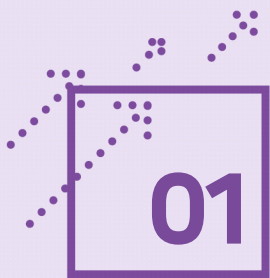
- 파이썬(Python) 기반 웹 크롤러(web crawler)를 개발하여 AURUM 홈페이지(<http://www.aurum.re.kr>)에 2024년도에 등록된 정책 보도자료의 제목, 본문, 게시일, URL 등을 수집
- 총 1,414건의 정책 보도자료를 수집하여 최종 분석데이터로 사용

■ 데이터 전처리

- 수집된 데이터에서 조사, 감탄사, 부사, 대명사, 영문, 특수문자 등의 요소 제거
- 자연어 처리를 위한 의미 최소 단위인 형태소 분석 수행(Python 라이브러리 중 konlpy의 Mecab 사용)
- 형태소 분석을 통해 ‘명사’만을 추출하여 분석 가능한 벡터 형태로 변환

■ 데이터 분석

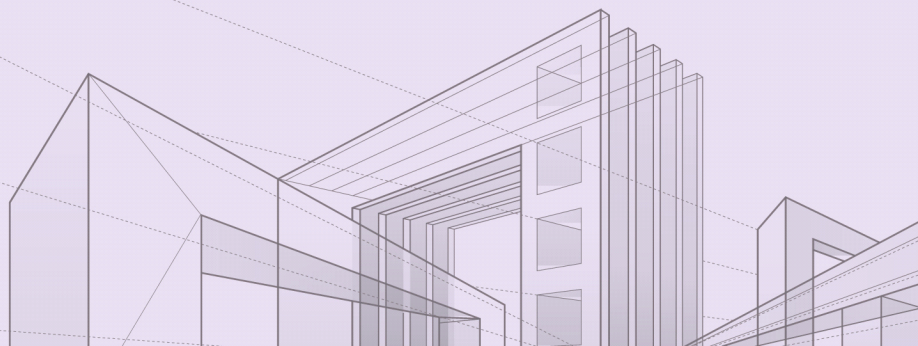
- 벡터 형태로 변환된 데이터를 대상으로 키워드 빈도 분석, TF-IDF 분석 수행
- 일관성 점수(Coherence Score)를 토대로 최적 토픽 수를 결정하여 LDA 토픽모델링 분석 수행



2024년도

건축·도시분야 정책동향

2024년도 중앙부처 및 지방자치단체의 정책 보도자료를 토대로
8개의 토픽을 도출하여 제공합니다.



[2024년도 건축·도시분야 정책동향]

키워드 빈도(TF) 및 TF-IDF로 본 건축·도시 정책동향

워드클라우드(WORD-CLOUD)

- 2024년도 건축·도시분야 정책 보도자료의 핵심 키워드를 살펴보기 위해 해당 용어의 출현 빈도 (TF)를 토대로 한 워드클라우드 생성
- “사업”, “계획”, “도시”, “주택”, “지역”, “지원”, “추진”, “시설”, “정비”, “조성” 등이 상위에 랭크



[그림] 2024년도 건축·도시분야 정책 보도자료 워드클라우드

TF 및 TF-IDF 분석

[용어설명]

- TF(Term Frequency)는 단어의 출현 빈도이며 TF-IDF(Term Frequency-Inverse Document)는 단어의 중요도를 반영한 가중치임
- TF가 높고 TF-IDF가 높다면, 정책 보도자료 전체에서 출현빈도가 높고, 그 전체 중 일부에서의 출현빈도가 매우 높음을 의미
- TF가 낮고 TF-IDF가 높다면, 정책 보도자료 전체에서 출현빈도가 낮으나, 그 빈도가 일부 정책 보도자료에 집중되어 나타남을 의미

2024년도 건축·도시분야 정책동향

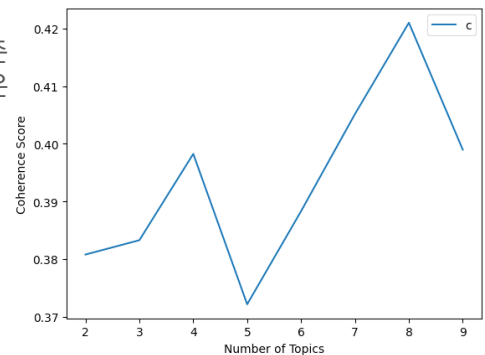
- TF(출현빈도)에서는 “사업(6314)”, “계획(5018)”, “도시(4832)”, “주택(4310)”, “지원(3250)” 등의 순으로 정책 보도자료에서 많이 언급
- TF-IDF 분석 결과, “사업(77.1)”, “도시(67.0)”, “주택(83.7)”, “계획(55.6)”, “정비(46.0)” 등의 순으로 중요도가 높게 도출되어, TF(출현빈도)와 비교 시 전반적으로 유사한 경향
- TF(출현빈도)가 높으며, TF-IDF도 높은 단어는 “사업”, “도시”, “계획”, “주택”, “지역” 등으로, 특정 정책 보도자료에서 매우 집중적으로 자주 나타나고 있어 건축·도시 분야의 정책 보도자료에서 도시 및 지역을 대상으로 사업 추진 및 계획 수립에 관한 내용이 많은 것으로 판단됨

[표] TF(출현빈도)와 TF-IDF의 상위 20위를 차지한 단어

TF(출현빈도)						TF-IDF					
순위	단어	TF	순위	단어	TF	순위	단어	TF-IDF	순위	단어	TF-IDF
1	사업	6314	11	공공	1850	1	사업	77.1	11	시설	34.0
2	계획	5018	12	개발	1803	2	도시	67.0	12	조성	33.0
3	도시	4832	13	관리	1774	3	주택	66.2	13	추진	32.1
4	주택	4310	14	건축	1711	4	계획	55.6	14	개발	30.3
5	지역	3308	15	건설	1693	5	정비	46.0	15	주거	29.3
6	지원	3250	16	예정	1691	6	지역	43.7	16	선정	29.1
7	추진	2494	17	선정	1579	7	지원	42.8	17	관리	28.7
8	시설	2317	18	국토	1564	8	건축	37.5	18	재생	28.3
9	정비	2185	19	주거	1559	9	건설	35.3	19	국토	28.1
10	조성	2111	20	공간	1523	10	공공	34.3	20	공간	27.9

토픽 수 결정

- 토픽 수는 단어 간 연결 정도를 평가하는 응집도를 기준으로 분석하여 응집도(Coherence Score)가 가장 높은 8개로 결정



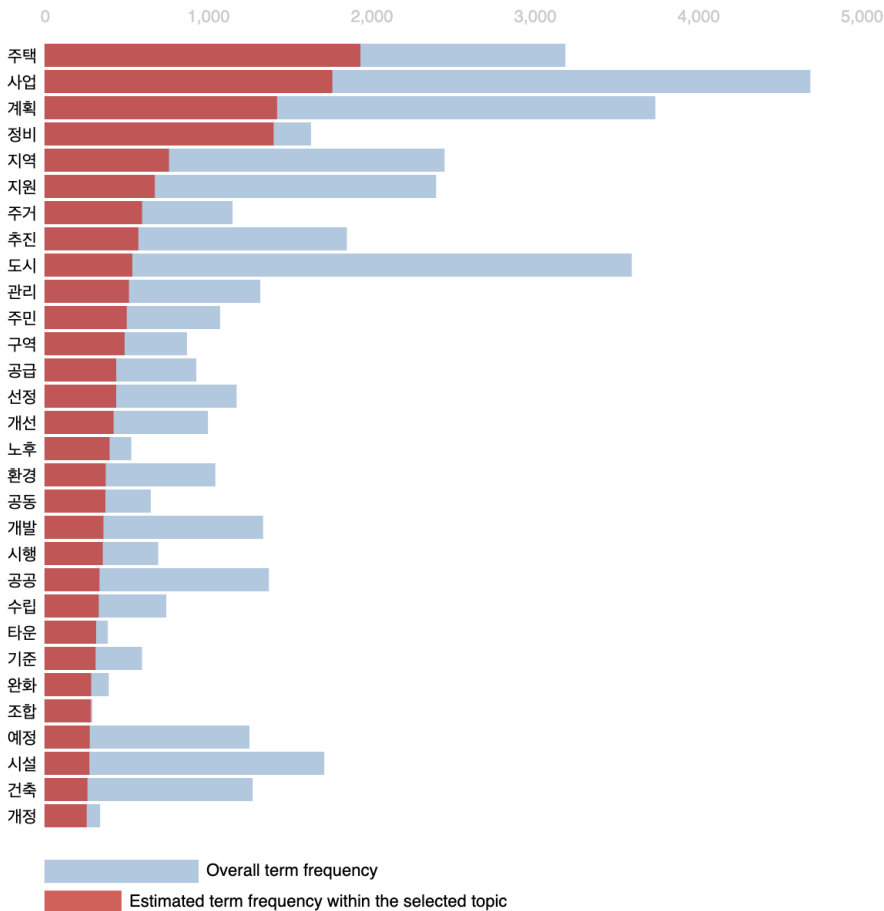
[그림] 토픽 수 별 일관성 점수

[2024년도 건축·도시분야 정책동향]

토픽 모델링으로 본 건축·도시 정책 8대 토픽

토픽1 노후 주거환경 개선

- 관련어는 “주택(0.033)”, “사업(0.030)”, “계획(0.024)”, “정비(0.024)”, “지역(0.013)”, “지원(0.012)”, “주거(0.010)”, “추진(0.009)”, “도시(0.009)”, “관리(0.009)” 등으로 다른 토픽과 가장 차별되는 관련어는 “주택”, “정비”



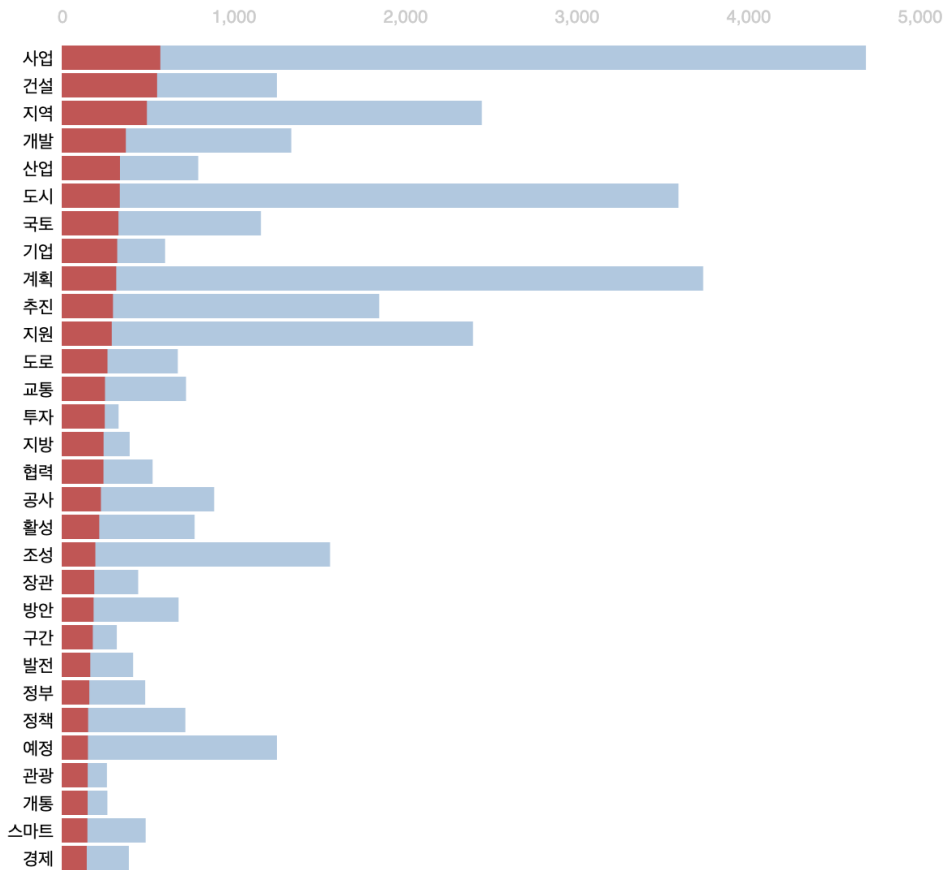
1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))]] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

[그림] 토픽1. 노후 주거환경 개선 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽2 지역산업 육성 및 기업 유치

- 관련어는 “사업(0.016)”, “건설(0.015)”, “지역(0.013)”, “개발(0.010)”, “산업(0.09)”, “도시(0.009)”, “국토(0.009)”, “기업(0.009)”, “계획(0.009)”, “추진(0.008)” 등으로 다른 토픽과 가장 차별되는 관련어는 “지역”, “산업”, “기업”



Overall term frequency
Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

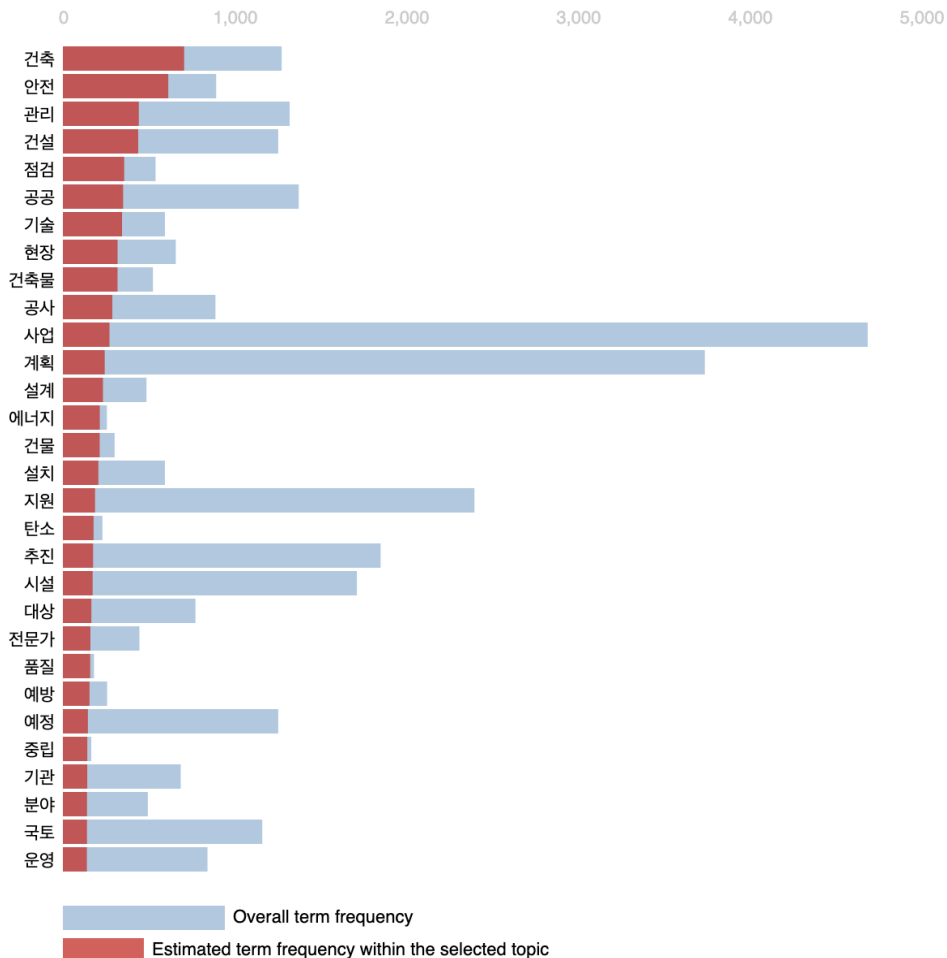
[그림] 토픽2.지역산업 육성 및 기업 유치 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽3

건축 안전관리

- 관련어는 “건축(0.021)”, “안전(0.018)”, “관리(0.013)”, “건설(0.013)”, “점검(0.010)”, “공공(0.010)”, “기술(0.010)”, “현장(0.009)”, “건축물(0.009)”, “공사(0.008)” 등으로 다른 토픽과 가장 차별되는 관련어는 “건축”, “안전”, “관리”, “점검”



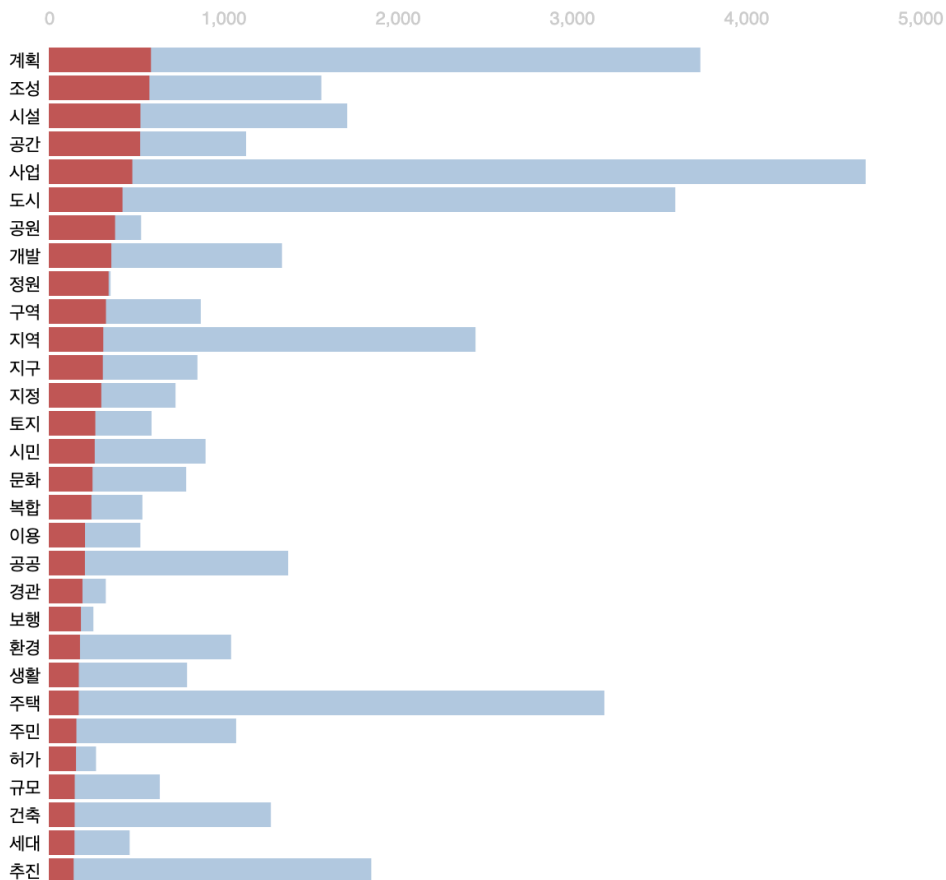
[그림] 토픽3. 건축 안전관리 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽4

생활인프라 조성

- 관련어는 “계획(0.017)”, “조성(0.017)”, “시설(0.015)”, “공간(0.015)”, “사업(0.014)”, “도시(0.012)”, “공원(0.011)”, “개발(0.010)”, “정원(0.010)”, “구역(0.010)” 등으로 다른 토픽과 가장 차별되는 관련어는 “시설”, “공원”, “정원”



Overall term frequency

Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)

2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

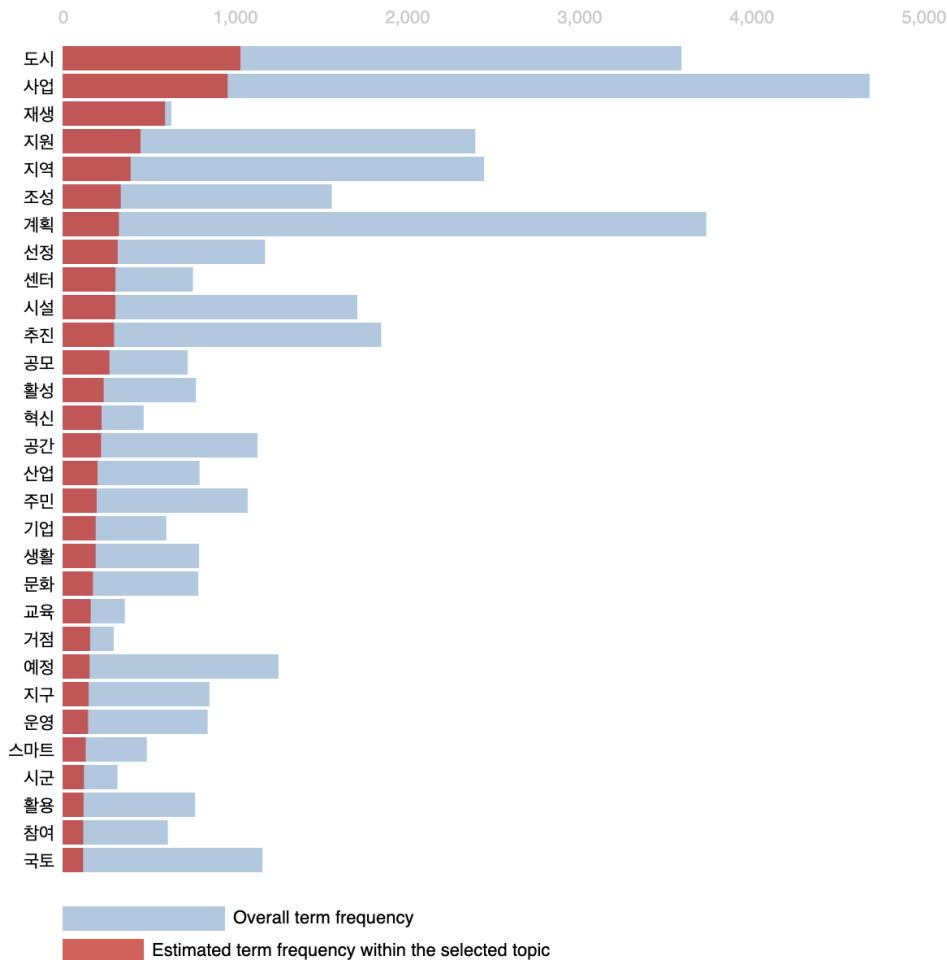
[그림] 토픽4. 생활인프라 조성 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽5

도시 재생

- 관련어는 “도시(0.031)”, “사업(0.029)”, “재생(0.018)”, “지원(0.013)”, “지역(0.012)”, “조성(0.010)”, “계획(0.010)”, “선정(0.010)”, “센터(0.009)”, “시설(0.009)” 등으로 다른 토픽과 가장 차별되는 관련어는 “재 생”



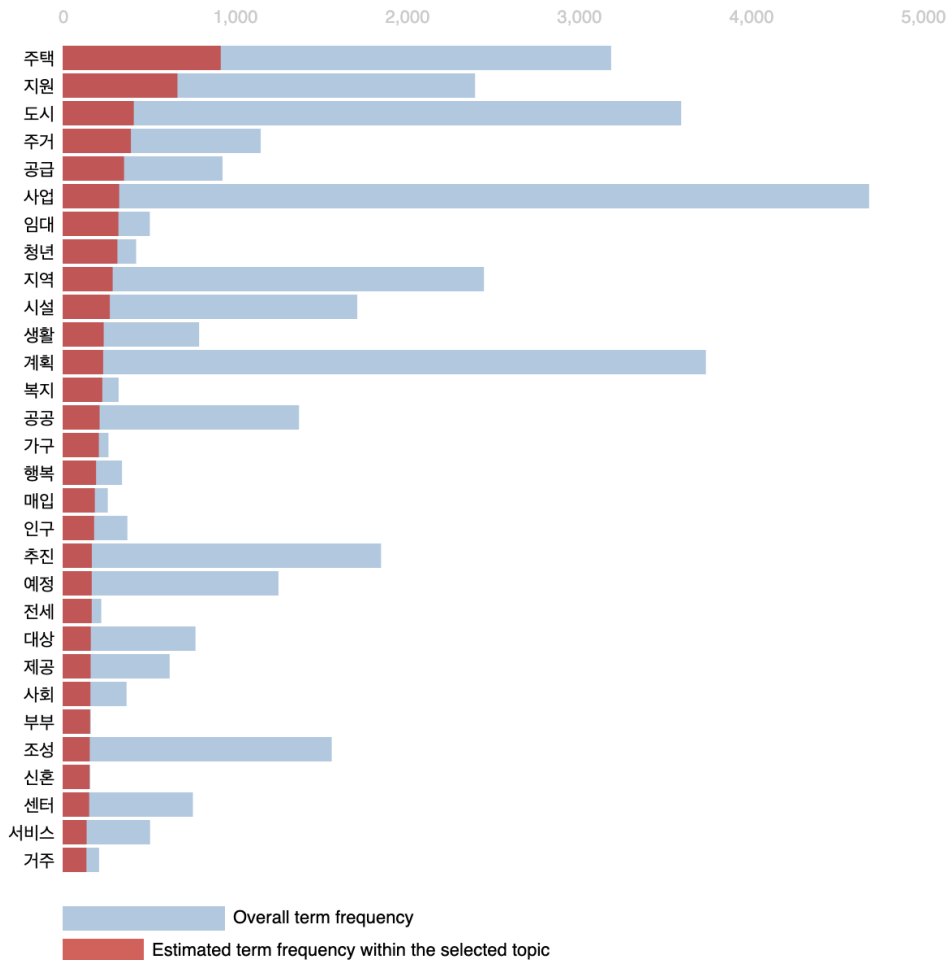
1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = $\lambda * p(w | t) + (1 - \lambda) * p(w | t)/p(w)$; see Sievert & Shirley (2014)

[그림] 토픽5. 도시 재생 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽6 주거취약계층 주거 안정화

- 관련어는 “주택(0.029)”, “지원(0.021)”, “도시(0.013)”, “주거(0.012)”, “공급(0.011)”, “사업(0.010)”, “임대(0.010)”, “청년(0.009)”, “지역(0.009)”, “시설(0.009)” 등으로 다른 토픽과 가장 차별되는 관련어는 “주택”, “임대”, “청년”, “신혼부부” 등

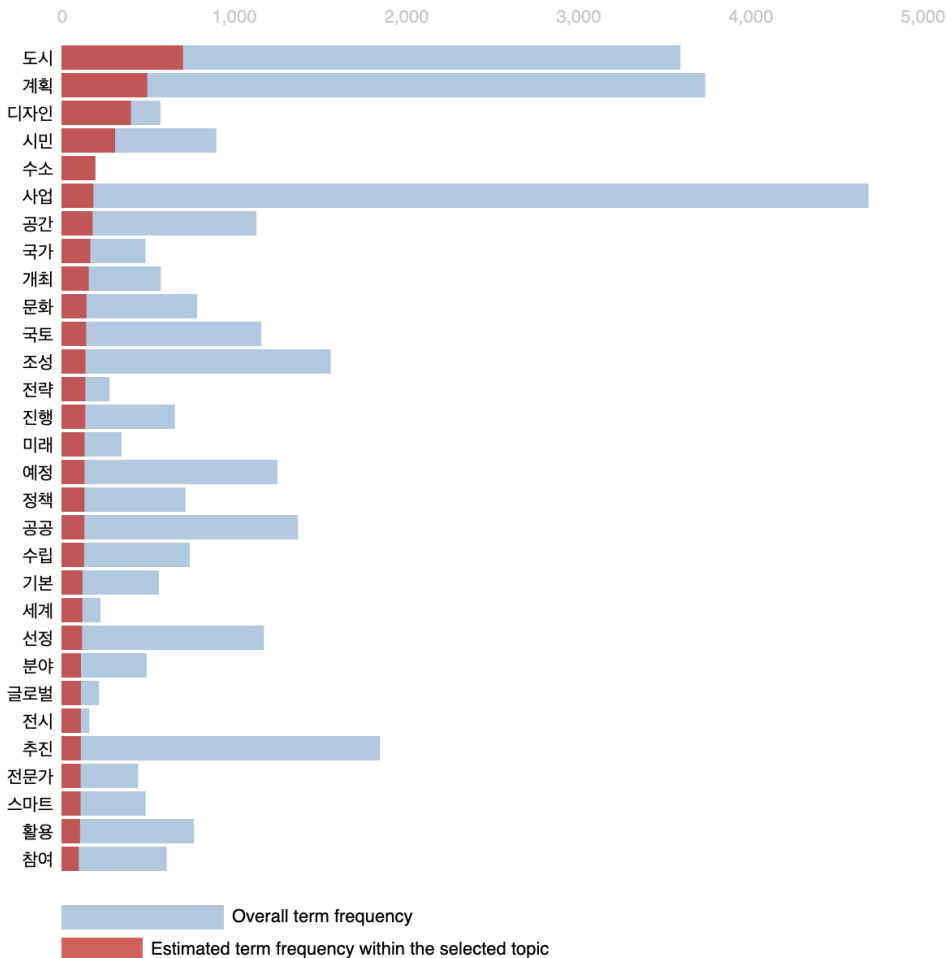


[그림] 토픽6. 주거취약계층 주택 공급 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽7 도시 경쟁력 강화

- 관련어는 “도시(0.027)”, “계획(0.019)”, “디자인(0.016)”, “시민(0.012)”, “수소(0.008)”, “사업(0.007)”, “공간(0.007)”, “국가(0.006)”, “개최(0.006)”, “문화(0.006)” 등으로 다른 토픽과 가장 차별되는 관련어는 “디자인”, “개최”, “문화”



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))]; for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

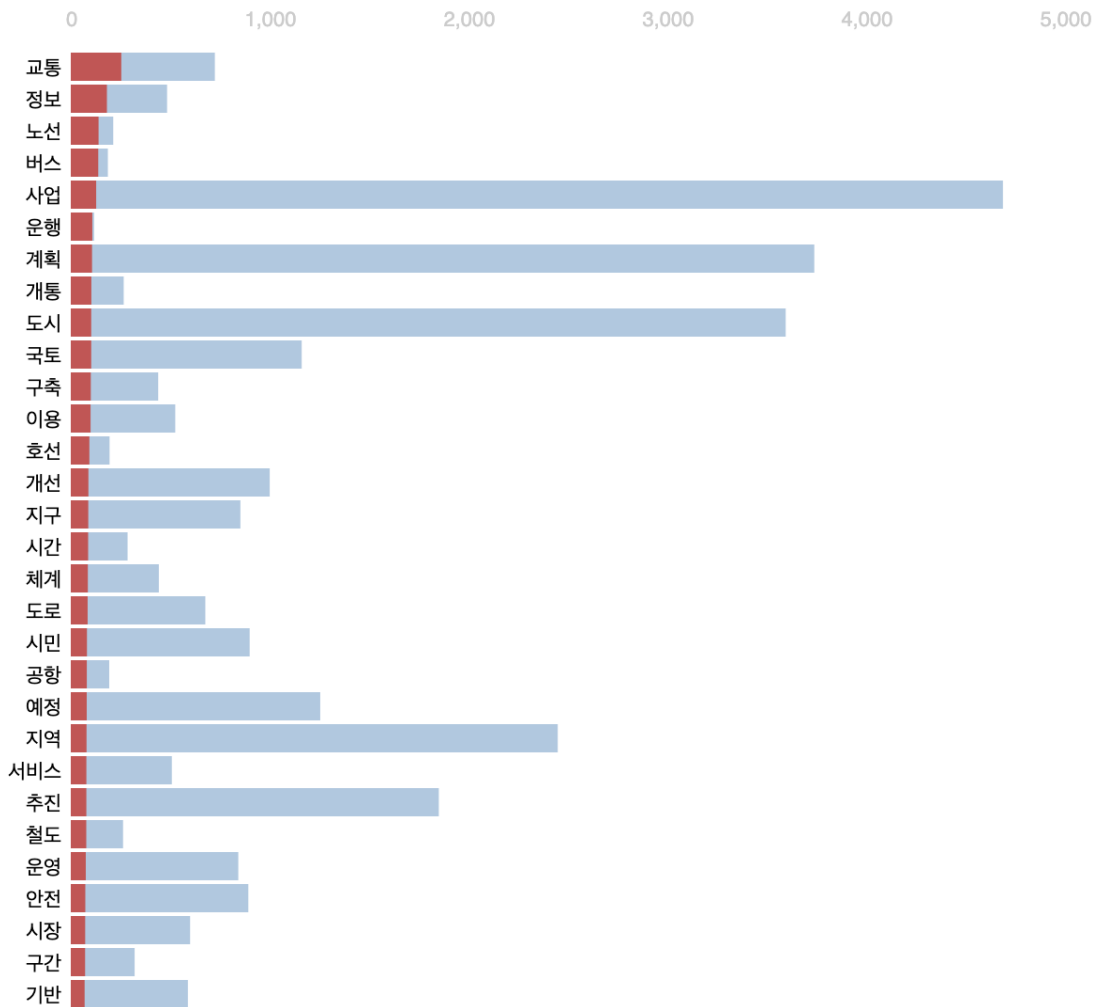
[그림] 토픽7. 도시 경쟁력 강화 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

토픽8

교통인프라 조성

- 관련어는 “교통(0.015)”, “정보(0.011)”, “노선(0.008)”, “버스(0.008)”, “사업(0.007)”, “운행(0.006)”, “계획(0.006)”, “개통(0.006)”, “도시(0.006)”, “국토(0.006)” 등으로 다른 토픽과 가장 차별되는 관련어는 “교통”, “노선”, “버스”, “운행”, “개통”



[그림] 토픽8. 교통인프라 조성 관련 단어 상위 30개

[2024년도 건축·도시분야 정책동향]

■ 토픽별 출현비중

- 토픽별로 등장할 확률을 분석한 결과, 토픽1. 노후 주거환경 개선이 21.4%로 가장 많이 등장하며, 그 다음으로 토픽2. 지역산업 육성 및 기업 유치(13.6%), 토픽3. 건축 안전관리(12.6%) 등의 순으로 나타남



[그림] 토픽 분포도

[표] 토픽별 출현비중

토픽	출현비중
토픽1, 노후 주거환경 개선	21.4%
토픽2, 지역산업 육성 및 기업 유치	13.6%
토픽3, 건축 안전관리	12.6%
토픽4, 생활인프라 조성	12.6%
토픽5, 도시 재생	12.3%
토픽6, 주거취약계층 주거 안정화	11.7%
토픽7, 도시 경쟁력 강화	9.5%
토픽8. 교통 인프라 조성	6.3%

건축도시 정책동향

건축·도시정책정보센터

건축·도시정책정보센터는 국내 건축도시 분야에서 생산되는 다양한 기록 자료와 정책 정보 및 학술 연구정보 구축을 통하여, 건축물 및 공간환경 정책 수립과 사업추진을 위한 지적 토대를 제공하기 위해 설립되었습니다.

2007년 건축기본법 제정 이후 건축물 및 공간환경 조성을 위하여 추진되는 사업과 관련한 각종 기록 자료에 대한 관리의 필요성이 증가함에 따라, 건축 정책수립을 지원하고, 학술문헌 및 연구정보 구축을 통해 건축·도시 분야의 관련 정보와 지식을 체계적으로 축적하고 보급하는 역할을 수행하고 있습니다.



건축공간연구원
Architecture & Urban Research Institute

30116 세종특별자치시 가름로 143, KT&G세종타워B 8층, 건축도시정책정보센터
architecture and urban policy information center,
KT&G Sejong Tower B 8F, Gareum-ro, Sejong-si, 30116, Korea