

강원권 주요 도시 역세권개발 담론 및 쟁점 분석 : 춘천, 원주, 강릉역세권을 중심으로*

Evolution and Issues in Station-Area Development Discourse in Major Cities of
Gangwon Province: Chuncheon, Wonju, and Gangneung

정 진 태 (Jeong, Jintae)**

김 서 정 (Kim, Seojeong)***

이 천 재 (Lee, Cheonjae)****

< Abstract >

This study examines how station-area development discourse has evolved in three major cities of Gangwon Province—Gangneung, Wonju, and Chuncheon—using news big data from 2010 to 2026. Articles were retrieved from the BIGKinds database of the Korea Press Foundation by combining the keyword “station area” (yeoksegwon) with each city name, yielding 8,991 articles. Because keyword retrieval also captures unrelated coverage, rule-based screening removed topically irrelevant articles and two alternative relevance criteria were applied to test sensitivity, producing a common corpus of 6,738 articles on which all analyses—TF-IDF, semantic network and frame analysis, and KNU-lexicon sentiment analysis—were conducted at the city-by-year level. Although collected under an identical keyword, discourse composition diverged markedly across cities. Gangneung shifted from the Olympics and KTX opening toward regional transport networks; Wonju sustained real estate and railway infrastructure discourses in parallel; Chuncheon dispersed into culture and tourism after the Gyeongchun Line opening before development agendas cyclically re-emerged. Neutral articles predominated in all three cities, with positive shares exceeding negative ones (Wonju 40.1%, Gangneung 35.3%, Chuncheon 31.3%); because absolute scores are sensitive to document length and lexicon bias, interpretation was confined to relative differences. The cities were accordingly classified as tourism-event, industrial-planned-city, and capital-region-accessibility growth types.

Keyword : Transit-Oriented Development(TOD), Discourse Analysis, TF-IDF, Sentiment Analysis, News Big Data

* 본 논문은 2026년도 한국부동산분석학회 상반기 학술대회(2026년 5월 29일, 연세대학교 미래캠퍼스)에서 발표한 논문을 수정·보완하였음.

** 강원대학교 부동산학과 박사과정, tae80@hanmail.net, 주저자

*** 강원대학교 부동산학과 석사과정, asdf0625@kangwon.ac.kr, 공동저자

**** 본 학회 종신회원, 강원대학교 부동산학과 조교수, cheonjae.lee@kangwon.ac.kr, 교신저자

I. 서론

역세권(驛勢圈)은 철도역을 중심으로 교통 접근성의 이점을 누릴 수 있는 공간적 범위를 가리키며¹⁾ 대중교통중심개발(Transit-Oriented Development, TOD)의 핵심 대상으로 주목받아 왔다. 국내에서 역세권 개발은 2000년대 이후 수도권 광역철도 확충, 고속철도 개통, 도시재생 정책과 맞물려 활발히 추진됐다. 최근 서울시의 '역세권 직·주·락 활성화' 전략(2022)이나 「역세권의 개발 및 이용에 관한 법률」(2024) 개정에서 이러한 흐름이 제도적으로 공고해지고 있음이 확인된다.

그러나 이러한 정책적 확장에도 불구하고 역세권이 실제 지역사회에서 어떻게 이야기되고 있으며 어떤 쟁점이 부각되고 어떤 정서가 형성되는지를 다룬 담론적 분석은 상대적으로 드물었다. 기존 역세권 연구는 대체로 부동산 가격 파급효과나 토지이용 밀도와 같은 계량적·공간적 접근에 집중해 왔다(성현곤, 2011). 이에 역세권 개발은 단순한 물리적 인프라 조성을 넘어 지역의 정체성과 주민 기대, 정치적 의제가 교차하는 사회적 과정이라는 점에서 재검토가 필요하다. 2010년대 중반 이후 텍스트마이닝 기법이 사회과학 전반에 확산하면서 뉴스 기사를 대량으로 수집·분석해 특정 주제의 담론 구조를 파악하는 연구가 늘어난 것도 이러한 문제의식과 무관하지 않다(이은별 외, 2017; 주영희 외, 2025).

본 연구는 역세권 담론이 철도 인프라의 유형만으로 결정되기보다 철도 개통 시점과 도시 기능, 주요 개발 사업 및 정책 환경이 결합하며 지역별로 상이하게 구성된다는 관점에서 시작되었다. TOD가 고정된 설계 원리를 넘어 정책과 계획으로 이동하며 지역적 맥락에 따라 재구성될 수 있다는 점에 착안했다(Jamme et al., 2019). 이러한 관점에서 강원권 거점도시인 강릉·원주·춘천의 역세권 개발 담론을 뉴스 빅데이터로 시계열 분석하고 지역 간 담론 구조의 차이와 변화 양상을 규명하는 것을 목적으로 한다.

세 도시를 분석 대상으로 선택한 이유는 철도 인프라의 유형과 도입 시점, 도시 기능이 각기 다르기 때문이다. 강릉은 경강선 KTX 개통(2017)과 평창동계올림픽(2018)이라는 국가적 이벤트를 경험한 도시다. 원주

는 혁신도시·기업도시 조성과 중앙선 복선전철화(2021)가 맞물린 계획도시의 성격을 지닌다. 춘천은 경춘선 복선전철 개통(2010)이 강원도 내에서 비교적 이른 시기에 이루어져 장기적 담론 변화를 관찰할 수 있다. 세 도시 모두 강원권의 주요 거점도시이면서도 철도 인프라의 성격과 도시 기능이 상이하다는 점에서 담론의 지역별 분화 양상을 비교하기에 적합한 사례로 판단된다.

본 연구의 공간적 범위는 강원도 강릉시·원주시·춘천시의 역세권으로 한정한다. 세 도시는 각각 고속철도(강릉 경강선 KTX), 복선전철(원주 중앙선), 광역철도(춘천 경춘선)라는 상이한 유형의 철도 인프라를 기반으로 역세권이 형성되어 있다. 따라서 교통 인프라 유형에 따른 담론의 차별적 구성을 비교하기에 적합하다. 시간적 범위는 경춘선 복선전철 개통(2010년 12월)부터 2026년 3월까지로 설정하였다. 세 도시의 실질적 분석 시작 시점은 자료의 가용 여부에 따라 춘천 2010년, 강릉 2012년, 원주 2016년으로 상이하다. 내용적 범위는 역세권 개발을 다룬 언론 보도의 담론적 구조와 정서적 방향성에 한정한다. 역세권의 물리적 공간 범위에 대한 정의나 부동산 가격 변동의 계량 분석, 주민 인식 설문 등 기존 역세권 연구의 주요 영역은 분석 대상에서 제외한다.

분석에는 TF-IDF와 KNU 한국어 감성사전 기반 감성분석을 결합하여 담론의 내용과 정서를 동시에 포착한다. 아울러 의미 연결망 분석과 엔트만(Entman) 프레임 분석으로 담론의 관계구조를 규명한다. 이로써 '무엇이 이야기되었는가'를 넘어 '어떤 관계와 프레임 속에서 이야기되었는가'까지 분석범위를 확장하고자 한다. 논문의 구성은 다음과 같다. II장에서는 역세권 개발과 TOD에 관한 이론적 배경과 텍스트마이닝의 방법론적 근거를 검토한다. III장에서는 자료수집·정제 절차와 분석기법 설계를 설명한다. IV장에서는 도시별·시기별 담론구조와 감성변화를 제시한다. V장에서는 연구 결과의 함의와 한계를 논의한다.

1) 「역세권의 개발 및 이용에 관한 법률」 제2조는 역세권을 철도역과 그 주변지역 중 국토교통부령으로 정하는 범위의 지역으로 정의한다. 본 연구는 이러한 법적 정의를 전제하되, 분석에서는 물리적 거리 범위가 아니라 언론 보도에서 역세권으로 지칭되는 담론적 범위를 대상으로 한다.

II. 이론적 고찰

1. 역세권 개발과 대중교통중심개발(TOD)

역세권 개발의 이론적 기반은 Calthorpe(1993)가 제안한 대중교통중심개발(TOD)에서 출발한다. TOD는 대중교통 결절점을 중심으로 고밀도·복합용도의 개발을 유도하고 보행 친화적 환경을 조성함으로써 자동차 의존도(automobile dependency)를 낮추는 계획 원리다. Cervero(2004)는 이를 밀도(density), 다양성(diversity), 설계(design)의 3D 원칙으로 체계화한 바 있다. 그러나 한국의 역세권 개발은 유럽·미국의 TOD 모형과는 상이한 맥락에서 전개됐다. 수도권에서는 광역전철 개통에 따른 부동산 가격 상승효과가 역세권 연구의 주된 관심사였던 반면(성현근, 2011), 비수도권에서는 KTX 개통이나 철도 고속화에 따른 지역개발 효과에 초점이 맞추어졌다. 조재욱·우명제(2014)는 KTX 경부선·경전선 개통이 정차 도시의 지역내총생산에 미친 영향을 분석하였다. 허재완 외(2018)는 KTX 정차역 입지가 수도권 인구 집중 완화에 기여하는지를 다중회귀모형으로 검증하였다. 강원권을 대상으로 한 이제용·김영표(2020)의 연구 역시 KTX 강릉선 개통이 관광산업과 지역 이미지에 미친 영향을 실증하였다. 다만 설문 기반의 인식 조사에 머물렀다는 한계를 지닌다.

이처럼 TOD는 본래 물리적 계획원리로 제시되었다. 그러나 국가와 지역의 정책 과정에 적용되면서 주택공급, 도시재생, 관광개발, 광역교통 및 부동산개발 등 다양한 목표와 결합해 왔다. Jamme et al.(2019)이 지적하듯 TOD는 정책 및 계획 담론 속에서 선택적으로 해석되고 재구성될 수 있는 개념이다. 이러한 관점에서 역세권은 거리 범위로만 규정되는 물리적 공간이 아니다. 교통시설과 개발사업, 지역경제와 주민 기대가 언론 및 정책 담론을 통해 결합되는 사회적 공간으로 이해할 수 있다. 이에 본 연구는 역세권 개발을 물리적 계획 대상이 아닌 담론적 구성물로 다룬다.

2. 텍스트 마이닝을 통한 담론 및 감성분석

뉴스 기사는 특정 시기에 사회적으로 부각된 의제와 쟁점을 반영하는 비정형 텍스트 데이터로서 담론분석의 주요 자료원이 된다. 최근 한국에서는 빅카인즈

(BIGKinds) 플랫폼을 활용하여 대규모 뉴스 코퍼스에 텍스트마이닝 기법을 적용하는 연구가 늘고 있다. 이은별 외(2017)는 외국인 마을 관련 보도를 분석해 특정 지역에 대한 매체의 편향적 보도 관행과 공간 담론의 구체화를 실증하였다. 주영희 외(2025)는 TF-IDF와 네트워크·CONCOR 분석을 적용해 순환경제 담론의 구조를 규명하였다.

TF-IDF는 단순 빈도(TF)에 역문서빈도(IDF)를 곱하여 특정 문서(또는 문서 집합)에서 상대적으로 변별력이 높은 키워드를 도출하는 가중치 산출 방식이다(Salton and Buckley, 1988). 전체 코퍼스에 보편적으로 등장하는 키워드의 가중치를 낮춘다는 점에서 특정 시기·지역에서 고유하게 부각된 의제를 포착하는데 유리하다. 본 연구에서는 ‘지역×연도’를 하나의 문서 단위로 묶어 TF-IDF를 산출함으로써 동일 지역 내에서도 연도에 따라 어떤 키워드가 차별적으로 부각되는지를 파악하고자 하였다. 의미연결망 분석은 동일 기사에 함께 등장하는 키워드의 관계를 네트워크로 구성하여 중심 키워드와 담론군집을 확인하는 방법이다. CONCOR 분석은 연결패턴이 유사한 키워드를 군집화함으로써 개별 단어의 빈도를 넘어 담론의 구조적 등위성을 파악한다. 본 연구는 이러한 군집을 Entman의 프레임 요소와 연결하여 무엇이 문제로 규정되고 원인이 어떻게 진단되며 어떤 정책처방이 제시되는지를 해석한다.

감성분석은 텍스트에 내포된 긍정·부정 등의 정서적 방향성을 자동으로 판별하는 기법이다. 사전 기반 접근과 기계학습 기반 접근으로 대별된다. 본 연구는 군산대학교 Data Intelligence Lab이 개발한 KNU 한국어 감성사전을 활용하였다(박상민 외, 2018). KNU 감성사전은 표준국어대사전의 뜻풀이를 Bi-LSTM 모델로 분류하여 14,843개의 감성 어휘에 극성과 정도값을 부여한 사전이다. 특정 도메인에 편향되지 않은 범용 사전이라는 점에서 뉴스 텍스트 분석에 적합하다. 다만 사전 기반 감성분석은 문맥을 고려하지 못한다는 근본적 한계를 지닌다. 특히 KNU 감성사전은 ‘해(害)·화(禍)·병(病)’ 등 한 글자 한자어를 부정어로 판정하는 경우가 빈번하다. 사건·사고·갈등 보도에 수반되는 부정어가 양적으로 우세한 뉴스 텍스트의 특성상 감성점수가 구조적으로 부정 쪽으로 치우칠 가능성이 있다. 이에 본 연구는 감성점수의 절대값을 주된 태도나 정책 성과로 직접 해석하지 않았다. 대신 기사 길이

보정과 인간 라벨링, 언어모델 비교를 거쳐 지역·시기 간 상대적 차이를 중심으로 해석 범위를 한정하였다.

3. 선행연구 및 연구의 차별성

앞서 검토한 연구를 종합하면(<표 1>), 기존 역세권 연구는 TOD 이론·설계 연구, 부동산 가격효과 연구, KTX 지역효과 연구, 뉴스 빅데이터 담론분석, 감성분석 기법 연구의 다섯 영역으로 구분된다. 이 중 앞의 세 영역인 TOD 이론, 부동산 가격효과, KTX 지역효과는 공통적으로 물리적 공간분석이나 계량적 가격·경제효과에 초점을 맞춘다. 그래서 역세권이 지역 담론 속에서 어떻게 의미화되고 시간에 따라 변화하는지를 추적한 사례는 드물다.

한편 뉴스 빅데이터를 활용한 담론분석은 TF-IDF와 네트워크·CONCOR 분석을 통해 특정 주제의 담론 구조를 파악해 왔다. 그러나 담론의 정서적 방향성까지 동시에 분석한 사례는 제한적이다. 감성분석 분야에서는 KNU 한국어 감성사전이 범용적 뉴스 텍스트 분석 도구로 활용되고 있다. 그러나 이를 역세권 담론에 적용한 연구는 현재까지 이루어지지 않았다. 본 연구의 차별성은 첫째, 역세권을 물리적 개발대상이 아니라 지역별로 상이하게 의미화되는 담론적 공간으로 분석한다. 둘째, TF-IDF에 의미연결망·CONCOR

·프레임 분석을 결합하여 키워드의 관계와 정책의제를 함께 규명한다. 셋째, 감성분석에 표본정제와 점수 정규화, 인간 라벨링 및 KLUE-BERT 비교검증을 적용해 방법론적 신뢰성을 확보하고자 시도했다는 점이다.

III. 연구방법론

1. 분석 대상 및 자료수집

본 연구의 분석 절차는 자료 구축, 담론구조 분석, 감성분석 및 검증, 통합 및 유형화의 네 단계로 구성되며 전체 흐름은 <그림 1>과 같다. 이하에서는 각 단계의 세부 절차를 순서대로 설명한다.

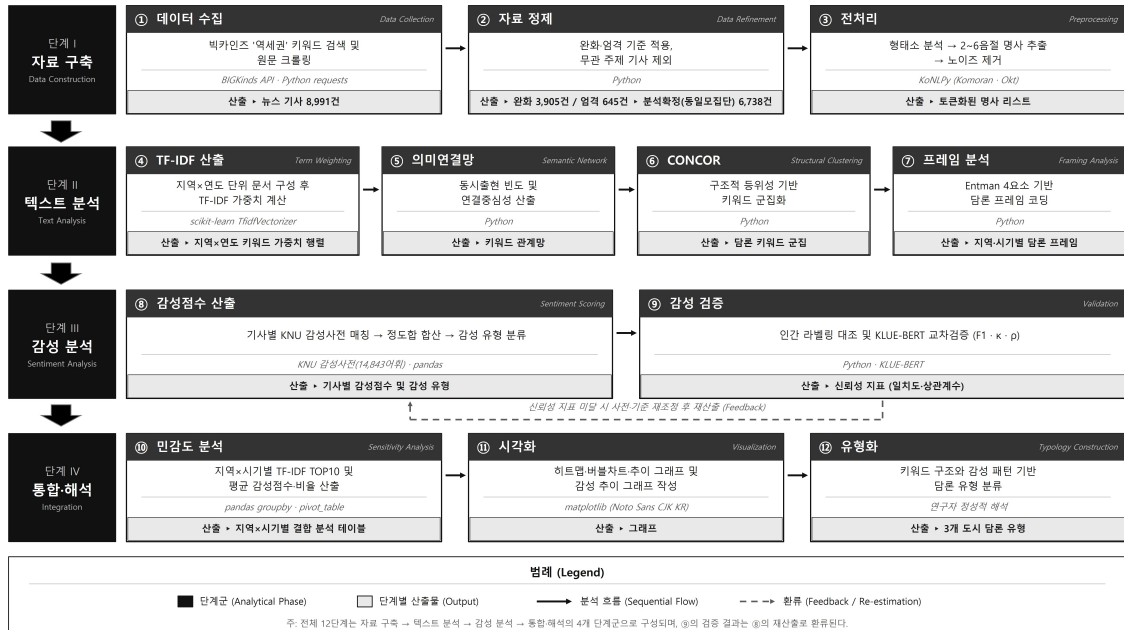
분석 대상은 강원도 강릉시·원주시·춘천시의 역세권 개발 관련 뉴스 기사이다. 빅카인즈(BIGKinds) 플랫폼에서 ‘역세권’, ‘역세권개발’ 키워드와 각 도시명(강릉/원주/춘천)을 AND 조건으로 검색하여 기사를 수집하였다.²⁾ 다만 빅카인즈 검색은 제목뿐 아니라 본문 전체를 대상으로 매칭이 이루어진다. ‘역세권’이라는 단어가 본문 중 한 번이라도 등장하면 검색 결과에 포함된다. 이 때문에 역세권 개발과 직접적 관련이 없는 기사도 함께 수집되는 문제가 발생할 수 있다.

<표 1> 선행연구 및 연구의 차별성

연구 영역	저 자	연구내용 및 방법
역세권 TOD	박천보(2016)	국외 사례 비교분석 - TOD 3유형 분류, 발전전략
	Jamme et al.(2019)	25년간 TOD 문헌 회고 - 설계 개념에서 정책·계획·실행으로의 이행
	Ibraeva et al.(2020)	체계적 문헌고찰 - TOD 개념·효과·계획 쟁점 종합
TOD와 부동산	성현곤(2011)	다수준 특성가격모형 - 서울9호선 아파트
	노재준·임재빈(2024)	공간계량모형 - 대전 1호선 17년 실거태가, 장기 공간파급효과
	Rojas(2024)	헤도닉 이중차분 - 스웨덴 말뾰 시티터널 신설역, 직접·간접효과 분리
KTX 지역효과	조재욱·우명제(2014)	개통 전후 지역경제·균형발전 실증분석 - 경부선·경전선
	이재용·김영표(2020)	설문 기반 영향관계 분석 - 강릉선 거리속성·관광·지역이미지
	하미란·진장익(2024)	실증분석 - KTX 개통의 지역 인구변화 효과
뉴스 빅데이터 담론	Kwak et al.(2020)	프레임 분류기(딥러닝) - NYT 150만 건 프레임 시계열
	조운승·이태경(2021)	TF-IDF, 네트워크, CONCOR - 빅카인즈 뉴스 기사
	주영희 외(2025)	빈도·TF-IDF·N-gram·네트워크·CONCOR - 빅카인즈
감성 분석	온병원·박상민·나철원(2018)	Bi-LSTM, 사전 기반 - 표준국어대사전 뜻풀이
	최창식·임영호(2021)	KNU 감성사전 - 10개 일간지 정치기사 감성지수

2) 빅카인즈는 한국언론진흥재단이 운영하는 뉴스 빅데이터 분석 시스템으로 1990년 이후 국내 언론사 기사를 제공한다.

<그림 1> 본 연구의 분석 절차



출처: 본 그림은 생성형 AI(Antropic, Claude Opus 5, 2026년 6월 활용)를 이용하여 초안을 생성한 후, 저자가 내용을 검토·수정하여 최종 작성하였다.

이 문제는 다음 절에서 다루는 정제 절차의 필요성으로 이어진다.

검색 기간은 2010년 1월부터 2026년 3월까지이다. 수집 대상 언론사는 경향신문 · 국민일보 · 내일신문 · 동아일보 · 문화일보 · 서울신문 · 세계일보 · 아시아투데이 · 조선일보 · 중앙일보 · 한겨레 · 한국일보 · 강원도민일보 · 강원일보 등 전국지 · 지역지와 KBS · MBC · OBS · SBS · YTN 방송사를 포함하였다. 검색 단계에서 강릉 4,585건, 원주 547건, 춘천 3,860건, 합계 8,991건(중복 제외 후)의 기사가 수집되었다. 다만 이는 검색 결과의 규모일 뿐이며, 이후 2절의 관련성 정제를 거쳐 확정된 분석 대상은 6,738건이다. 본 연구가 제시하는 모든 결과는 원본이 아니라 이 정제 모집단에서 산출된 값이다.

빅카인즈에서 제공하는 뉴스 메타데이터와 본문 텍스트는 엑셀 파일로 내려받았다. 빅카인즈가 본문 일부만 제공할 때는 각 언론사 웹사이트에서 원문을 별도로 크롤링하여 보완하였다. 다만 강원도민일보 등 일부 언론사는 웹사이트의 HTML 구조상 본문이 JavaScript로 로딩되는 방식이어서 정적 크롤링으로는 본문 전체를 추출하기 어려웠다. 이러한 기사는 빅카인즈가 제공하는 제목과 요약문, 부분 본문만으로

분석에 포함하였다.

분석 시기는 강원권 철도 인프라의 주요 변화 시점을 기준으로 네 구간으로 구분하였다. 담론 형성 시기(2010~2011)는 경춘선 복선전철 개통 전후로 춘천에만 해당한다. 1기 기반 형성기(2012~2017)는 경강선 KTX 건설 · 개통과 2018 평창동계올림픽 준비기에 해당한다. 다만 원주는 2016년 이전 기사가 거의 없어 실질적 분석 기간이 2016~2017년으로 단축된다. 2기 전환 · 확장기(2018~2021)는 평창동계올림픽 개최와

<표 2> 분석 시점 구분

시기	기간	지역	주요 사건 · 정책
담론 형성 시기	2010 ~ 2011	춘천	경춘선 복선전철 개통(2010.12)
1기 기반 형성기	2012 ~ 2017	강릉 · 원주 · 춘천	경강선 착공(2012), 경강선KTX 개통(2017.12), 평창올림픽 준비
2기 전환 확장기	2018 ~ 2021	강릉 · 원주 · 춘천	평창동계올림픽(2018.2), 중앙선 복선전철 개통(2021.1), COVID-19
3기 재구조화 기	2022 ~ 2026	강릉 · 원주 · 춘천	강원특별자치도 출범(2023), 레고랜드 개장(2022), 동해선 연장, 서원주-여주 복선전철 논의

중앙선 원주~제천 복선전철 개통이 집중된 시기다. 3기 재구조화기(2022~2026)는 강원특별자치도 출범과 동해선 연장, 서원주~여주 복선전철 논의 등이 진행되는 현재진행형의 시기다(<표 2> 참조).

2. 자료 정제

명사 추출에는 KoNLPy의 Komoran과 Okt 형태소 분석기를 병용하였다(Park and Cho, 2014). 2~6음절 명사를 추출 대상으로 삼되 1음절 명사는 맥락 없이 의미를 판별하기 어려운 경우가 많아 제외하였다. 뉴스 사이트 인터페이스에서 반복 추출되는 노이즈 키워드(예: '보기', '검색', '구독')는 해석 대상에서 배제하되³⁾ TF-IDF 수치 자체는 원자료 그대로 유지하여 재현 가능성을 확보하였다.

정제 절차의 필요성은 원자료를 점검하는 과정에서 분명해졌다. '역세권'이라는 단어가 기사 제목 또는 본문에 실제로 포함된 기사는 세 도시 모두 전체의 6~10% 수준에 그쳤다. 이는 빅카인즈 검색이 형태소 확장 및 관련어 매칭을 통해 폭넓게 이루어지기 때문이다. 이에 두 단계의 정제 기준을 마련하였다. 엄격 기준은 '역세권' 단어가 직접 포함된 기사만을 선별한다. 완화 기준은 역세권·철도·역·KTX·복선전철 등 관련어를 포함하는 기사에서 마약·선거·유세·

마라톤 등 명백히 무관한 주제를 규칙 기반으로 제외한다.⁴⁾ 두 기준을 병행하여 원본(8,991건), 엄격(645건), 완화(3,905건)의 세 데이터 세트에 대한 민감도를 검토하였다. 이 과정에서 최종적으로 확정된 분석 확정(동일모집단) 자료는 6,738건이다(<표 3> 참조). 이하 TF-IDF, 의미연결망, 감성분석은 이 동일모집단을 공통 자료로 사용한다.

3. TF-IDF 분석 설계

TF-IDF는 문서 내 단어 출현 빈도(TF)와 전체 코퍼스에서 나타나는 희소성(IDF)을 곱한 가중치 지표다(김세영, 2023). 전체 코퍼스에 보편적으로 등장하는 단어의 가중치는 낮추고 소수의 문서에만 등장하는 단어의 가중치는 높이는 방식으로 산출된다. 본 연구에서는 '지역×연도'를 하나의 문서 단위로 설정하여 예컨대 '강릉×2017'에 해당하는 모든 기사의 명사를 하나의 문서로 취급하고 전체 지역×연도 조합(43개 문서)을 코퍼스로 삼아 가중치를 산출하였다. 이 방식은 개별 기사 수준의 분석에 비해 키워드의 연도별 변동을 안정적으로 포착할 수 있다는 장점이 있다. 다만 동일 연도 안에서 기사 간 이질성이 평균화된다는 한계도 함께 지닌다. 아울러 지역×연도 문서에 포함된 기사 수가 적을 경우 특정 기사나 개별 사업 보도 하나가 해당 셀의 상위 키워드 순위를 좌우할 수 있다. 이에 본 연구는 기사 수가 미달하는 셀에 대해서는 단년도 상위 키워드를 독립적 해석의 근거로 삼지 않고, 시기 단위로 병합한 결과만을 해석 대상으로 삼았다. 원주의 2016년 이전 연도가 이에 해당한다.

가중치 산출에는 Python scikit-learn의 TfidfVectorizer를 사용하였다. L2 정규화를 적용해 문서 길이의 영향을 보정하였다. 시기 구분별 분석에서는 동일 시기에 속하는 연도를 다시 하나의 문서로 묶어 시기별 TF-IDF를 재산출하였다(<표 4> 참조).

<표 3> 역세권 관련성 필터링

해당 지역	원본	엄격기준 (역세권직접)	완화기준 (관련어+무관계외)	분석확정 (동일모집단)
강릉	4,585	356	2,192	3,701
원주	547	55	308	409
춘천	3,860	234	1,405	2,628
합계	8,991	645	3,905	6,738

주: 엄격기준과 완화기준은 관련성 판정의 민감도를 검토하기 위한 두 개의 대안적 기준이며, 분석확정 모집단은 두 기준의 검토 결과를 반영하여 별도로 설정한 독립 필터링 단계의 산출물이다. 따라서 네 수치는 순차적 축소 관계가 아니다. 본 연구의 정제는 해석 단계에서 특정 키워드를 배제하는 방식이 아니라 자료 단계에서 기사를 제거하는 방식이다. TF-IDF·의미연결망·감성분석의 산출값은 모두 원본 8,991건이 아닌 분석확정 6,738건에서 계산되었다.

3) 적용한 불용어는 총 113개로 '앵커', '리포트', '포토', '재생', '제공', '구독' 등 방송 리포트 및 뉴스 인터페이스에서 반복 추출되는 형식어가 대부분. 이를 적용한 결과 전체 토큰 252,749개 중 1.77%(4,478개)가 제거되었고, 고유 명사수는 13,155개에서 13,121개로 감소하였다. 정제 전후 지역별 TF-IDF 상위 10개 키워드는 세 지역 모두 구성과 순위가 동일하여, 불용어 제거가 상위 담론 키워드의 구조를 변화시키지 않았음을 확인하였다.

4) 규칙 기반 제외는 제목에 마약·선거 유세·마라톤·부고·스포츠 경기 결과 등 역세권 개발과 무관기사를 대상으로 하였다. 다만 본문에 해당 주제가 포함되었더라도 역세권 관련어가 함께 등장할 때는 제외하지 않았다.

<표 4> TF-IDF 분석 설계

항목	설정
문서 단위(d)	지역× 연도(예: 강릉×2017)
코퍼스 규모(N)	43개 문서(강릉15 + 원주11 + 춘천17)
토큰화	KoNLPy Komoran · Okt 병용, 2~6음절 명사
불용어 처리	뉴스 인터페이스 노이즈113개+ 수작업 제거
가중치 산출 도구	scikit-learn TfidfVectorizer (L2 정규화 적용)
시기별 분석	동일 시기 연도를 병합하여 재산출

4. 의미연결망, CONCOR, 프레임 분석 설계

TF-IDF가 키워드의 상대적 중요도를 보여준다면 의미연결망 분석은 키워드 간의 관계 구조를 드러낸다. 지역·시기별 핵심어를 대상으로 동일 기사 내 동시출현 빈도를 계산하여 키워드 간 행렬을 구축하고 연결중심성이 높은 단어를 해당 담론의 주요 결절점으로 해석하였다. CONCOR 분석은 연결패턴의 구조적 유사성을 기준으로 키워드 군집을 도출하는 데 활용하였다(이천재, 2024). 군집명은 군집 내부의 중심 키워드와 대표기사를 함께 검토하여 결정함으로써 자동 산출된 군집을 연구자가 사후적으로 임의 명명하는 문제를 최소화하고자 하였다. 마지막으로 Entman의 프레임 구성요소(문제 정의, 원인 진단, 도덕적 평가, 처방 제안)에 따라 각 군집의 상위 키워드와 대표기사 원문을 검토하여 프레임을 분류하였다(박재명, 2025). 복수 요소가 나타나는 경우 가장 지배적으로 확인되는 요소를 중심으로 해석하였다.

5. 감성분석 설계 및 신뢰성 검증

각 기사의 본문 텍스트를 형태소 분석한 뒤 KNU 한국어 감성사전(14,843개 감성 어휘)과 매칭하여 기사별 감성점수를 산출하였다. 감성점수는 매칭된 감성 어휘의 정도값(-2~+2)을 합산하여 산출한다. 0보다 크면 긍정, 작으면 부정, 매칭되지 않거나 합이 0이면 중립으로 분류하였다. 다만 기사 길이가 길수록 매칭되는 감성 어휘 수가 늘어나 절대값이 커지는 경향이 있으므로, 감성점수를 감성어 매칭 수로 나눈 길이보정 감성점수(length-normalized sentiment score)를 함께 산출하여 병기하였다. 이는 문서 길이에 따라

부정 어휘가 누적되는 구조적 편향을 통제하기 위한 정규화(normalization) 절차이며, 지역·시기 비교에는 원점수와 정규화 점수를 항상 병기한다. 지역×시기 집계에서는 감성점수의 절대값보다 지역 간·시기 간 상대적 차이를 중심으로 해석하였다(<표 5> 참조). 사전 기반 감성분석의 한계를 보완하기 위해 정규화점수, 인간 라벨링, KLUE-BERT 비교라는 세 단계의 검증을 설계하였다. 첫째, 문서 길이 편향을 통제하기 위해 감성어 매칭 수로 정규화한 길이보정 감성점수를 병기하였다. 둘째, 지역·시기를 기준으로 층화한 표본을 인간 라벨링과 대조하여 사전 판정의 정확도와 일치도(Cohen's κ)를 검토하였다. 셋째, 동일 표본에 KLUE-BERT 문장 단위 감성 분류를 적용하여 KNU 정규화 점수와 순위상관(Spearman's ρ)을 검토하였다.

<표 5> KNU 감성분석 설계

항목	설정
감성사전	KNU 한국어 감성사전(14,843개 어휘)
정도값 범위	-2(매우 부정) ~ +2(매우 긍정)
매칭 단위	1-gram, 2-gram, 관용구, 문형, 축약어
기사별 점수 산출	매칭된 감성 어휘 정도값의 합산
감성유형 분류 기준	$S > 0$: 긍정, $S < 0$: 부정, $S = 0$: 중립
집계 단위	지역× 시기(10개 조합)
분석 도구	Python (pandas, KoNLPy)

마지막으로 TF-IDF와 감성분석 결과는 지역×시기 별로 대조하여 결합 해석한다. 각 지역·시기 조합의 TF-IDF 상위 키워드와 해당 시기의 평균 감성점수·긍정/부정 비율을 하나의 매트릭스로 통합함으로써 담론의 내용과 정서적 방향이 동일한 시공간 단위에서 어떻게 연동되는지를 살펴본다. 다만 이러한 결합 해석은 상관관계의 서술에 해당한다. TF-IDF와 감성점수가 동일한 기사 집합에서 독립적으로 산출된 지표라는 점에서 특정 키워드의 부상이 감성 변화의 원인이라는 인과적 주장으로 확대 해석될 수 없음을 미리 밝혀둔다.

IV. 분석 결과

1. 도시별 담론 구조 분석

강릉·원주·춘천 세 도시는 동일한 검색 키워드로 수집되었음에도 담론이 조직되는 축과 그 축이 전환되는 방식에서 뚜렷한 차이를 보였다. 강릉은 국가적 이벤트에서 광역 교통망으로 담론의 중심이 이동하는 단계적 전환을, 원주는 부동산과 철도라는 두 축이 병존하는 구조를, 춘천은 교통 이벤트 이후 문화·관광으로 분산되었다가 개발 의제로 회귀하는 순환적 구조를 각각 나타냈다. 이하에서는 이러한 차이가 시기별로 어떻게 형성되었는지를 살펴본다.

1) 강릉: 국제행사에서 광역 철도망으로의 담론 이행

강릉의 담론은 시기에 따라 축이 뚜렷하게 전환된다는 점에서 세 도시 가운데 가장 선명한 전환 양상을 보인다.

1기 기반 형성기(2012~2017)는 KTX 경강선 개통과 평창동계올림픽이라는 두 개의 대형 이벤트가 담론의 골격을 이룬 시기다. ‘강릉역’이 이 시기 누적 TF-IDF에서 압도적 비중을 차지했고 ‘철도’, ‘구간’, ‘올림픽’, ‘복선전철’ 등이 뒤를 이었다. 특히 ‘구간’, ‘지하’와 같은 건설 관련 기술적 용어의 가중치가 높게 나타난 점은 이 시기 보도의 초점이 노선 결정과 공사 진행, 개통 일정 등 인프라 구축 자체에 맞추어져 있었음을 시사한다.

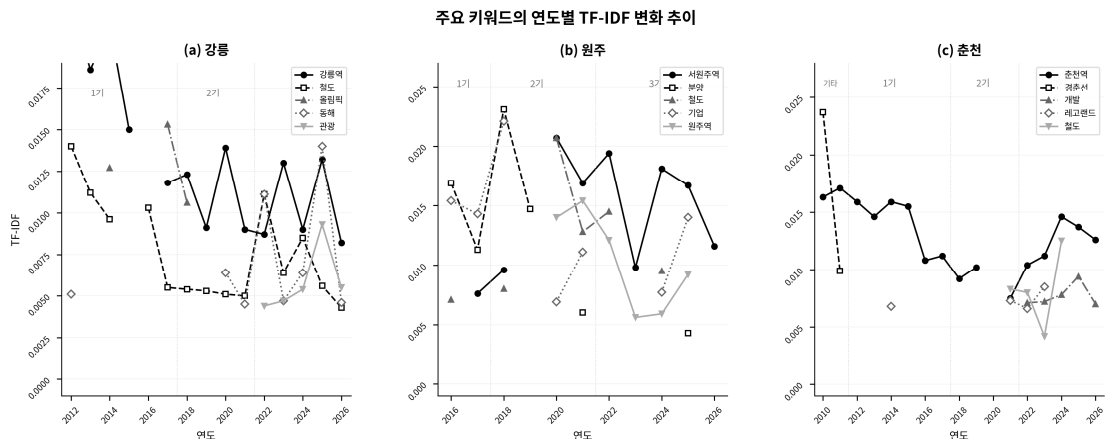
2기 전환·확장기(2018~2021)에 이르러 담론의 성

격은 크게 달라진다. 팬데믹 관련 키워드(‘코로나’, ‘환자’, ‘접종’)가 상위권에 새롭게 진입하는 한편 ‘코레일’, ‘승객’ 등 철도 운영 관련 키워드도 함께 부각되었다. 이는 KTX 개통 이후 일상적 이용 과정에서 운행 지연과 안전 관리에 대한 보도가 늘어난 결과로 해석된다. 1기에 두드러졌던 ‘올림픽’이 2기에도 일정한 비중을 유지하고 있으나 운영·안전 담론에 비해서는 상대적으로 위축되었다. 요컨대 2기의 강릉 담론은 올림픽·개통이라는 기대 중심의 서사에서 운영 품질의 문제로 중심이 이동하는 경향을 보였다.

3기 재구조화기(2022~2026)는 담론의 공간적 지평이 확장되는 시기다. ‘동해’, ‘관광’, ‘관광객’이 새롭게 상위권에 진입했으며 동해선을 통한 부산 방면 철도 연결(‘부산’, ‘이음’, ‘부전역’)이 담론의 새로운 축으로 부상하였다. ‘자율주행’이 3기 상위권에 등장하기도 하였으나, 이는 특정 연도의 보도에 국한된 단발적 관찰이므로 담론 구조의 전환을 뒷받침하는 근거로는 삼지 않는다. ‘운행’, ‘이용’ 등의 키워드가 3기 전반에 걸쳐 고르게 분포하는 양상은 담론의 초점이 건설에서 운영·이용으로, 단일 노선에서 광역 교통망으로 확대되었음을 보여준다.

이러한 시기별 전환은 의미연결망과 프레임 분석에서도 일관되게 확인된다. 강릉역은 의미 연결망에서 최상위 허브로 나타났으며 철도·이용·구간·운행이 그 주변에 결합되어 있었다. 프레임은 1기의 ‘시설 구축과 국가행사 지원’에서 2·3기의 ‘운영 품질과 광역 연결’로 이행하였다. 이는 곧 ‘건설에서 운영으로’의 프레임 이동으로 요약할 수 있다.

<그림 2> 주요 키워드 TF-IDF 추이



2) 원주: 부동산 개발에서 철도 결절점으로

원주의 담론 구조는 강릉과 뚜렷이 구별된다. 원주 데이터는 2016년부터 수집되어 1기에 해당하는 연도가 2개(2016~2017)에 불과하므로 이하 해석에는 이러한 자료상의 제약을 함께 고려할 필요가 있다.

1기(2016~2017)의 담론은 강릉과 달리 처음부터 민간 부동산 시장에 가까운 성격을 띠었다. ‘원주’, ‘가구’, ‘기업’, ‘분양’, ‘단지’, ‘아파트’ 등 분양 관련 용어가 상위권을 지배했으며 2017년에는 ‘반도’, ‘아이비’, ‘파크’와 같은 특정 아파트 브랜드명까지 상위에 진입하였다. 강릉의 1기가 올림픽·철도라는 공공 인프라 담론이었던 것과 대조적으로, 원주의 초기 담론은 처음부터 투자 대상으로서의 성격을 내포하고 있었던 셈이다.

2기(2018~2021)는 원주 담론의 구조적 전환점에 해당한다. ‘서원주역’이 처음으로 높은 TF-IDF를 기록하면서 특정 역사(驛舍)를 중심으로 한 공간적 담론이 형성되기 시작했으며 ‘철도’, ‘중앙’, ‘제천’, ‘복선’, ‘개통’ 등 중앙선 복선전철화 사업과 직접 관련된 키워드가 집중적으로 등장했다. 그러나 이 시기에도 ‘안애’, ‘카운티’ 등 부동산 관련 키워드가 여전히 유지되었다는 점은 원주의 담론에서 공공 철도 인프라와 민간 부동산 개발이라는 두 축이 대체되지 않고 중첩되어 있었음을 보여준다.

3기(2022~2026)에서는 서원주역의 TF-IDF가 더욱 상승하여 원주 역세권 담론의 핵심 결절점으로 자

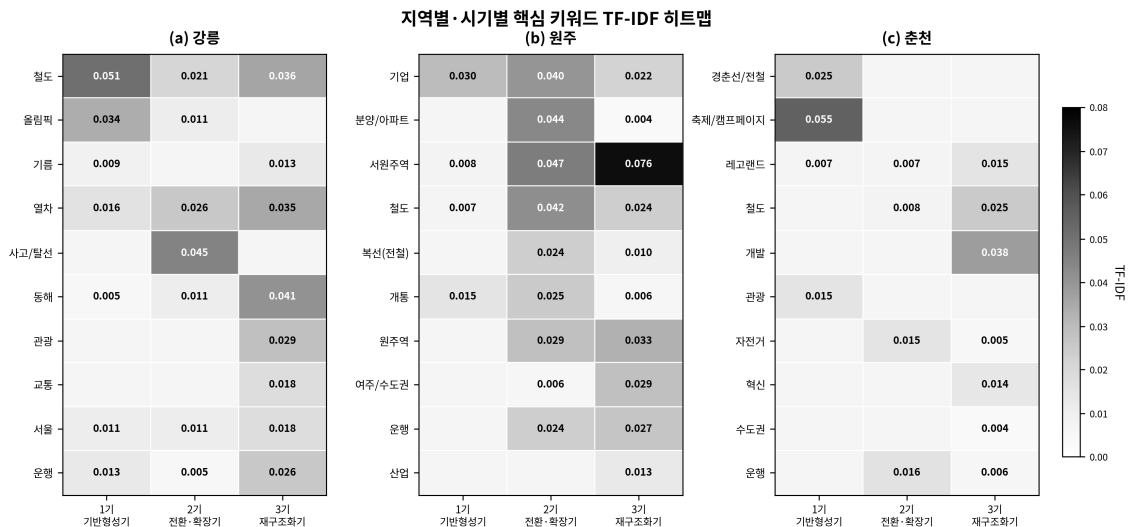
리 잡는다. ‘여주’, ‘복선전철’, ‘수도권’이 상위에 등장한 것은 서원주~여주 간 복선전철을 통한 수도권 연결 논의가 새로운 담론 축으로 형성되고 있음을 시사한다. 의미 연결망에서도 원주시에 이어 서원주역이 높은 중심성을 보였으며 개통·중앙·수도권이 그 주변에 결합되어 있었다. Entman의 프레임으로 보면 원주 담론의 문제는 수도권 접근성과 운영 편의 부족으로 규정되고, 원인은 철도망의 미완결과 배차·환승 체계로 진단되며, 처방은 여주~서원주 연결 및 광역교통 개선으로 제시된다. 강릉의 ‘건설에서 운영으로’라는 프레임 이동과 달리, 원주의 담론은 처음부터 결절점을 중심으로 한 ‘연결 확장’의 성격을 일관되게 유지했다고 볼 수 있다.

3) 춘천: 경춘선 개통에서 도시개발 담론으로

춘천은 세 지역 가운데 가장 이른 시점(2010년)부터 자료가 수집되어 철도 개통 전후의 담론 변화를 가장 긴 시계열에서 관찰할 수 있는 사례다.

담론 형성 시기(2010~2011)는 경춘선 개통이라는 단일 사건에 담론이 집중된 시기다. ‘경춘선’, ‘복선전철’, ‘개통’, ‘전철’ 등이 상위를 차지하는 가운데 ‘관광객’, ‘수도권’ 등 개통 이후의 이용 패턴과 관광 효과에 대한 관심 역시 이미 반영되어 있었다. 1기(2012~2017)에 접어들면서 담론의 성격은 빠르게 바뀐다. 철도 관련 키워드의 비중이 줄어드는 대신 ‘축제’, ‘캠핑’, ‘행사’ 등 문화·관광 관련 키워드가 부상하

<그림 3> 지역별·시기별 TF-IDF 상위 키워드 히트맵



였으며 2014년에는 이후 춘천 담론의 핵심어로 자리 잡게 될 ‘레고랜드’가 처음 상위권에 진입하였다. ‘경춘선’과 ‘복선전철’이 개통 이후 불과 2년 만에 상위권에서 밀려난 것은 철도 인프라 자체에 대한 담론적 관심이 급격히 약화되고 그 인프라를 활용한 관광·문화 콘텐츠로 초점이 이동했음을 뜻한다.

2기(2018~2021)는 춘천의 역세권 담론이 상대적으로 약화된 시기다. 기사 수가 1기의 1,011건에서 494건으로 절반 이하로 줄었을 뿐 아니라 내용상으로도 뚜렷한 상위 키워드가 다른 지역에 비해 드러나지 않았다. 역세권 고유의 공간적·교통적 담론보다는 지역의 다양한 이슈가 혼재된 양상에 가까웠다.

3기(2022~2026)에 이르러 춘천 담론은 새로운 국면을 맞는다. ‘개발’이 처음으로 높은 누적 TF-IDF를 기록하며 상위권에 진입했고 ‘철도’ 역시 2기에 비해 크게 상승했다. ‘레고랜드’는 2022년 개장과 맞물려 다시 부각되었으며 ‘강원도’, ‘특별자치도’는 강원특별자치도 출범(2023)을 반영한다. 다만 이 시기에는 지역 내 마약 범죄 보도가 집중되어 역세권 담론과 무관한 기사가 잔존하였다. 이에 제목 기준 규칙에 더해 본문 기준으로 해당 기사를 추가 제거한 뒤 춘천 3기 TF-IDF를 재산출하였으며, 상위 5개 키워드의 구성과 순위는 변동 없이 유지되었다.

의미연결망에서는 춘천역·경춘선·개발·시설·관광이 서로 결합되어 나타났으며 프레임은 ‘개통을 통한 수도권 접근성’에서 ‘관광 활용’으로, 다시 ‘역세권 도시개발과 광역철도’로 이어지는 순환적 구조를 보였다. 이는 강릉의 단계적 전환이나 원주의 병존적 지속과 구분되는, 춘천 담론 고유의 특징이라 할 수 있다.

2. 지역 간 비교 및 유형화

세 도시의 담론구조를 개별적으로 살펴본 결과는 이를 나란히 놓고 비교할 때 비로소 온전한 함의를 갖는다. 이 절에서는 TF-IDF 키워드 구성의 차이, 의미연결망과 프레임의 차이를 각각 검토한 뒤 이를 종합하여 세 도시의 담론을 유형화한다.

1) TF-IDF 키워드 비교

세 지역의 TF-IDF 결과를 종합하면 동일한 검색 키워드로 수집된 기사임에도 담론의 내적 구성은 도시

마다 상이한 방향으로 분화되었음이 드러난다(<표 6> 참조).

강릉은 시기별 담론 전환이 가장 선명한 사례다. 1기에는 ‘올림픽’과 ‘동계올림픽’이 결합하여 높은 비중을 차지했던 반면, 3기에는 이러한 국가 행사 관련 키워드가 완전히 자취를 감추고 ‘동해’, ‘관광’, ‘부산’이 그 자리를 대신했다. 담론의 공간적 범위가 강릉이라는 단일 도시에서 동해안이라는 광역 축으로 확장된 셈이다.

원주는 부동산 개발 담론과 철도 인프라 담론이 시종일관 병존한다는 점에서 강릉·춘천과 뚜렷이 구별된다. 다른 두 지역에서는 아파트 브랜드명이 TF-IDF 상위권에 오르는 경우가 거의 없었던 반면, 원주에서는 1기부터 3기까지 ‘분양’, ‘아파트’, ‘가구’와 특정 브랜드명이 꾸준히 높은 가중치를 유지했다. 이는 원주의 역세권 담론이 교통 결절점으로서의 공적 기능뿐 아니라 부동산 투자 대상으로서의 성격을 함께 내포하고 있음을 뜻한다.

춘천은 경춘선 개통이라는 단일 교통 이벤트 이후 담론이 문화·관광·축제 방향으로 빠르게 분산되었다는 점에서 특징적이다. 강릉이 올림픽 이후에도 철도·교통 관련 키워드를 지속적으로 상위권에 유지한 것과 달리, 춘천에서는 경춘선 관련 키워드가 개통 2년 만에 상위권에서 이탈했다. 3기에 ‘개발’과 ‘철도’가 다시 상승한 것은 춘천~속초 철도와 역세권 도시개발 사업이 새로운 담론 동력으로 작용하고 있음을 시사한다.

<표 6> 지역별·시기별 TF-IDF 상위 5개 키워드 (노이즈 제외)

시기	강릉	원주	춘천
형성(2010~2011)	-	-	경춘선, 복선전철, 개통, 전철, 전동차
1기(2012~2017)	강릉역, 철도, 올림픽, 구간, 복선전철	원주, 가구, 기업, 분양, 단지	춘천역, 춘천시, 문화, 축제, 캠페이지
2기(2018~2021)	코로나, 코레일, 승객, 올림픽, 현송월	서원주역, 철도, 중앙, 제천, 복선	(역세권 고유 키워드 약화, 기사 수 급감)
3기(2022~2026)	동해, 관광, 부산, 이음, 자율주행	서원주역, 여주, 수도권, 복선전철, 원주역	개발, 철도, 레고랜드, 강원도, 특별자치도

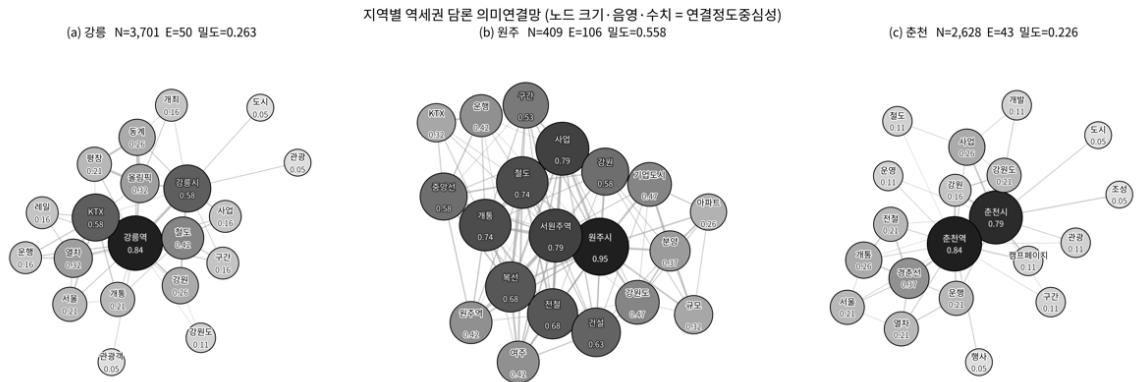
2) 의미연결망, 프레임 비교

TF-IDF가 보여주는 키워드의 상대적 비중을 넘어 담론이 어떤 관계 구조 속에서 조직되는지를 파악하기 위해 TF-IDF · 감성분석과 동일한 분석확정 모집단(6,738건)에서 명사 간 동시출현 의미연결망을 구성하고 연결 정도 중심성을 산출하였다(<그림 4> 참조). 노드는 지역별 문서빈도 상위 20개 명사로 하되 해당 지역명은 전 기사에 편재하여 판별력이 없으므로 제외하였으며 옛지는 세 지역에 공통으로 적용한 상대기준(해당 지역 기사의 4.0% 이상에서 공출현)으로 선별하였다. 강릉과 춘천에서는 역 자체(강릉역 .84, 춘천역 .84)가 최상위 허브로 나타나 담론이 물리적 결절점을 중심으로 조직되는 양상을 보였다. 반면 원주에서는 행정주체인 '원주시'(.95)가 '서원주역'(.79)보다 앞서 담론이 특정 역사(驛舍)보다는 사업 추진 주체를 중심으로 조직되는 다소 상이한 양상이 관찰되었다. 이는 원주의 역세권 담론이 서원주역 · 복선전철 건설사업이라는 진행형 사업을 축으로 형성되어 온 것과 무관하지 않은 것으로 보인다. 네트워크 밀도에서도 원주는 .558로 강릉(.263) · 춘천(.226)의 두 배를 넘어 세 지역에 동일한 상대기준을 적용했음에도 원주 담론이 단일 축으로 응집되어 있음을 보여준다.

각 도시 담론이 이 허브를 중심으로 어떤 의제와 결합하는지도 도시마다 달랐다. 강릉은 철도 시설과 운영 관련 키워드가 밀집하여 '건설에서 운영으로'의 프레임을 형성했고, 원주는 서원주역 · 개통 · 수도권 이 결합하여 '결절점 연결' 프레임을, 춘천은 경춘선 · 관광 · 개발이 순환적으로 결합하여 '교통에서 도시개발로'의 프레임을 각각 형성하였다(<표 7> 참조). 이러한 프레임의 차이는 앞서 TF-IDF 비교에서 확인한 담론 축의 차이와 대응한다는 점에서 두 분석 결과가 서로를 교차 검증하는 관계에 있다고 볼 수 있다.

세 도시 모두 지방정부가 최다 행위자로 나타났으나 그 다음 순위에서 차이가 확인된다. 강릉과 춘천은 지역 주민이 중앙정부에 앞서거나 근접하는 반면, 원주는 중앙정부(33)가 지역주민(5)을 크게 앞선다. 이는 원주 역세권 담론이 중앙선 복선전철이라는 국가 재정 사업을 축으로 형성되어 주민 참여의 계기가 상대적으로 적었음을 시사한다. 갈등구조에서는 세 도시 모두 중앙-지방 축이 지배적이며 춘천에서 교통편의-재정부담 갈등(17)이 두드러지는 점은 경춘선 개통 이후 운영 · 유흥지 활용을 둘러싼 비용 논의가 지속되어 온 사정과 대응한다(<표 8> 참조).

<그림 4> 지역별 역세권 담론 의미 연결망



<표 7> 지역별 연결중심성 상위 키워드

지역	연결중심성 상위(값)	프레임 해석(Entman)
강릉	강릉역(.84) · 강릉시(.58) · KTX(.58) · 철도(.42) · 올림픽(.32)	'이벤트 계기 → 광역교통 연계' 프레임: 동계올림픽을 계기로 개통된 고속철도가 역 중심 접근성 의제로 이어짐
원주	원주시(.95) · 서원주역(.79) · 사업(.79) · 철도(.74) · 개통(.74)	'사업 추진 → 결절점 형성' 프레임: 서원주역 · 복선전철 건설사업이 단일 축으로 수렴하는 응집형 담론
춘천	춘천역(.84) · 춘천시(.79) · 경춘선(.37) · 개통(.26) · 사업(.26)	'교통 → 관광 · 도시개발 분산' 프레임: 경춘선 개통 이후 의제가 관광 · 유흥지 개발로 분화하는 다축 구조

<표 8> 주요 행위자·갈등구조

지역	주요 행위자(빈도)	주요 갈등구조(빈도)
강릉	지방정부(69), 지역주민(27), 중앙정부(16), 공기업·교통기관(7), 정치권(7)	중앙-지방(30), 사업자-주민(6), 교통편의-재정부담(5), 개발-보전(4)
원주	지방정부(66), 중앙정부(33), 정치권(11), 공기업·교통기관(7), 지역주민(5)	중앙-지방(25), 교통편의-재정부담(2), 사업자-주민(1)
춘천	지방정부(112), 중앙정부(32), 지역주민(27), 정치권(10), 공기업·교통기관(5)	중앙-지방(50), 교통편의-재정부담(17), 개발-보전(2)

<표 9> 강원 역세권의 담론 유형

구분	강릉	원주	춘천
담론 유형	관광·이벤트 기반 성장거점형	산업·계획도시 기반 성장거점형	수도권 연계·광역접근성 강화형
담론촉발 계기	올림픽, KTX개통	계획도시 개발	경춘선 개통
의미연결망 최상위 허브	강릉역(.84)	원주시(.95)	춘천역(.84)
동일 모집단(기사수)	3,701건	409건	2,628건
길이보정 정규화 감성점수	0.26	0.41	0.23
전체 평균 감성점수	0.44	0.59	0.36
전체 긍정비율	35.3%	40.1%	31.3%
TF-IDF 핵심 키워드	강릉역, 철도, 동해, 열차, 관광	원주, 서원주역, 기업, 분양, 철도	춘천, 춘천역, 춘천시, 문화, 개발
담론 전환 특성	올림픽→운영·안전→ 광역교통망(단계적 전환)	부동산+철도 담론 병존 (긍정 우위 유지)	경춘선→문화관광→담론공백→ 개발재부상(순환적)
감성 변화 방향	완만한 굴곡(2기저점·3기회복)	전 시기 긍정 우위(긍정 유지)	완만·안정(중립우세)

3) 담론의 유형화

TF-IDF 키워드 구조와 의미연결망·프레임 분석을 종합하면, 세 도시의 역세권 담론은 감성점수의 크기가 아니라 담론이 조직되는 내용적 축과 그 전환 양상을 기준으로 서로 다른 유형으로 구분된다(<표 9> 참조).

강릉은 국가적 이벤트로 시작된 담론이 운영 단계를 거쳐 광역 교통망으로 확장되었다는 점에서 ‘관광·이벤트 기반 성장거점형’으로, 원주는 부동산 개발과 철도 결절점 담론이 대체되지 않고 지속적으로 병존했다는 점에서 ‘산업·계획도시 기반 성장거점형’으로, 춘천은 경춘선 개통 이후 문화·관광으로 분산되었다가 도시개발 의제로 순환적으로 재부상했다는 점에서 ‘수도권 연계·광역접근성 강화형’으로 각각 유형화된다. 이 세 유형은 단순히 담론의 주제가 다르다는 사실을 넘어, 철도 인프라가 지역사회에서 의미화되는 경로 자체가 도시의 산업구조와 개발 이력에 따라 상이함을 보여준다.

3. TF-IDF, 감성분석 결합 분석

지금까지의 분석이 담론의 내용적 구조를 분석하는데 초점을 두었다면, 이 절에서는 그 담론이 어떤 정서적 방향성을 동반하는지를 살펴본다. 이를 위해 먼저 감성분석의 방법론적 신뢰성을 검증한 결과를 제시하고 이어 지역별·시기별 감성 특성을 TF-IDF 결과와 대조하여 논의한다.

1) 감성분석 신뢰성 검증 결과

감성분석의 신뢰성을 검증하기 위해 정규화 점수, 인간 라벨링, KLUE-BERT 비교의 세 단계를 수행한 결과는 <표 10>과 같다. 검증 표본은 지역·시기를 기준으로 총화한 301건이며 부동산·도시 분야 전문가 2인이 KNU 판정과 감성점수를 제시받지 않은 블라인드 조건에서 독립 코딩하였다.⁵⁾ 검증 표본의 감성점수 절대값 분포는 전체 모집단과 통계적으로 구분되지 않

5) 코더에게 제공된 시트는 표본ID·지역·시기·연도·제목 다섯 열만으로 구성하였으며, KNU 감성유형·감성점수·감성이 매칭 수는 모두 제거하였다. 두 코더의 기사 제시 순서는 서로 다르게 무작위 배열하였고, 불일치 사례의 협의 조정 단계에서도 KNU 판정을 제시하지 않았다.

아(Mann-Whitney U 검정 $p=.241$, Kolmogorov-Smirnov 검정 $p=.892$), 판정이 용이한 기사로 표본이 치우치지 않았음을 확인하였다.

코더 간 일치도는 $\kappa=.740$ 으로 높은(substantial) 수준이었으며⁶⁾ 불일치 48건 중 38건(79.2%)이 긍정과 중립의 경계에서 발생하였다. 이는 개발 성과를 전하는 보도를 성과에 대한 평가로 볼지 사실 전달로 볼지에 대한 판단 차이에서 비롯된 것으로, 뉴스 텍스트 감성 판정에 내재한 구조적 모호성을 보여준다. 불일치 사례는 KNU 판정을 제시하지 않은 상태에서 두 코더의 협의를 거쳐 조정하였다.

합의 라벨을 기준으로 KNU 사전 판정의 정확도는 .864, Macro-F1은 .862, Cohen's κ 는 .780(95% CI [.715, .837])으로 높은(substantial) 수준이었다. 코더별로는 코더A .914($\kappa=.857$), 코더B .900($\kappa=.836$)이었다. 반면 KLUE-BERT는 정확도 .488, Macro-F1 .470, $\kappa=.215$ 에 그쳤고 지역·시기 셀 단위에서 KNU 정규화 점수와의 순위상관도 유의하지 않았다($\rho=.273$, $p=.446$). 다만 셀 표본 수가 15건 이상인 7개 셀로 한정하면 상관의 부호가 역전되어($\rho=.679$, $p=.094$), 소표본 셀이 이 추정에 미치는 영향이 작지 않음을 함께 밝혀 둔다.

주목할 점은 오분류의 방향이 일관되게 한쪽으로 쏠

기사를 사전이 중립으로 분류한 사례였으며 그 반대 방향의 오분류는 관찰되지 않았다. 이 41건의 90.2%는 KNU 감성사전과 매칭되는 어휘가 하나도 없어 점수 합이 0으로 산출된 기사였다. 그 결과 검증 표본에서 사전은 중립을 53.2%로 판정한 반면 인간은 40.2%로 판정하여 13.0%포인트의 차이를 보였다.

사전 기반 방법이 문장 단위 딥러닝 모델보다 인간 판정과의 일치도가 높게 나타난 것은 일반적 기대와 다른 결과이므로 후속 연구에서 원인을 규명할 필요가 있다. 이에 본 연구는 KLUE-BERT 비교 결과를 감성 점수 해석의 보강 근거로 삼지 않으며 감성분석 결과는 KNU 사전 체계 내에서의 지역·시기 간 상대 비교에 한정하여 해석한다. 다만 인간 라벨을 기준으로 지역별 긍정비율을 재산출한 경우에도 원주(77.8%)·강릉(41.0%)·춘천(40.2%)의 순위는 사전 판정 기준의 순위와 동일하게 유지되어 본 연구가 제시하는 지역 간 상대 비교의 방향성은 판정 방법의 차이에 강건한 것으로 확인된다.

2) 지역별 감성 특성 차이

TF-IDF와 동일한 정제 모집단(6,738건)에서 재산출한 결과, 세 도시의 감성 구조는 사전 판정기준으로 중립 기사가 가장 큰 비중을 차지한다는 공통점을 보였다. 중립 비율은 강릉 47.4%, 원주 48.4%, 춘천 53.5%였으며 긍정 기사 비율(강릉 35.3%, 원주 40.1%, 춘천 31.3%)이 부정 기사 비율(강릉 17.3%, 원주 11.5%, 춘천 15.3%)을 모든 도시에서 앞섰다. 이는 정제 이전 원자료에서 부정 기사 비율이 70~78%에 달했던 것과 크게 대비되는 결과다. 무관 사건·사고 보도를 제거한 동일모집단에서 부정 비율이 15~17% 수준으로 낮아진 점, 그리고 잔존 무관 기사를 추가로 제거해도 부정 비율 변화가 1%포인트 안팎에 그친 점을 함께 고려하면, 정제 이전에 관찰되던 높은 부정 비율은 역세권 담론 고유의 특성이라기보다 무관 기사의 과다 포함과 문서 길이에 따른 부정어 누적이라는 방법론적 요인이 결합된 결과였을 가능성이 크다.

지역 간 상대 순위에서는 원주가 상대적으로 긍정적인 경향을 보였다. 부트스트랩(2,000회 재표집) 95% 신뢰구간으로 검토하면 강릉 [33.7, 36.9], 원주 [35.5, 44.7], 춘천 [29.5, 33.1]로, 춘천은 강릉·원주와 구간

<표 10> 감성분석 신뢰성 검증 결과

검증영역	검증지표	비교대상	값
표본구성	검증표본수	지역·시기 총화표본	301건
KNU 분류성능	정확도		.864
	Macro-F1	합의라벨-KNU	.862
	Cohen's κ		.780
KLUE-BERT 분류성능	정확도		.488
	Macro-F1	합의라벨 -KLUE-BERT	.470
	Cohen's κ		.215
방법 간 관계	Spearman's ρ	KNU정규화	.273
	p-value	-KLUE-BERT	.446

주: 1) N = 301(지역·시기 총화표본)

2) κ = Cohen's kappa(우연일치 통계 지표)

3) 유의수준은 $\alpha = .05$ 기준

렸다는 것이다. 사전 판정이 인간 판정과 어긋난 41건 가운데 39건(95.1%)이 인간이 긍정 또는 부정으로 본

6) Cohen's κ 의 수준 해석은 Landis and Koch(1977)의 기준을 따랐다. .81 이상 매우 높음(almost perfect), .61~.80 높음(substantial), .41~.60 보통(moderate), .21~.40 낮음(fair), .20 이하 매우 낮음(slight)로 구분된다.

이 겹치지 않아 낮은 긍정비율이 표본 우연으로 보기 어려운 반면, 원주와 강릉은 구간이 부분적으로 겹쳐 두 도시 간 차이는 경계적인 수준에 머물렀다. 길이보정 정규화 감성점수에서도 원주(+0.41) > 강릉(+0.26) ≥ 춘천(+0.23)의 순위가 유지된다는 점은 이러한 지역 간 상대 차이가 문서 길이라는 교란 요인을 통제한 뒤에도 존속함을 시사한다.

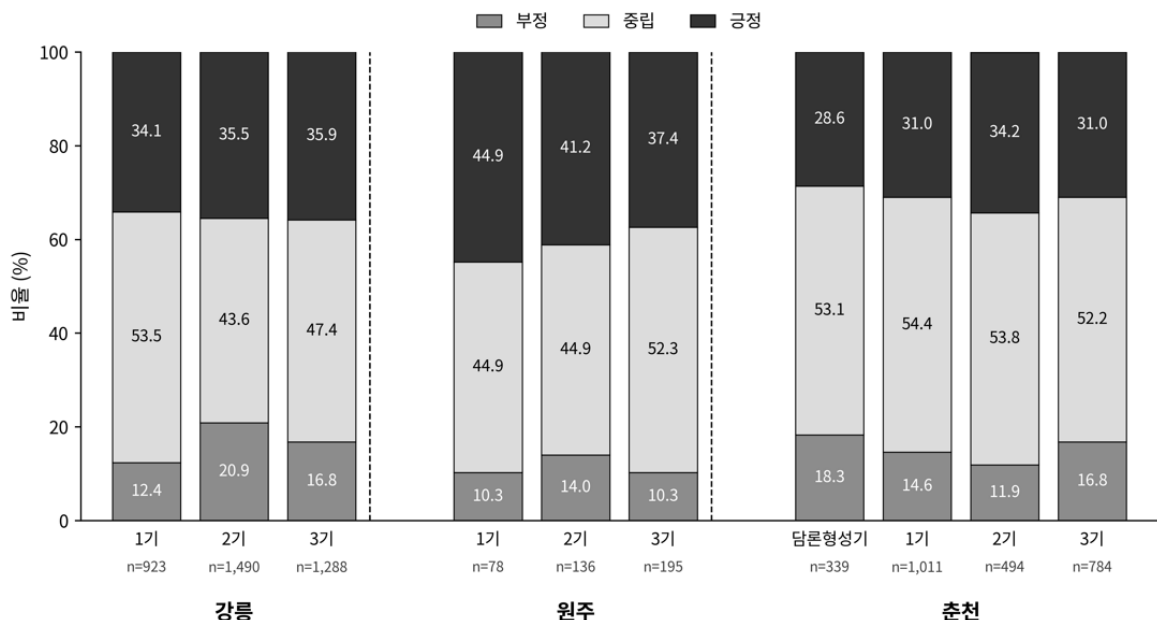
강릉의 시기별 평균 감성점수는 1기 +0.57, 2기 +0.33, 3기 +0.47로, 2기에 낮아졌다가 3기에 다소 회복되는 궤적을 보였다(길이보정 시에도 +0.32→+0.21→+0.29로 동일한 방향). 부정 기사 비율 역시 2기에 20.9%로 세 시기 중 가장 높았다. 이 2기의 저점은 같은 시기 TF-IDF 상위어에 ‘코로나·사고·열차’가 부각되고 팬데믹 및 철도 운영 초기의 안전·자연 보도가 늘어난 시점과 겹친다. 이는 앞서 1절에서 확인한 담론 축의 전환, 즉 올림픽에서 운영 품질 문제로의 이동과도 맞물려 있다. 3기에는 관광·광역철도 담론이 부각되면서 긍정 기사 비율이 35.9%로 회복되었다. 다만 이러한 대응은 상관 기술에 해당하며 특정 키워드의 부상이 감성 변화를 초래했다는 인과적 해석으로 확장될 수 없다.

원주의 시기별 평균 감성점수는 1기 +1.05, 2기 +0.28, 3기 +0.62였다(길이보정 시 +0.56→+0.37→

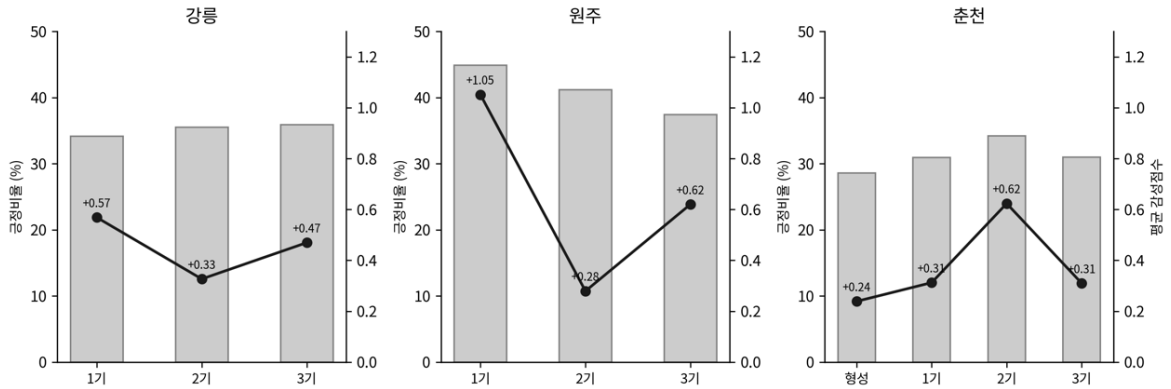
+0.37). 추정값만 보면 1기가 두드러지게 높았으나 부트스트랩 95% 신뢰구간(1기 [34.6, 56.4], 2기 [33.1, 49.3], 3기 [30.8, 44.1])이 서로 크게 겹쳐 시기 간 차이를 통계적으로 유의한 것으로 단정하기는 어렵다. 원주는 전체 409건, 1기는 78건에 불과해 표본 규모의 제약이 뚜렷하므로 시기별 감성 변화는 확정적 패턴이 아니라 제한적 사례로 해석하는 것이 타당하다. 다만 지역 수준에서 원주의 상대적 긍정 우위(긍정 40.1%)는 춘천과 신뢰구간이 겹치지 않아 유의하게 관찰된다. 이는 1기의 부동산 분양 담론과 2·3기의 철도 결절점 담론이 대체 없이 병존해 온 구조와 무관하지 않은 것으로 보인다.

춘천은 담론형성기부터 3기까지 가장 긴 분석 기간을 포괄하며 평균 감성점수는 담론형성기 +0.24, 1기 +0.31, 2기 +0.62, 3기 +0.31로 원주·강릉에 비해 변동 폭이 완만했다(길이보정 시에도 동일하게 완만). 네 시기 모두 중립비율이 52~54%로 세 도시 가운데 가장 높았다는 점은 절대 수준의 해석은 유보하더라도 춘천의 담론이 상대적으로 사실전달 중심의 성격을 강하게 띠어 왔음을 보여준다. 2기의 소폭 상승은 레고랜드 개장 전후의 문화·관광 보도와, 3기 부정 기사 비율의 소폭 상승(16.8%)은 개발·사업 재부상기의 갈등 보도와 각각 시기적으로 맞물리는 것으로 해석된다.

〈그림 5〉 지역별·시기별 감성 비율



<그림 6> 지역별 감성점수 내 긍정 비율 분석 결과



이상의 결합 해석은 상관관계의 서술에 해당하며 TF-IDF와 감성점수가 동일한 기사 집합에서 독립적으로 산출된 지표라는 점에서 특정 키워드의 부상이 감성 변화를 초래했다는 인과적 주장으로 일반화될 수 없다는 점을 다시 한 번 밝혀둔다.

4. 소결

도시별 담론 구조 분석부터 TF-IDF·감성분석 결합 분석까지의 결과를 종합하면 세 층위의 분석이 서로 다른 측면에서 하나의 결론으로 수렴함을 확인할 수 있다.

첫째, 담론이 조직되는 축 자체가 도시마다 다르다는 점은 이미 확인되었으나 이 차이는 감성 구조에서도 대응 관계를 이룬다. 강릉의 담론 축이 국가행사에서 운영·안전 문제를 거쳐 광역 교통망으로 전환된 궤적은 같은 시기 감성점수가 1기 고점에서 2기 저점을 거쳐 3기에 회복되는 과정과 시기적 대응한다. 이는 담론의 내용적 전환과 정서적 전환이 독립적으로 산출된 지표임에도 동일한 시간 구조를 공유한다는 점에서 두 층위의 분석이 서로를 뒷받침하는 관계에 있음을 시사한다.

둘째, 도시 간 비교에서 도출된 유형으로서 강릉의 단계적 전환, 원주의 병존적 지속, 춘천의 순환적 재부상은 감성 패턴과도 정합적이다. 원주는 부동산과 철도라는 두 축이 대체되지 않고 지속되었듯 감성 역시 표본 규모의 한계 속에서도 세 시기 내내 상대적 긍정 우위를 유지했다. 춘천은 담론 축의 순환만큼이나 감성 변동 폭도 완만하여 사실 전달 중심의 담론 성격이 정서적 층위에서도 일관되게 나타났다.

셋째, 이러한 해석에는 방법론적 제약이 전제되어야 한다. 감성분석 결과는 KNU 사전과 인간 라벨링 간의 높은 일치도에 근거해 해석되었으나 KLUE-BERT를 통한 독립적 교차검증에서는 통계적 유의성이 확인되지 않았다. 따라서 위에서 언급한 담론-감성 간의 대응 관계는 KNU 사전 체계 내에서의 상대 비교로 그 해석 범위를 한정해야 한다.

V. 결론

1. 연구 요약 및 정책적 함의

본 연구는 강원권 세 거점도시(강릉·원주·춘천)의 역세권 개발 담론이 2010~2026년에 걸쳐 어떻게 구성되고 변화해 왔는지를 입증하고자 하였다. 분석은 뉴스 빅데이터를 바탕으로 TF-IDF·의미연결망·프레임·감성분석을 활용해 이루어졌다. 주요 연구 결과를 정리하면 다음과 같다.

먼저 동일한 검색 키워드로 수집된 기사임에도 세 도시의 담론은 서로 다른 방식으로 조직되어 왔다. 강릉은 국가적 이벤트 중심의 담론에서 출발해 운영·안전 문제를 거쳐 광역 교통망 담론으로 이행하는 단계적 전환을 보였다. 원주는 부동산 개발과 철도 결절점이라는 두 축이 대체 없이 병존하는 구조를, 춘천은 단일 교통 이벤트 이후 문화·관광으로 분산되었다가 개발 의제로 순환적으로 회귀하는 구조를 각각 나타냈다. 이러한 차이는 TF-IDF 핵심 키워드뿐 아니라 의미연결망의 중심성 구조와 Entman 프레임에서도 일관되게 확인되었다. 이는 강릉을 ‘관광·이벤트 기반 성

장거점형', 원주를 '산업·계획도시 기반 성장거점형', 춘천을 '수도권 연계·광역접근성 강화형'으로 유형화하는 근거가 되었다.

셋째, 이러한 해석에는 방법론적 제약이 전제되어야 한다. 감성분석 결과는 KNU 사전과 인간 라벨링 간의 높은 일치도에 근거해 해석되었으나 KLUE-BERT를 통한 독립적 교차검증에서는 통계적 유의성이 확인되지 않았다. 그러므로 위에서 언급한 담론-감성 간의 대응 관계는 KNU 사전 체계 내의 상대 비교로 그 해석 범위를 한정해야 한다.

감성분석에서는 정제된 동일모집단(6,738건)을 기준으로 세 도시 모두 사전 판정에서 중립 기사가 가장 큰 비중을 차지하는 가운데 긍정 기사 비율이 부정 기사 비율을 앞섰다. 지역 간 상대 순위는 원주(긍정 40.1%)·강릉(35.3%)·춘천(31.3%) 순으로 나타났다. 정제 이전 원자료에서는 부정 우세(70~78%)가 관찰되었다. 이는 담론 자체의 속성이라기보다 무관 기사 포함과 문서 길이 편향에서 비롯된 방법론적 산물이었을 가능성이 있다. 다만 이러한 감성 결과의 신뢰성은 KNU 사전과 인간 라벨링 간의 높은 일치도($\kappa=.780$)에 근거한 것이다. KLUE-BERT를 통한 독립적 교차검증에서는 통계적 유의성이 확인되지 않았다($p=.273$, $p=.446$).

이상의 분석 결과는 강원 지역의 역세권 정책에 몇 가지 시사점을 준다. 다만 이는 강원권 세 도시 사례에 근거한 것이어서 타 지역에 적용할 때는 추가적인 검증이 필요함을 전제한다.

첫째, 강릉과 원주에서 공통으로 관찰된 '초기 고점-중간 시기 저점-후기 회복'이라는 감성 궤적은 개통 직후의 운영 안정화 단계에 정책적 관심이 집중될 필요가 있음을 시사한다. 강릉의 2기 저점은 팬데믹과 철도 운영 초기의 안전·지연 보도와 시기적으로 맞물렸다. 이 점을 고려하면 배차 간격 최적화, 역사 인근 생활 인프라 확충, 이용자 만족도 모니터링 같은 운영 지원이 담론의 정서적 회복과 연동될 가능성이 있다. 다만 이는 인과관계가 아닌 시기적 상관에 근거한 제언임을 밝혀둔다.

둘째, 원주 사례에서 확인되듯 부동산 분양 담론이 역세권 담론에 지속적으로 병존·침투하면 역세권은 교통 결절점으로서의 공공적 가치와 투자 대상으로서의 성격을 동시에 내포하게 된다. 아파트 브랜드명이 역세권 담론의 상위 키워드로 등장한다는 점은 이러한

유형의 역세권에 개발사업의 공공성 비율 설정이나 민간투자 우선원칙의 재검토 같은 균형 방안이 함께 논의되어야 함을 보여준다.

끝으로 춘천의 사례는 철도 개통이라는 일회적 계기가 지나간 뒤 역세권 담론의 공적 관심이 자연스럽게 줄어들 수 있음을 보여준다. 춘천의 담론 구조는 네 시기에 걸쳐 세 도시 가운데 가장 높은 중립 비율을 유지했다. 역 주변의 도시 기능을 꾸준히 갱신할 중장기 의제가 마련되지 않는다면 역세권은 반복적으로 담론의 공백기를 맞을 수 있다.

2. 연구의 한계 및 향후 과제

본 연구에는 다음과 같은 한계가 있다. 첫째, 감성분석의 방법론적 검증이 부분적으로만 완결되었다는 점이 가장 큰 한계다. 총화표본 301건을 대상으로 전문가 2인이 블라인드 독립 코딩을 수행한 결과 KNU 사전 판정은 인간 판정과 $\kappa=.780(95\% \text{ CI } [.715, .837])$ 수준으로 일치하였으나 KLUE-BERT 문장 단위 분류와의 순위상관은 유의하지 않았다($p=.273$, $p=.446$). 사전 판정이 인간 판정을 상당 부분 재현하였으므로 지역 간 상대 비교의 근거로는 활용할 수 있으나 독립적 방법론을 통한 완전한 교차검증에는 이르지 못하였다.

이 한계가 결과에 미치는 방향과 정도는 다음과 같이 특정된다. 검증 표본의 49.8%(150건)는 KNU 감성 사전과 매칭되는 여회가 하나도 없어 점수 합이 0으로 산출되었고 자동으로 중립으로 분류되었다. 사전과 인간 판정이 어긋난 41건 가운데 39건이 이러한 강제 중립에서 비롯되었으며 반대 방향의 오분류는 나타나지 않았다. 그 결과 본 연구의 중립 비율은 실제보다 과대 추정되었을 가능성이 높다. 그 크기는 검증 표본 기준 약 13%포인트로 추정된다. 본 연구가 중립 비율의 절대 수준에 대한 해석을 제시하지 않고 지역 간 상대 비교에 한정된 것은 이러한 편향을 고려했기 때문이다. 인간 라벨 기준으로 재산출한 지역별 긍정비율 순위가 사전 판정 기준과 동일하게 유지된다는 점에서 상대 비교의 방향성 자체는 유지되는 것으로 판단된다.

다음으로 원주는 표본 규모가 작아(전체 409건, 1기 78건) 시기별 비교의 신뢰도가 제한적이다. 본 연구는 이를 부트스트랩 신뢰구간으로 명시하고 제한적 사례로 해석하였으나 이러한 표본 제약은 원주를 포함한

유형화의 일반화 가능성에도 영향을 미칠 수 있다.

마지막으로 일부 언론사(강원도민일보 등)의 기사는 HTML 구조상 본문 전체를 추출하기 어려워 제목과 부분 본문만으로 분석에 포함되었다. 이러한 기사에서는 감성어 매칭 수가 과소 산출되어 중복으로 분류될 여지가 있다.⁷⁾

향후 연구에서는 검증 표본을 확대해 KNU와 KLUE-BERT 간 불일치의 원인을 입증하고 두 방법론 간 괴리가 두드러지는 사례를 분석함으로써 방법론적 신뢰성을 보강할 필요가 있다. 아울러 역세권과 무관한 기사를 사전 필터링하는 작업을 주제 분류 모델과 결합한다면 분석의 정밀도를 한층 높일 수 있다.

이러한 한계에도 불구하고 본 연구는 TF-IDF와 감성분석을 동일한 정제 모집단에서 결합함으로써 키워드 빈도만으로는 포착하기 어려운 담론의 정서적 차원을 상대 비교의 형태로 가시화하였다. 이는 역세권이 물리적 개발 대상을 넘어 지역별로 다르게 구성되는 담론적 공간임을 실증한다. 나아가 이 연구는 역세권 정책의 수립과 평가에 정량적·정성적 근거를 함께 제공하는 분석 틀로 기능할 수 있다는 점에서 의의를 지닌다.

논문접수일 : 2026년 5월 12일

논문심사일 : 2026년 7월 13일

게재확정일 : 2026년 8월 24일

7) 본문 전체 추출이 불가하여 제목과 요약문만으로 분석된 기사는 1,583건(전체의 23.5%)이며, 이 집단의 중복 분류 비율은 52.1%로 전체 평균 49.8%보다 2.3%포인트 높다. 다만 이러한 기사의 비율은 지역별로 균등하지 않아 강릉 18.4%, 원주 13.0%, 춘천 32.3%로 최대 19.3%포인트의 격차가 있다. 해당 기사를 제외하고 재산출하여도 지역 간 중복 비율의 순위는 유지되므로, 지역 간 상대 비교에 미치는 영향은 제한적인 것으로 판단된다.

참고문헌

- 곽희중, “텍스트마이닝을 활용한 지역별 도시재생 이슈 분석-빅카인즈 지역일간지를 대상으로”, 『부동산분석』 제8권 제3호, 2022, pp. 53-71
- 김세영, “TF-IDF 분석과 토픽 모델링을 활용한 AI 기반 개별화 학습 국내의 연구동향 분석”, 『정보교육학회논문지』 제27권 제4호, 2023, pp. 453-464
- 김재훈, “강원지역의 산업구조 변동과 지역 연결망: 2000년대를 중심으로”, 『사회과학연구』 제59집 제1호, 2020, pp. 43-79
- 김형민·장희순, “지방도시내 KTX 역세권으로의 주거이동 특성”, 『주거환경』 제18권 제4호, 2020, pp. 77-89
- 노재준·임재빈, “도시철도 개통이 역세권 아파트 가격에 미치는 장기적 공간파급효과 분석”, 『부동산학연구』 제30권 제4호, 2024, pp. 7-26
- 박경, “글로벌 금융위기 전후 한국 사회 공정성 담론의 변화: 빅데이터 기법을 활용한 미디어 담론 분석”, 『사회과학연구』 제48권 제2호, 2022, pp. 31-54
- 박보현, “한국 지방정부의 헤게모니 프로젝트 II: 강원도의 축적전략과 2018 평창동계올림픽”, 『한국스포츠사회학회지』 제26권 제2호, 2013, pp. 85-105
- 박상민·나철원·최민성·이다희·온병원, “Bi-LSTM 기반의 한국어 감성사전 구축 방안”, 『지능정보연구』 제24권 제4호, 2018, pp. 219-240
- 박재명, “국내 포털 뉴스 빅데이터 내용 분석을 통한 생활체육 달리기 트렌드 분석 연구”, 『스포츠사이언스』 제43권 제2호, 2025, pp. 203-210
- 박천보, “TOD(대중교통중심)형 도시개발 특성 및 발전전략에 관한 연구”, 『한국산학기술학회논문지』 제17권 제2호, 2016, pp. 635-641
- 서울특별시, 『역세권 직·주·락 활성화 전략』, 2022
- 성현곤, “대중교통 중심 개발(TOD)이 주택가격에 미치는 잠재적 영향”, 『지역연구』 제27권 제2호, 2011, pp. 63-76
- 송희원·김상혁, “고속철도 개통이 관광수요에 미치는 영향: KTX 강릉선을 중심으로”, 『이벤트컨벤션연구』 제18권 제1호, pp. 203-221
- 이선우·오종열, “사회네트워크 분석을 이용한 KTX 역세권 개발사업의 지연 요인 연구: 서울특별시 역세권개발사업을 대상으로”, 『한국도시설계학회지 도시설계』 제19권 제1호, 2018, pp. 51-63
- 이은별·전진오·백지선, “서울의 다문화 공간 연구: 뉴스 빅데이터 분석 시스템 ‘빅카인즈’를 이용한 국내 언론의 외국인 마을 보도(1990-2016) 분석”, 『미디어 경제와 문화』 제15권 제2호, 2017, pp. 7-43
- 이제용·김영표, “KTX 강릉선 개통 이후 거리 속성이 관광산업 활성화, 지역 발전, 지역 이미지에 미치는 영향”, 『한국철도학회 논문집』 제23권 제10호, 2020, pp. 989-1000
- 이천재, “텍스트 마이닝을 활용한 한국지적정보학회지의 연구동향 탐색”, 『한국지적정보학회지』 제26권 제2호, 2024, pp. 22-40
- 임동일, “강원도 3대 도시의 인구이동 결정요인 분석: 춘천, 원주, 강릉을 대상으로”, 『한국콘텐츠학회논문지』 제11권 제1호, 2011, pp. 411-421
- 정명진, “‘지방소멸’담론의 한계와 권력 효과: 지역일간지의 사설 보도를 중심으로 한 비판적 담론분석”, 『한국언론정보학보』 제132호, 2025, pp. 185-226
- 정현·전희정, “한국 지역균형발전의 담론에 관한 탐색적 연구: LDA 토픽모델링 분석을 활용하여”, 『한국행정연구』 제34권 제4호, 2025, pp. 148-179
- 조운승·이태겸, “언론매체 빅데이터 분석을 통한 민선자치시대의 한강공원 정책과 사회적 담론 변화 연구”, 『한국지적정보학회지』 제23권 제1호, 2021, pp. 173-187
- 조재욱·우명제, “고속철도 개통이 지역경제 및 균형발전에 미치는 영향: 대한민국 KTX 경부선·경전선을 중심으로”, 『국토계획』 제49권 제5호, 2014, pp. 263-278
- 주영희·정사랑·Wu Jun, “텍스트 마이닝 기법을 활용한 순환경제 빅데이터 분석: 빅카인즈 국내 언론 보도 중심으로”, 『미래기술융합논문지』 제4권 제5호, 2025, pp. 145-153
- 하미란·진장익, “KTX 개통이 인구변화에 미치는 영향에 관한 연구”, 『한국지역개발학회지』 제36권 제1호, 2024, pp. 117-138
- 허재완·김갑성·유예진, “고속철도(KTX)의 입지효과 및 개통으로 인한 수도권 인구집중 완화효과 분석”, 『국토계획』 제53권 제4호, 2018, pp. 107-122
- 홍성조·김용건, “고속철도 개통에 따른 강원도 지역경제 변화 예측”, 『도시행정학보』, 제30권 제3호, 2017, pp. 47-72
- 한국언론진흥재단 빅카인즈(BIGKinds), <https://www.bigkinds.or.kr/>
- Calthorpe, P., The Next American Metropolis: Ecology, Community, and the American Dream, Princeton Architectural Press, 1993
- Cervero, R., S. Murphy, C. Ferrell, N. Goguts, Y. H. Tsai, G. B. Arrington, J. Boroski, J. Smith-Heimer, R. Golem, P. Peninger, E. Nakajima, E. Chui, R. Dunphy, M. Myers, S. McKay, and N. Witenstein, “Transit-oriented development in the United States: Experiences, challenges, and prospects,” TCRP Report 102, Transportation Research Board, 2004.
- Cournoyer-Gendron, M., “The transit-oriented development model in Montreal (Canada): Mobilizing a concept and negotiating urban development at the local and metropolitan scale,” Environnement Urbain/Urban Environment, Vol. 12, 2017
- Douay, N. Olivier Roy-Baillargeon, “Transit-Oriented Development (TOD): Vector or Mirage of Metropolitan Planning and Governance Transformations in the Greater Montreal Area?”, Flux, Vol. 101102 No. 2, 2015, pp. 29-41
- Eom, J. K., K. S. Lee, J. Y. Song, and J. Lee, “Analysis of mobile phone data to compare mobility flows and

- hotspots before and after the opening of high-speed railway: case study of honam KTX in Korea,” *Applied Sciences*, Vol. 10 No. 14, 2020, Article 5009.
33. Gunder, M., “Lacan, planning and urban policy formation,” *Urban Policy and Research*, Vol. 23 No. 1, 2005, pp. 87-107
34. Han, J., and Yunho Ji, “The effect of the opening of the KTX Gyeonggang highspeed railway line in improving perceptions of PyeongChang in Korea leading to sustainable tourism,” *African Journal of Hospitality, Tourism and Leisure*, Vol. 9 No. 2, 2020, pp. 1-18
35. Ibraeva, A., G. H. A. Correia, C. Silva, and A. P. Antunes, “Transit-oriented development: A review of research achievements and challenges,” *Transportation Research Part A: Policy and Practice*, Vol. 132, 2020, pp. 110-130
36. Jamme, H. T., J. Rodriguez, D. Bahl, and T. Banerjee, “A twenty-five-year biography of the TOD concept: From design to policy, planning, and implementation,” *Journal of Planning Education and Research*, Vol. 39 No.4, 2019, pp. 409-428
37. Jeong, H., and Haejoo Chung, “Public Opinion on Healthy Cities: Topic Modeling of South Korean News Articles (2004-2024),” *Journal of Urban Health*, 2025
38. Kim, J. H. “Spatial planning and economic development in border region: The experiences of Gangwon Province, Korea,” *Ordnungspolitische Diskurse*, No. 2010-05, 2010
39. Kim, J. S., and Nina Shin, “Planning for Railway Station Network Sustainability Based on Node-Place Analysis of Local Stations,” *Sustainability*, Vol. 13 No. 9, 2021, Article 4778.
40. Kwak, H., Jisun An, and Yong-Yeol Ahn, “A systematic media frame analysis of 1.5 million New York Times articles from 2000 to 2017,” *Proceedings of the 12th ACM Conference on Web Science (WebSci '20)*, 2020, pp. 305-314
41. Park, E. L., and Sungzoon Cho, “KoNLPy: Korean natural language processing in Python,” *Proceedings of the 26th annual conference on human & cognitive language technology*, 2014, vol. 6, pp. 133-136
42. Rojas, A., “Train stations’ impact on housing prices: Direct and indirect effects,” *Transportation Research Part A: Policy and Practice*, Vol. 181, 2024, Article 103979.
43. Salton, G., and Christopher Buckley, “Term-weighting approaches in automatic text retrieval,” *Information Processing & Management*, Vol. 24 No. 5, 1988, pp. 513-523
44. Seo, M., and Dongyoub Lee, “Typological differences in railway station areas according to locational characteristics: A nationwide study of Korea,” *Sustainability*, Vol. 13 No. 8, 2021, Article 4310
45. Shin, Dong-Chun, “Recent experience of and prospects for high-speed rail in Korea: Implications of a transport system and regional development from a global perspective,” *Working Paper*, No. 2005-02, 2005

<국문요약>

강원권 주요 도시 역세권개발 담론 및 쟁점 분석
: 춘천, 원주, 강릉역세권을 중심으로

정 진 태 (Jeong, Jintae)
김 서 정 (Kim, Seojeong)
이 천 재 (Lee, Cheonjae)

본 연구는 강원권 주요 도시(강릉·원주·춘천)의 역세권개발 담론이 시기에 따라 어떻게 변화해 왔는지를 뉴스 빅데이터를 통해 규명하였다. 빅카인즈에서 ‘역세권’과 도시명을 결합해 2010~2026년 기사 8,991건을 수집한 뒤, 규칙기반 정제와 두 가지 관련성 기준의 민감도 검토를 거쳐 확정된 동일 모집단 6,738건에 TF-IDF, 의미연결망·프레임 분석, KNU감성사전 기반 감성분석을 결합·수행하였다. 분석 결과 동일 키워드로 수집되었음에도 세 도시의 담론 구성은 상이하였다. 강릉은 올림픽·KTX 개통에서 광역 교통망으로 담론 축이 이동하였고, 원주는 부동산 분양과 철도 인프라 담론이 병존하였으며, 춘천은 경춘선 개통 이후 문화·관광으로 분산되었다가 개발 의제가 재부상하는 순환적 양상을 보였다. 감성분석에서는 세 도시 모두 중립 기사의 비중이 가장 높은 가운데 긍정이 부정을 앞섰고, 원주(40.1%)·강릉(35.3%)·춘천(31.3%) 순의 차이가 관찰되었다. 다만 감성점수의 절대값은 문서 길이와 사전 편향에 민감하므로 해석은 상대 비교에 한정하였다. 이를 바탕으로 강릉을 ‘관광·이벤트 기반 성장거점형’, 원주를 ‘산업·계획도시 기반 성장거점형’, 춘천을 ‘수도권 연계·광역접근성 강화형’으로 역세권개발 담론을 유형화하였다.

주 제 어 : 역세권개발, 담론분석, TF-IDF, 감성분석, 뉴스 빅데이터