



**STARTUP
ALLIANCE**

**뉴스 기사를 통해 본
지역 스타트업 생태계 담론**

뉴스 기사를 통해 본 지역 스타트업 생태계 담론

스타트업얼라이언스 이지영 전문위원^{01 02}

01 정보시스템학 박사

02 이 글은 스타트업얼라이언스의 공식 입장이 아닌, 필자의 주관적인 견해를 밝힙니다.

그동안 우리 사회에서 국토균형발전 및 지역경제 활성화를 위한 논의가 다방면에서 이뤄져왔다. 이러한 논의에 기반해 정부는 지역 특성에 맞는 혁신 생태계의 필요성을 파악하고, 다양한 정책적·제도적 지원을 통해 각 지역에 혁신 성장의 토대가 될 스타트업 생태계를 조성하기 위해 노력해왔다. 본 연구는 뉴스 기사 텍스트마이닝을 통해 그간 지역 스타트업 생태계에 대해 사회적 담론이 어떠한 관점에서 이뤄져왔는지 파악하고자 한다. 분석대상은 2015년 1월 1일부터 2023년 12월 31일 사이에 작성된 뉴스 기사 중에서 ‘지역 스타트업 생태계’ 키워드를 포함한 뉴스 기사 18,474건이다. LDA 토픽모델링 기법을 통해 도출된 토픽은 ‘중앙 주도 육성’, ‘지역 주도 육성’, ‘금융 방안’, ‘글로벌’, ‘특화 산업’, ‘교육 및 인적자원’ 6개로 나타났다. 본 연구결과는 우리 사회가 지역 스타트업 생태계 조성과 활성화를 위해 지속해야 할 담론이 무엇인지, 논의를 좀 더 집중해야 할 부분은 무엇인지 파악하는데 필요한 기초자료로 활용될 수 있을 것이다.

목차 Table of Contents

	한 눈에 보는 이슈페이퍼	4
1	들어가며	7
	1-1 배경	
	1-2 연구목적	
2	연구방법	12
	2-1 연구설계 및 분석 데이터	
	2-2 데이터 분석방법	
3	연구결과	16
	3-1 기사량 추이 및 키워드 분석	
	3-2 토픽 분석	
4	나가며	23
	참고문헌	26

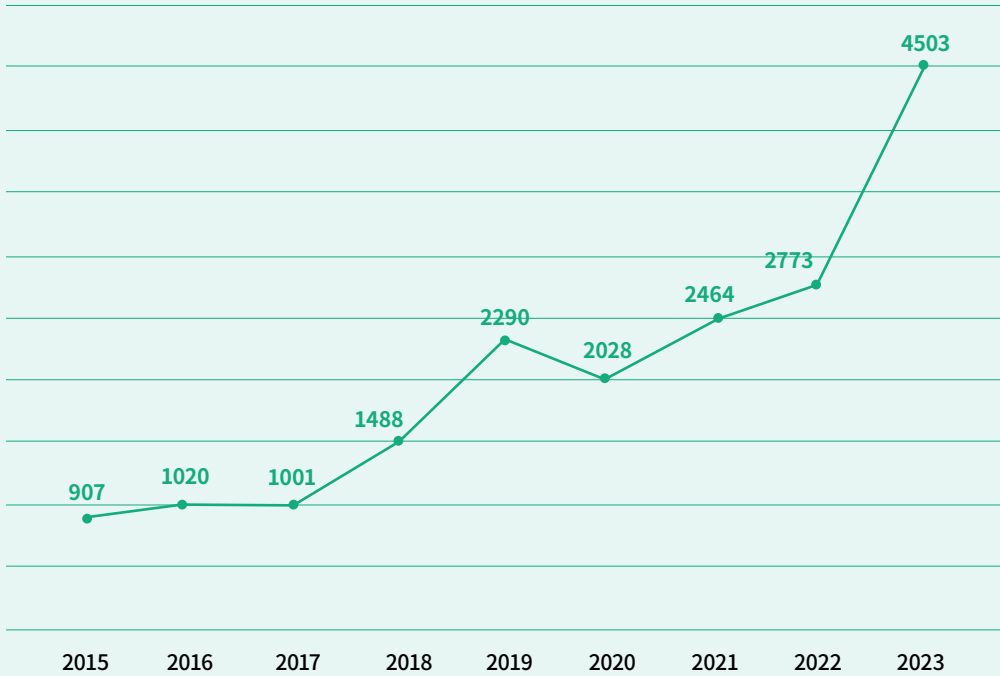
한눈에 보는 이슈페이퍼

1 지역 스타트업 생태계 관련 뉴스기사는 2015년부터 2023년까지 연평균 22.2% 증가함

○ 지역 스타트업 생태계에 대한 사회적 관심이 지속적으로 늘어나고 있음

지역 스타트업 생태계 뉴스 기사량 추이

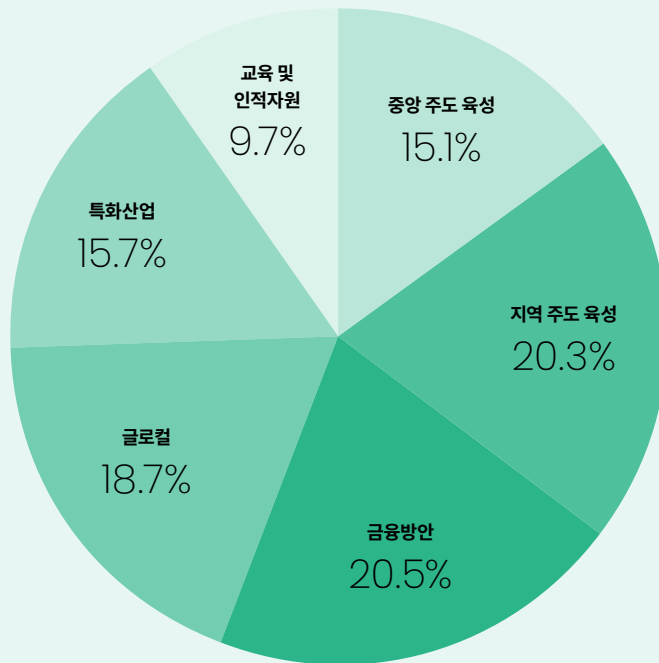
(단위: 건)



2 ‘지역 스타트업 생태계’ 키워드를 포함한 뉴스 기사에서 주로 논의된 토픽은 1) 중앙 주도 육성, 2) 지역 주도 육성, 3) 금융 방안, 4) 글로벌, 5) 특화 산업, 6) 교육 및 인적자원 이었음

- 그동안 지역 스타트업 생태계에 대한 사회적 논의가 토픽 내용과 키워드 빈도 면에서 중앙정부가 주도하는 방향으로 이뤄져왔다는 사실을 뉴스 기사를 통해 확인할 수 있음
- 중앙정부의 지역 스타트업 생태계에 대한 관심과 지원을 바탕으로 지역 혹은 민간 주도의 자생적인 활동이 이어지는 선순환 구조를 갖추는데 집중해야 할 것임

지역 스타트업 생태계 뉴스 토픽



1

들어가며

1-1

배경

1-2

연구목적

1 들어가며

1-1 배경

그동안 우리 사회·경제·문화적 환경에 따라 국토균형발전 및 지역경제 활성화 위한 노력이 다방면에서 진행됨

- 1980년대 경제성장과 함께 지역균형발전에 대한 정책목표가 수립되기 시작했으며, 1990년대 들어서 비로소 지역균형발전 정책이 국가 주요 정책의제에 오르게 됨
- 1997년 IMF 외환위기 이후 전환점을 맞은 지역 정책은 지역별 환경과 산업의 특성을 고려해 지역특화산업을 육성하는 방향으로 바뀜
- 2000년대부터는 전세계적으로 지식 기반 경제가 대두되었고, 각 광역시도에 TP, 벤처기업단지가 조성되는 등 지역도 지식 기반 경제에서 글로벌 경쟁력을 갖추는 방향으로 지역발전전략이 추진됨
- 2010년 이후에는 혁신 스타트업이 새로운 국가경제 성장동력이 되면서 광역시도에 창조경제혁신센터가 설립되었고, 기술 기반 스타트업과 더불어 로컬 크리에이터, 로컬 브랜드 등과 같이 지역의 다양성에 기반한 창조적 창업 활동도 확산되기 시작함

현재 정부는 지역 특성에 맞는 혁신 생태계의 필요성을 파악하고 다양한 정책적·제도적 지원을 통해 각 지역에 혁신성장의 토대가 될 스타트업 생태계를 조성하기 위해 노력하고 있음

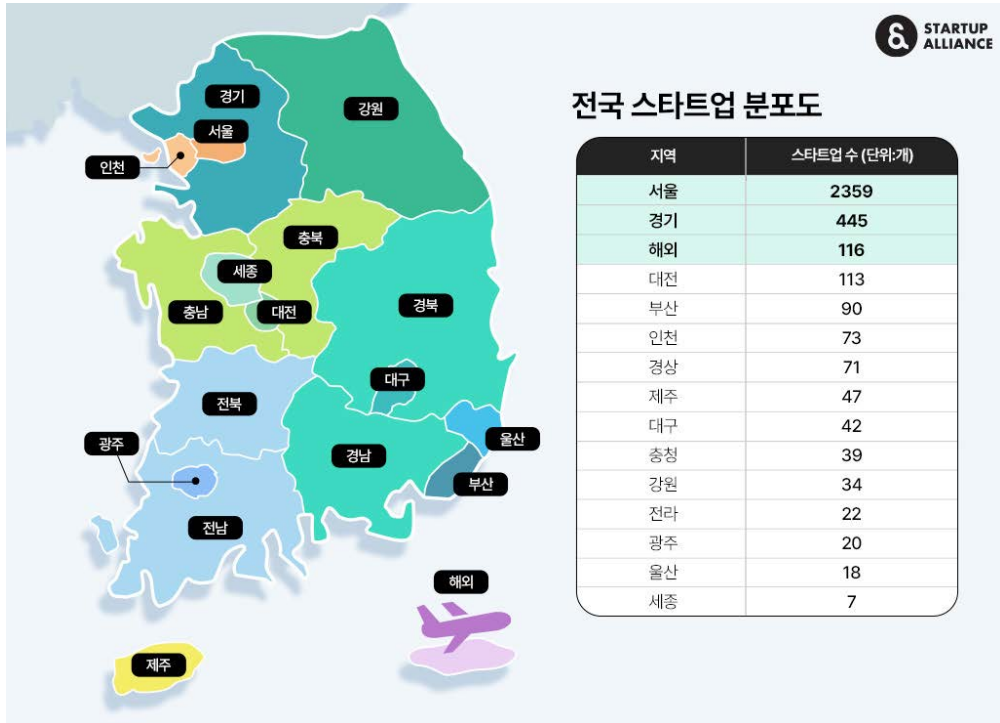
- 특히 2015년 광역시도에 창조경제혁신센터가 설립되면서 본격적으로 지역 스타트업 생태계에 대한 관심과 지원이 늘어나는 계기가 됨
- 그럼에도 불구하고 수도권과 지역의 스타트업 생태계의 격차는 여전히 큰 상황이며, 스타트업의 수와 투자유치 등 현황으로 그 사실을 알 수 있음

스타트업얼라이언스 스타트업 주소 분석 리포트⁰³에 따르면, 벤처투자를 받은 국내 스타트업의 82.3%(2877개)가 수도권에 위치해있음

- 수도권에 이어서 부산·울산·경상 지역이 5.1%(179개), 대전·충청 지역이 4.3%(152개)을 차지함
- 투자를 받은 스타트업의 수 뿐만 아니라, 대규모 투자를 유치한 스타트업도 서울 지역에 많았음. 1000억원 이상 투자받은 스타트업의 78.3%(47개)는 서울에 위치해 있으며, 300억원 이상 1000억원 미만 투자받은 스타트업의 69%(116개), 100억원 이상 300억원 미만 투자받은 스타트업의 71.5%(328개)가 서울에 위치
- 서울에 이어 경기 지역에도 대규모 투자를 받은 스타트업이 많이 위치함. 1000억원 이상 투자받은 스타트업의 6.7%(4개), 300억원 이상 1000억원 미만 투자받은 스타트업의 11.9%(20개), 100억원 이상 300억원 미만 투자받은 스타트업의 11.1%(51개)가 경기도에 위치

03 스타트업얼라이언스. (2023). 스타트업 주소 분석 리포트. URL: <https://brunch.co.kr/@startupalliance/435>

그림1. 스타트업 지역 분포



출처: 스타트업얼라이언스(2023)

1-2

연구목적

지역 스타트업 생태계의 필요성이 대두되면서, 이와 관련한 다양한 사회적 논의가 있었고 관련 뉴스도 많이 보도되었음

- 뉴스는 대중과 사회가 커뮤니케이션하는 방법 중 하나로서, 지역 스타트업 생태계에 대해서도 뉴스를 통해 사회적 담론이 형성되어 왔다고 볼 수 있음
- 관련 뉴스 기사를 수집해 분석한다면, 민간, 공공, 정책, 산업 등 어떤 관점에서 사회적 논의가 형성되어 왔는지 파악할 수 있음

따라서 뉴스 기사 토픽모델링을 통해 지역 스타트업 생태계를 위한 발전적 논의에 필요한 기초자료를 마련하고자 함

- 그동안 우리 사회에 지역 스타트업 생태계를 어떠한 관점에서 바라보고 사회적 논의가 이어졌는지 실증적으로 파악하고자 함
- 앞으로 지역 스타트업 생태계의 조성과 발전을 위해 지속해야 할 담론은 무엇이며, 새롭게 논의를 시작하고 집중해야 할 부분은 무엇인지 파악하고자 함

2

연구방법

- 2-1 연구설계 및 분석데이터
- 2-2 데이터 분석방법

2 연구방법

2-1 연구설계 및 분석 데이터

(연구대상) 지역 스타트업 생태계 관련 뉴스 기사 분석을 위해 다음과 같이 연구대상을 선정

- 온라인 플랫폼에서 서비스되는 뉴스 기사를 연구의 대상으로 함
- 검색어: ‘지역 스타트업 생태계’ 키워드를 포함한 뉴스 기사를 수집
- 보도시기: 각 지역에 창조경제혁신센터가 활발히 설립되기 시작한 2015년부터 보도된 뉴스 기사 (2015년 1월 1일~2023년 12월 31일)

(데이터 수집) 위에서 제시한 기준으로 데이터 수집을 진행함

- 뉴스 기사는 한국언론진흥재단의 뉴스 기사 데이터베이스인 빅카인즈⁰⁴에서 수집
- 개별 검색어를 포함한 뉴스 기사를 수집하고, 중복된 기사는 취합 후 삭제
- 최종적으로 총 18,474건의 뉴스 기사가 수집됨

⁰⁴ 한국언론진흥재단의 빅카인즈(BIGKinds)는 빅데이터 분석 서비스를 제공하는 뉴스 기사 데이터베이스로, 종합일간지, 경제지, 지역일간지, 방송사 등 54개 언론사의 뉴스 기사 메타데이터를 제공함(<https://www.bigkinds.or.kr>)

표1. 데이터 수집 개요

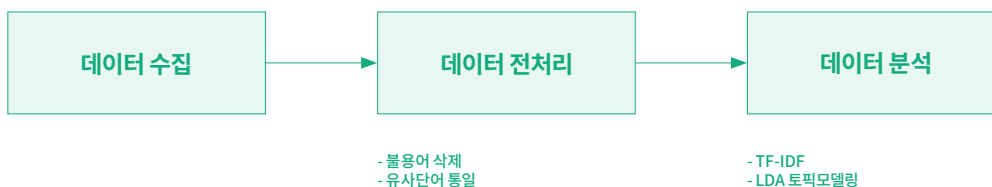
검색어	지역 스타트업 생태계
보도시기	2015년 1월 1일~2023년 12월 31일
수집출처	빅카인즈(https://www.bigkinds.or.kr)
분석대상	뉴스 기사 18,474건
분석툴	파이썬(Python)

2-2 데이터 분석방법

(분석과정) 본 연구는 데이터 수집-데이터 전처리-데이터 분석 순으로 진행⁰⁵

- 데이터 전처리 단계에서는 의미가 불분명한 단어 등 불용어를 삭제하고, 의미가 유사한 단어들을 통일하는 데이터 정제를 진행하며, 데이터 정제는 데이터 분석과정과 함께 반복적으로 실행해 분석결과의 정확도를 높임
- 데이터 분석단계에서는 기초 통계 분석, TF-IDF 분석, 토픽모델링 분석 등을 통해 기사 텍스트에 숨겨진 의미를 도출함

그림2. 데이터 분석과정



⁰⁵ 뉴스 기사 데이터베이스인 빅카인즈는 언론사의 뉴스 기사 메타데이터와 함께 기사에서 단순 숫자를 제외한 모든 명사를 추출한 결과도 제공하고 있어, 본 연구에서 뉴스 기사의 토큰화(tokenization), 명사 추출의 과정은 생략되었음

TF-IDF(Term Frequency-Inverse Document Frequency; 단어 빈도-문서 역빈도) 분석법은 상대적으로 중요한 키워드 도출을 위해 활용함

- TF-IDF 값은 TF(단어 빈도)와 IDF(문서 역빈도)를 곱한 값으로 계산함
- 문서에서 흔하게 나타나지 않는 단어가 특정 문서에서 자주 사용될수록 큰 값을 가짐

본 연구에서는 토픽모델링 알고리즘 중에서 LDA(Latent Dirichlet Allocation, 잠재 디레클레 할당) 기법을 통해 토픽을 추출

- 토픽모델링은 방대한 양의 문서 내에서 일정한 패턴을 찾아, 같은 집합이라고 추론되는 단어들을 추출하는 알고리즘
- LDA는 문서를 구성하는 단어들이 독립적이지 않다는 가정 하에 확률적으로 동일한 토픽에 해당할 가능성이 높은 단어들의 집합을 결과값으로 추출하는 알고리즘⁰⁶
- 일관성 점수(coherence score), 복잡성 점수(perplexity score)를 기반으로 토픽 수를 조정하며 최적화된 모델을 결정

06 Blei, D. M. (2012). Probabilistic topic models. Communications of the ACM, 55(4), 77-84.

3

연구결과

- 3-1 기사량 추이 및 키워드 분석
- 3-2 토픽분석

3 연구결과

3-1 기사량 추이 및 키워드 분석

(기사량 추이) 지역 스타트업 생태계 관련 뉴스 기사는 2015년 1월부터 2023년 12월까지 지속적인 증가세를 보임

- 2015년부터 2023년까지 뉴스 기사량의 연평균 증가율은 22.2%였으며, 2023년에 특히 기사량이 급증함
- 2023년에는 정부가 스타트업 코리아 종합대책(관계부처 합동, '23.08.)」을 통해 벤처·스타트업이 디지털 경제를 선도하는 글로벌 창업대국을 실현한다는 비전을 선포했으며, '지역창업 클러스터 활성화 및 지역 벤처투자 환경 보완'이 핵심전략에 포함됨
- 이 밖에도 2023년 발표된 여러 지역균형발전 정책⁰⁷에서 인구를 유입시키고 정주요건을 마련하고자 스타트업 생태계 조성을 주요 전략으로 삼고 있음
- 정책과 예산이 마련됨에 따라 관계부처, 지자체, 창업지원기관 등에서 지역 스타트업 생태계와 관련한 사업을 진행했고, 결과적으로 언론의 관심도 늘어난 것으로 판단됨

⁰⁷ 제1차 지방시대 종합계획 2023-2027년(지방시대위원회, '24. 11.), 제2차 혁신도시 종합발전계획 2023-2027년(지방시대위원회, '24. 11.) 등

2015년 이후 지역 스타트업 생태계에 대한 사회적 관심이 크게 늘어났음에도 불구하고 지역 창업기업 수는 증가하지 않음

- 지역 스타트업 생태계 기사량은 연평균 22.2% 증가한 반면, 수도권 외 지역 창업기업 수는 연평균 성장률이 0%에 수렴하는 상황

그림3. 연간 기사량 및 창업기업 증감 추이

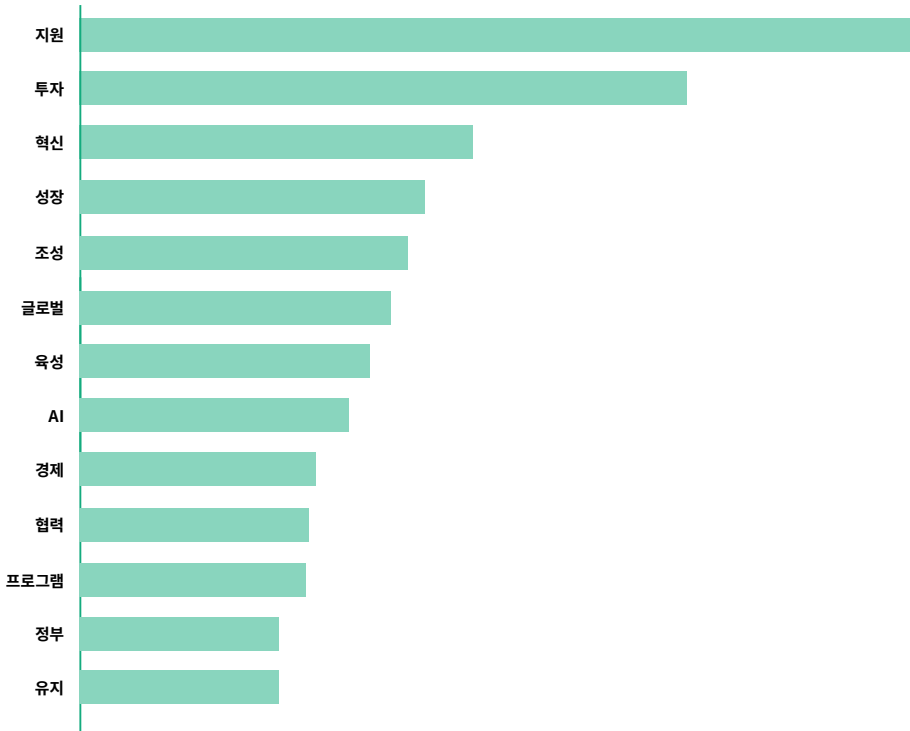


출처: 중소기업벤처부(2016~2023)

(키워드 분석) 지역 스타트업 생태계 뉴스 기사에 등장하는 키워드를 빈도에 따라 개별적으로 추출함

- 지원, 투자, 혁신, 성장, 조성, 글로벌, 육성, AI, 경제, 협력, 프로그램, 정부, 유치순으로 활용 빈도가 높은 것으로 나타남
- 상위 빈도를 차지하는 키워드가 대부분 스타트업 생태계의 트렌드보다는 공공주도의 상황을 표현하는 용어인 경향이 있음

그림4. 키워드 분석 결과



3-2 토픽 분석

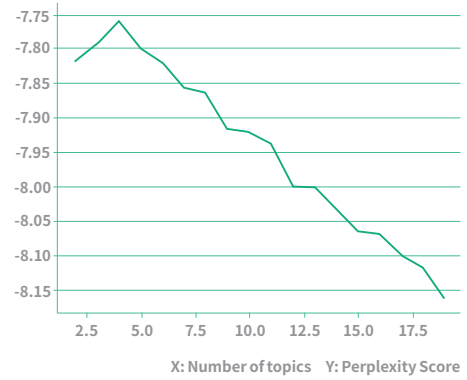
토픽모델링 모델을 평가하고 최적화하기 위해 토픽 수에 따른 일관성 (coherence)과 복잡성(perplexity) 점수를 확인함

- 일관성 점수가 높을수록 토픽이 의미론적으로 일관성이 높으며, 토픽 수가 3개 일 때 일관성 측면에서 가장 의미있는 결과를 가지고 5개 이상일 때 일관성 점수가 다시 높아짐
- 복잡성 점수가 낮을수록 해당 모델이 결과를 정확하게 예측하며, 토픽 수가 많을수록 복잡성은 낮아지는 것으로 나타남

그림 5. Coherence score



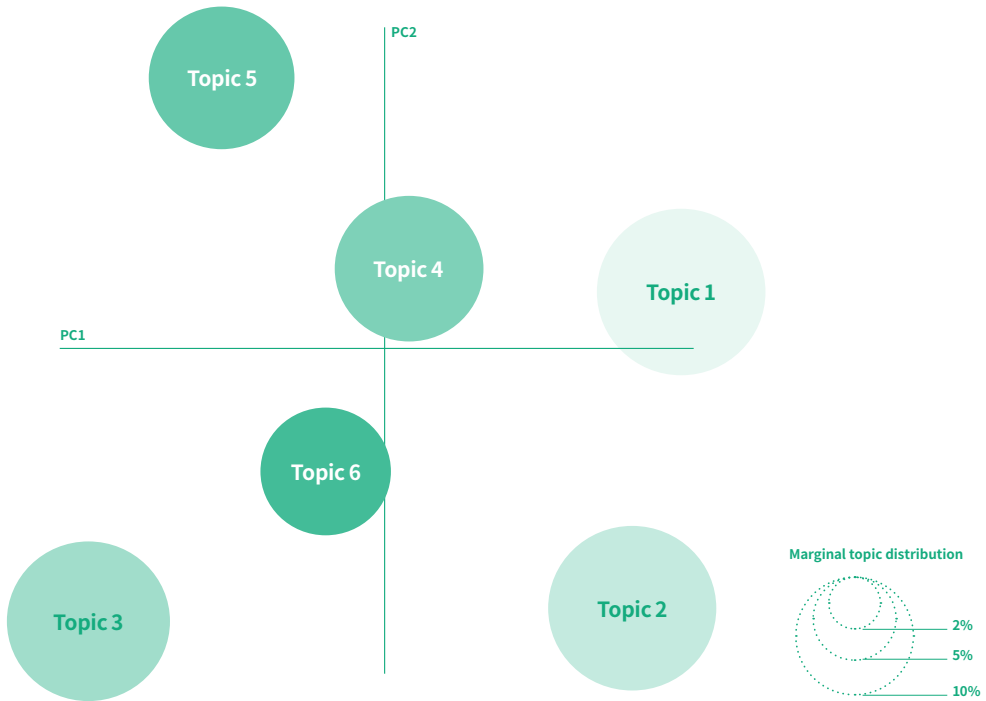
그림 6. Perplexity score



본 연구에서는 일관성 점수가 높은 모델 중 복잡성이 낮은 토픽 수 6개로 결정함

- 6개 토픽들의 구조적 관계와 연관성을 파악하기 위해 2차원 시각화를 실행한 결과, 토픽 간 유사성이 낮고 고르게 분포한 것을 확인함
 - 원의 크기: 각 토픽에 해당하는 단어들이 어떻게 분포되어 있는지 의미함
 - 원의 거리: 토픽 간의 유사성, 원이 겹칠수록 토픽 간 유사성이 높음

그림7. 토픽모델링 시각화 결과



최종적으로 지역 스타트업 생태계에 대한 뉴스 기사 토픽모델링 결과는 표3.과 같음

- (토픽 1) 중앙 주도 육성, (토픽 2) 지역 주도 육성, (토픽 3) 금융 방안, (토픽 4) 글로벌, (토픽 5) 특화 산업, (토픽 6) 교육 및 인적자원, 총 6개 토픽을 추출함
 - 토픽 1: 중앙 주도 육성
중앙정부에서 발표하거나 주도하는 지역 스타트업 생태계 정책 및 사업을 의미함 (예시 기사제목: 정부, 신성장산업에 100조원...정책금융 총 180조원 지원)
 - 토픽 2: 지역 주도 육성
지자체, 지역 창업지원기관 등 지역에서 추진하는 사업을 의미함 (예시 기사제목: 부산시, 2021년까지 500억 투자 VR·AR산업 육성)
 - 토픽 3: 금융 방안

지역에 스타트업 생태계를 조성하고 창업과 투자의 선순환을 이끌어내기 위한 금융적 측면의 방안으로, 펀드 조성, 자금 지원, 투자활동 등의 내용을 포함함 (예시 기사제목: 청년창업펀드 9개사 100억 투자받았다)

— 토픽 4: 글로벌

글로벌 경쟁력을 갖춘 지역 스타트업 생태계를 조성하기 위해 스타트업의 해외 진출과 해외 투자 유치 등을 지원하는 내용을 포함함 (예시 기사제목: 아시아 스타트업 연결되면 더 큰 가능성 만들 수 있어)

— 토픽 5: 특화 산업

지역별 강점을 살릴 수 있는 특화 산업에 대한 내용을 포함함 (예시 기사제목: 녹슨 철강 대신 로봇-생명과학 피츠버그에 '새 피'가 돈다)

— 토픽 6: 교육 및 인적자원

지역 스타트업 생태계에 필요한 인재를 육성하고 외부의 인재를 모으기 위한 정책에 대한 내용을 포함함 (예시 기사제목: 대학이 지역 창업 거점으로 정부도 정책도 '지역 협력' 방점)

- 토픽별 비중을 보면, 금융 방안(토픽 3) 20.5%, 지역 주도 육성(토픽 2) 20.3%, 글로벌(토픽 4) 18.7%, 특화 산업(토픽 5) 15.7%, 중앙 주도 육성(토픽 1) 15.1%, 교육 및 인적자원(토픽 6) 9.7% 순으로 나타남

표2. 토픽모델링 결과

토픽	주요 키워드
토픽 1 중앙 주도 육성 (15.1%)	경제, 정부, 정책, 지역, 산업, 대통령, 일자리, 중소기업, 사회, 국민
토픽 2 지역 주도 육성 (20.3%)	창업, 지원, 부산, 혁신, 광주, 진출, 인천, 프로그램, 성장, 유치
토픽 3 금융 방안 (20.5%)	펀드, 조성, 금융, 경제, 투자, 지원, 진출, 확대, 글로벌, 미래
토픽 4 글로벌 (18.7%)	시장, 글로벌, 아시아, 한국, 중국, 미국, 투자, 행사, 경제, 지원
토픽 5 특화 산업 (15.7%)	산업, 지역, 스마트, 에너지, AI, 융합, IoT, 드론, 자동차, 경제
토픽 6 교육 및 인적자원 (9.7%)	대학, 협약, 발굴, 교육, 프로그램, 콘텐츠, 육성, 지원, 청년, 조성

4

나가며

4 나가며

본 연구에서는 뉴스 기사 토픽모델링을 통해 그간 지역 스타트업 생태계에 관해 어떠한 사회적 논의가 이뤄져왔는지 파악하고자 했음

- 2015년 1월 1일부터 2023년 12월 31일까지 온라인 플랫폼을 통해 유통된 뉴스 기사 중에서 지역 스타트업 생태계와 관련한 검색어를 포함한 기사 18,474건을 분석함
- 키워드 분석을 위해서는 빈도, 상대적 중요도(TF-IDF) 값을 도출했고, 토픽 분석을 위해서는 LDA 토픽모델링을 활용함

지역 스타트업 생태계 관련 뉴스 기사는 2015년부터 2023년까지 지속적인 증가세를 보임

- 뉴스 기사량의 연평균 증가율은 22.2%였으며, 2023년에 특히 기사가 급증한 것으로 나타남
- 그 이유는 2023년 정부에서 선포한 글로벌 창업대국을 실현하기 위한 핵심전략에 ‘지역창업’과 ‘지역 벤처투자’가 포함되었고, 같은 해 발표된 여러 지역균형발전 정책에서도 지역 스타트업 생태계 조성을 주요 전략으로 삼으며 다양한 사업들이 진행됐기 때문으로 보임

지역 스타트업 생태계에 대한 사회적 관심과 다르게, 지역 창업기업의 수는 2016년부터 2023년까지 연평균 증가율이 0%에 수렴하는 등 지역 생태계가 성장했다고 보기 어려움

- 정부가 정책적 목표를 가지고 지역 스타트업 생태계를 지속적으로 지원해왔으나, 아직까지는 지역 현장에 실질적인 변화를 이끌어내지 못한 상황으로 생각됨

뉴스 미디어를 통해 형성된 지역 스타트업 생태계에 대한 사회적 논의를 파악하기 위해 뉴스 기사 토픽모델링을 진행한 결과, 6개의 토픽이 제시됨

- (토픽 1) 중앙 주도 육성, (토픽 2) 지역 주도 육성, (토픽 3) 금융 방안, (토픽 4) 글로벌, (토픽 5) 특화 산업, (토픽 6) 교육 및 인적자원 토픽을 추출함
- 현재 지역에서 정책·사업을 운영하는 데 있어, 중앙정부의 재정 지원이 큰 비중을 차지하고 독자적인 전략 수립과 추진이 어려운 상황임을 감안하면, 토픽 2가 온전히 지역 주도적인 움직임이라고 보기는 어려움
- 상위 빈도를 차지하는 키워드들이 대부분 스타트업 생태계의 트렌드보다는 공공 주도의 상황을 반영하는 용어인 경향이 있음
- 결과적으로 그동안 지역 스타트업 생태계에 대한 사회적 논의는 중앙정부가 주도하는 방향으로 이뤄져왔다는 사실을 뉴스 기사를 통해 확인할 수 있음

지역 스타트업 생태계가 조성되기 위한 핵심은 중앙정부의 지원없이 지속될 수 있는 자생력에 있음

- 지난 8년 간 중앙정부로부터 지역 스타트업 생태계에 대한 많은 관심과 지원이 이어진 점은 긍정적이며, 그동안의 관심과 지원을 바탕으로 지역 혹은 민간 주도로 자생적인 선순환 구조를 갖추는데 집중해야 할 것임
- 지역사회에 있어 스타트업 생태계를 통한 선순환 구조는 지속가능한 지역경제를 가능하게 하고, 지역소멸을 막을 수 있는 핵심요인으로 작용할 수 있음
- 서울이나 수도권과 다른 각 지역만의 스타트업 생태계 모델을 만들어 나갈 필요가 있으며, 이를 위해서는 지역별 편더멘탈과 특성에 대한 면밀한 진단도 선행되어야 함

- 지금까지 중앙정부의 주도로 이뤄진 사회적 담론과 노력이 지역과 학계는 물론 산업계 플레이어들과 함께 다양한 관점으로 확장되어 지역 스타트업 생태계를 위한 협업 체계까지 구축되어야 할 것임

참고문헌

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77-84.

CB Insights. (2023). State of Venture Q2'23 Report

스타트업얼라이언스. (2023). 스타트업 주소 분석 리포트. URL: <https://brunch.co.kr/@startupalliance/435>

스타트업얼라이언스. (2024). 스타트업, 지역의 미래일까?: 지역과 스타트업이 서로에게 묻다

중소기업벤처부. (2016). 창업기업동향

중소기업벤처부. (2017). 창업기업동향

중소기업벤처부. (2018). 창업기업동향

중소기업벤처부. (2019). 창업기업동향

중소기업벤처부. (2020). 창업기업동향

중소기업벤처부. (2021). 창업기업동향

중소기업벤처부. (2022). 창업기업동향

중소기업벤처부. (2023). 창업기업동향



STARTUP ALLIANCE

발행인 스타트업얼라이언스

기획/제작 스타트업얼라이언스

발행처 스타트업얼라이언스

서울특별시 강남구 봉은사로 215,

11층(논현동, KTS 빌딩)

<https://startupall.kr>

발행일 2024.8.31.

ISSN 2982-4834

이 책의 저작권은
스타트업얼라이언스에 있으며
무단복제와 전재를 금합니다.

