

특허 정보를 활용한 지역 유망 기술 도출: 인천 지역 바이오 산업을 중심으로

김 건¹ · 김동관² · 최호철³ · 윤호열^{3*}

¹한국기계연구원 기업지원실, ²인천테크노파크 기업성장센터,

³한국화학연구원 전략기술정책센터

Identifying Regional Emerging Technologies Using Patent Information: Focus on Incheon Bio Industry

Keon Kim¹ · Dongkwan Kim² · Hochull Choe³ · Ho-Yeol Yoon^{3*}

¹Korea Institute of Machinery & Materials, ²Incheon Techno Park,

³Korea Research Institute of Chemical Technology

■ Abstract ■

This study proposes a scientific methodology for regional technology strategy development by identifying emerging technologies through the analysis of patent information. Given the resource constraints often encountered by regions, it is essential to establish evidence-based strategies for efficient R&D planning. To overcome the limitations of previous research, we applied a combined approach utilizing topic modeling and CPC network analysis techniques, while also enhancing the reliability of our findings through validation by relevant experts. The results of our analysis identified key emerging technologies in the bio industry of Incheon, specifically in fields such as anti-cancer therapies, food, and biosimilars. These technology areas are currently in high demand on a global scale and are closely aligned with the strategic industrial infrastructure of Incheon. The findings of this study are expected to support the development of regional technology strategies and science and technology policies based on scientific evidence. Furthermore, the proposed methodology provides a practical framework for identifying emerging regional technologies, which can be adapted for use in other regions facing similar challenges. This framework not only facilitates informed decision-making for policymakers but also provides a systematic approach for regions aiming to enhance their technological competitiveness and innovation capabilities in a rapidly evolving global landscape.

Keywords : Regional Emerging Technology, Topic Modeling, LDA, Network Analysis, Technology Strategy

논문접수일 : 2024년 09월 03일 논문게재확정일 : 2024년 09월 20일

논문수정일 : 2024년 09월 13일

* 교신저자, hoyeol@kricr.re.kr

1. 서 론

지역의 기술 역량에 상응할 수 있는 전략적 미래 유망 기술 발굴은 지역경쟁력 강화를 위한 중요한 과제이다[3, 7, 15]. 지역 자체적으로 Bottom-up 방식의 유망 기술을 발굴하여 중앙부처에 제안하고 이를 정책화하는 것은 지역의 특수성을 반영한 실효성 있는 기술 발굴, 발굴 과정에서 지역 혁신기관 간 연계협력 강화, 공동 사업 추진을 통한 중앙-지방 간 협력 등 다양한 이점을 제공한다. 정부는 기존 Top-down 방식 지역 정책의 한계점을 극복하고자 지역 주도 혁신 생태계 구축과 산업 육성을 골자로 하는 혁신 체계의 개편을 시도하고 있다[19].

지역 주도의 과학기술 및 산업 육성 계획(지방과학기술종합계획 등) 수립을 위해서는 중점 투자 유망 기술 분야를 전략적으로 선정하는 것이 필요하다. 성공적인 지역과학기술정책을 추진하기 위해 지역은 스스로 강점 역량을 분석하고 최대화하기 위한 자발적인 혁신의 추진이 필요하다[16]. 이를 바탕으로 지역 내 유망 기술의 성장을 지속적으로 확인하고, 해당 기술이 지역 산업과 얼마나 밀접하게 연계되어 발전하는지를 검토하는 절차가 지속되어야 한다. 이러한 접근은 지역의 과학기술 발전을 체계적으로 지원하며, 지속 가능한 혁신 생태계를 조성하는 데 기여할 수 있다[13]. 또한, 기획 과정에서 동일한 분야가 다수의 지역에 중복으로 투자되는 것을 회피하고 R&D 효율성을 강화하기 위해서는 지역별 특색에 적합한 기술 분야를 선정하고 중점 투자하는 것이 중요하다. 기술 전문화 전략은 기존 기술을 활용하여 고도화에 집중하는 전략인데[20], 한정된 자원을 활용해야 하는 지역의 특성 상 유용하다.

이러한 측면에서 증거 기반 정책의 추진은 지역 과학기술로 육성하기 위한 중점 분야 선정에 필수적인 요소이다. 특히, 투자 과정에서 불확실성을 낮추기 위해 데이터 기반 R&D 기획의 중요성은 지속해서 증가하고 있다[25]. 증거 기반 정책은 가장 효율적인 집행을 추구하는데, 그로 인해 증거나 데이터는 가능한 한 포괄적이어야 한다[37]. 데이터와 증거

에 기반하여 지역의 기술 전략 방향성을 제안한 관련 연구에서는 연구개발사업[2, 4, 14, 15], 특허[3, 9, 10, 13, 18] 등 다양한 기술 정보 데이터를 활용하고 있다. 이 중 특허는 기술 지식의 중요한 정보를 포함하는 데이터로 R&D 투자 결정을 위한 기술경영에 중요한 원천으로 활용된다[27, 31]. 특히, 기술의 잠재적 경쟁력을 평가하여 전략적 기술 기획에 활용되는 매우 유용한 지표다[27, 28].

인천은 과거 전통적인 제조업 중심 경제 구조를 가지고 있었으나, 최근 급변하는 기술혁신 환경에 따라 주력 산업의 기술변화 대응 방안 마련이 중요한 과제로 떠오르고 있다. 제조업 비중은 감소하고 있으며, 산업 구조의 서비스화 등 근본 구조가 변화하는 추세에 놓여있다[12]. 이에 인천시는 2015년 '인천 미래 성장동력 8대 전략산업 육성방안'을 선정하여 육성하였으며, 최근에는 4차 산업혁명과 디지털, 탄소중립 사회의 변화를 반영하여 2023년 바이오, 반도체, 로봇, 디지털·데이터, 미래차, 항공으로 주요 전략산업을 재선정, 이를 기반으로 산업육성정책을 추진하고 있다[13]. 인천의 핵심전략산업 중 하나인 바이오는 바이오시밀러 및 바이오의약품 제조 분야에서 두드러진 성과를 보이고 있으며, 글로벌 기업들이 바이오 생산 허브로서 역할을 수행하고 있다. 바이오 산업체의 집중으로 인해 자연스럽게 클러스터가 형성되었고, 지방자치단체의 지원을 통해 대형 바이오 클러스터로 성장하였다. 그 결과, 많은 대기업들이 송도에 위치하게 되었고, 인천은 세계 최대의 바이오 생산 도시로 도약하여 명실상부한 바이오 도시로 자리매김하고 있다[2].

이에 본 연구는 인천 전략산업 중 바이오 산업을 중심으로 특허 데이터를 활용하여 주요 핵심 기술을 도출하고 지역과학기술정책 및 기술 전략의 방향성을 제안하고자 한다. 특허 출원 지역을 중심으로 인천 데이터를 추출하였으며, 한국표준산업분류표를 기준으로 바이오 산업과 매칭을 통해 분석 데이터를 구축하였다. 비정형 데이터인 특허 자료가 내포하고 있는 의미를 도출하기 위해 토픽 모델링과 네트워크 분석을 활용하였으며, 인천 바이오 산업의 유망 기

술을 도출하고 증거에 기반한 지역 정책 및 전략 방향성을 제안하였다.

2. 선행 연구

유망 기술은 급진적 참신성, 빠른 성장, 일관성 등 특징을 갖는 기술로 다양한 문헌에서 정의되고 있다 [24]. 선행 연구에서는 유망 기술을 도출하기 위해 논문, 특허 등 다양한 데이터 접근 방법과 기술 수명 주기 분석, 시계열 분석, 과학계량학, 기계학습 등 예측을 위한 방법론이 활용되고 있으나[6, 21, 22, 38], 지역 단위 유망 기술은 일반적인 유망 기술 도출과 다르게 접근된다. 지역 유망 기술은 지역이 잘할 수 있는 특화된 기술을 중심으로 혁신성장동력의 축을 육성하는데 목적을 둔으로서 미래 예측적인 분석 방향보다 현재 강점이 갖고 있는 분야 도출에 중점을 둔다[17]. 김현우 외[7]는 부산 지역 미래 유망 산업 도출을 위해 특허 정보를 분석하였다. 지역 정책적으로 육성 중인 산업 분야의 관련 특허를 과학계량학 기법으로 분석하여 향후 미래 유망산업으로 육성 가능한 후보군을 제안하였다. 한 장협, 김채복[18]은 경북 지역 ICT 유망 기술 도출을 위해 특허 데이터를 분석하였다. 특허 지표를 활용한 지역 기술 동향 분석과 포트폴리오 분석을 통해 유망 기술을 도출하였다. 연구 결과의 신뢰성 제고를 위한 후속 연구 방향으로 특허 정량 분석 결과와 지역 전문가 검증을 혼합하는 방법을 제안하였다. 박장훈, 옥영석[10]은 창원 지역 수소 전기차 분야 유망 기술 도출을 위해 특허 정보를 활용하였다. 특허 출원 성장률과 출원 빈도를 통해 지역 주력 산업의 공백 기술과 유망 기술을 도출하였으며, 이를 토대로 지역산업에 대한 시사점을 제안하였다.

최근 인천은 산업 환경 변화 등 지역적 특성으로 인해 데이터 기반 정책 방향성 제안을 위한 연구가 추진되어 왔다. 오준병, 허원창[12]은 지역혁신정책 방향성 제안을 위해 인천 지역의 지식 재산 현황을 분석하였다. 지역에 입지한 기업 특허를 분석 대상으로 하였으며, 1980년부터 2014년까지 장기간 데이터

를 활용하였다. 분석 결과를 토대로 지역의 기술혁신 정책 방향성에 대해 제안하였다. 김동관, 이정영[3]은 인천 지역 기술 동향을 분석하기 위해 2000년부터 2022년까지 인천에서 출원된 특허 데이터를 분석하였다. Wipson DB를 활용하였으며, 네트워크 분석 기법을 적용하여 협력적 특허분류(Cooperative Patent Classification; 이하 CPC)¹⁾ 간 관계를 분석하여 주요 융합 기술들을 도출하였다. 도출된 결과를 토대로 지역 과학기술정책의 방향성에 대해 제안하였다. 유광민, 김동관[13]은 인천 지역 6대 전략산업(디지털·데이터, 로봇, 바이오, 반도체, 미래차, 항공)의 핵심 기술과 기술 융합을 분석하였다. 2018년부터 2022년 기간 특허 데이터의 국제특허분류(International Patent Classification; 이하 IPC)²⁾ 데이터를 활용하여 네트워크 분석 기법을 적용해 전략산업 간 기술 융합 네트워크를 도출하였다. 융합기술 중심의 투자 전략 방향성을 제안하였으나, 세부적인 수준의 분석이 수행되지 못한 한계를 언급하였다. 박소현 외[9]는 경제지리학적 관점에서 특허 데이터를 활용하여 인천 바이오 산업의 경로 창출 과정을 분석하였다. 2000년부터 2021년까지 KIPRIS 특허 데이터를 활용하였으며, 주요 기술로 제약·화장품 혁신, 스마트 진단, 의료기기 및 살균 솔루션, 식음료 가공 기술 등 4개 토픽을 도출하였다. 분석 결과를 토대로 산업 간 경로 관계를 분석하고 다양한 정책 방안을 제안하였다. 다만, 산업 경로에 초점을 두어 IPC 네트워크 분석 등은 수행하지 못한 한계를 남겼다.

종합하자면, 지역 유망 기술은 지역이 기술적 강점을 갖고 있으며, 지속해서 성장하는 분야로 주력 및 전략산업과 연관있는 분야가 제안되고 있다. 또한, 지역의 기술 전략이기 때문에 특허가 분석 데이터로서 주로 활용되는데, 기존 유망 기술 도출을 위한 미래 예측 분석보다 특허의 성장성, IPC, CPC를 활용한 융합성 분석 등 현재 강점 도출을 위한 분석이 주를 이루고 있다. 특히, 본 연구의 분석 대상인

1) 국제특허분류(IPC)보다 세분화된 특허분류체계.

2) 발명 기술 분야를 표기하는 국제적으로 통일된 분류 체계.

인천 지역으로 한정할 경우, 다양한 연구가 제안되었으나 바이오 산업의 세부 중점 유망 기술 도출을 위한 복합적인 방법론을 적용한 연구는 제한적이다. 토픽모델링은 문서별 단어의 분포를 중심으로 토픽을 추정한다[5]. 토픽모델링만 활용할 경우, 유망 기술 분야를 도출할 수 있으나, 단어를 기반으로 하기 때문에 세부 기술을 파악하기 어렵다는 한계를 가진다[9, 33]. 또한, 특허분류체계 네트워크 분석에 기반한 유망 기술 도출은 세부 기술 분야 도출이 가능하나, 세부 기술 분야는 다양한 응용 분야에서 활용될 수 있기에 해당 기술을 중심으로 구성된 유망 기술의 구체적인 선정이 어렵다는 한계가 있다[3, 13]. 특히, 기존 특허분류체계 기반 연구는 4자리 분류 체계를 주로 활용하고 있다. 이에 본 연구는 선행 연구의 한계를 극복하고, 인천 지역 전략 산업인 바이오 산업의 세부 유망 기술을 도출하기 위해 특허 데이터를 적용하였으며, 분석 방법론으로 토픽모델링과 CPC 6자리 네트워크 분석을 함께 적용한다. 마지막으로 분석 결과에 대한 지역 정책 전문가의 검증을 통해 연구 결과의 신뢰도를 제고하였다.

3. 데이터 및 분석 절차

3.1 분석 데이터

본 연구의 목적은 인천 지역의 바이오 산업에서 주요 유망 기술을 도출하는 것으로, 국내 특허를 대상으로 최근 5년간(2018년~2022년)을 분석 기간으로 설정하였다. 데이터 수집 및 선별은 다음과 같은 절차로 진행되었다. 첫째, 특허 검색 엔진인 Keywert를 활용하여 2018년 1월 1일부터 2022년 12월 31일까지 출원된 국내 특허 중 출원인의 주소가 인천광역시인 특허를 1차 선별하였다. 둘째, 한국표준산업분류표에서 바이오에 해당하는 표준 산업명을 선별하고, Raw Data에서 각 특허의 메인 IPC를 추출하였다. 최종적으로 특허청에서 제공하는 산업(한국표준산업분류, Korean Standard Industrial Classification; 이하 KSIC)-특허(IPC) 연계표에 근

거하여 각 특허의 메인 IPC를 기준으로 매칭한 후 바이오 산업 분야로 그룹화하는 방법으로 수행하였다. 데이터 수집 절차를 통해 3,314건의 특허가 분석 대상으로 활용되었다.

3.2 분석 절차

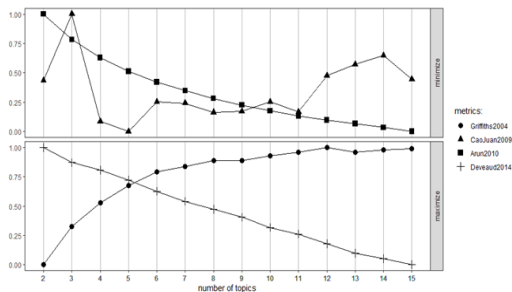
특허 정보를 활용한 인천 바이오 산업 핵심 기술 도출을 위해 본 연구에서는 분석 대상 특허에 수록된 국문 초록을 기준으로 토픽모델링을 실시하였고, 이를 통해 분류된 각각의 토픽별 특허의 CPC에 대한 의미연결망분석을 실시하였다. CPC는 IPC보다 확장된 분류 체계로 더 많은 기술 범주를 포함하며, 패밀리 특허 분석에 유용하다[26, 32].

토픽모델링은 비정형 텍스트 데이터에 내포된 핵심 주제를 도출하고 의미를 도출하는 텍스트 분석 기법이다[1, 5, 8]. 토픽모델링은 R(4.2.3)을 통해 텍스트 데이터인 특허 국문 초록에 대한 전처리, 토픽 수 결정, 토픽모델링 분석 및 분석 결과 해석 순으로 실시되었다.

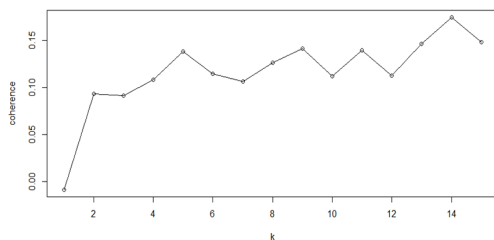
첫 번째 단계인 텍스트 데이터 전처리를 위해 가장 먼저 특허 국문 초록에 포함된 숫자, 영어, 기호를 제거하였다. 이후 특허 국문 초록에 존재하는 명사를 추출한 후 불용어(Stopword)를 제거하였다. 본 연구에서 불용어로 판단된 단어들은 특허 초록에 일반적으로 등장하는 단어(e.g. 발명, 상기, 기술 등)들과 단어 빈도분석을 통해 특허 초록에 빈번히 등장하나, 분석 자체에는 유의미하지 않다고 판단되는 단어(ex: 효능, 개선, 등)들이 포함되었다. 그리고 이렇게 전처리가 완료된 텍스트 데이터를 문서-단어 행렬(Document term matrix)로 변환하였다.

전처리 다음 단계로, 적정 토픽 수를 설정하기 위한 작업을 실시하였다. 토픽모델링 과정에서 적정 토픽 수를 설정하는 것은 매우 중요한 단계이다. 이는 토픽 수를 적정 수준보다 작은 수로 설정하는 경우 하나의 토픽에 여러 내용들이 함께 포함되어 일관성이 없어지거나, 많은 수로 설정하는 경우에는 각 토픽들이 해석이 불가능할 정도로 과도하게 세분

화될 수 있기 때문이다[34]. 이에 본 연구에서는 토픽 수 설정을 위한 대표적인 통계지표인 복잡도(Perplexity)와 응집도(Coherence) 두 가지 지표를 확인하여, 적정 토픽 수를 설정하였다[11, 34]. 복잡도는 토픽 내부의 무질서 정도를 측정하는 것으로, 하나의 토픽에 연관 단어들이 포함된 경우 무질서 정도가 낮아지고, 반대 상황에서는 무질서 정도가 높아짐을 가정하는 지표이다. 하지만 이는 토픽 수가 많아지면 많아질수록 무질서 정도가 낮게 산정된다는 한계점이 존재한다[35]. 즉 복잡도만을 기준으로 토픽 수를 설정한 경우 과도하게 많은 토픽 수가 분석의 대상이 될 수 있는 것이다. 이에 본 연구에서는 이러한 복잡도의 한계점을 보완하고자 토픽에 포함된 단어들이 의미적으로 얼마나 잘 연결되어 있는가를 의미하는 응집도를 함께 확인하였다. 그 결과 아래의 <그림 1> 및 <그림 2>와 같이 토픽 수가 5개인 지점에서 산출된 복잡도와 응집도 지표가 토픽을 분류함에 있어 가장 적절한 값으로 나타났다. 이에 본 연구에서는 토픽모델링을 위한 토픽 수를 5개로 결정하였다.



<그림 1> 토픽 수에 따른 복잡도



<그림 2> 토픽 수에 따른 응집도

토픽 수 5개를 기준으로 토픽모델링은 잠재 디리클레 할당(Latent Dirichlet Allocation; 이하 LDA) 방식으로 실행되었다(Parameter 추정: Gibbs sampling). 토픽모델링의 여러 유형들 중 LDA는 생성 확률 모델(Generative probabilistic model)을 활용한 방식이다. 이는 수많은 단어의 확률분포로 표현되는 토픽들의 집합이 문서로서 나타남을 가정한다. 이에 문서-단어행렬에 기초하여 각 문서에서 나타난 단어들의 조합을 통해 토픽을 설정, 토픽에 단어가 포함되는 확률을 디리클레(Dirichlet) 분포를 통해 계산하는 방식이다[23, 29]. 이러한 LDA는 다른 토픽모델링 방식들에 비해 알고리즘이 단순하고 확장성이 높으며, 논문이나 특허의 초록, 신문 기사 등 장문의 텍스트 분석에 효율적이기에 사회과학 분야에서 널리 사용되고 있다[29]. 이에 본 연구에서도 LDA 방식을 통해 토픽모델링을 실시하였다. 그리고 이를 통해 분류된 토픽에 포함된 단어들을 확인하여 토픽별 명칭을 부여하였다.

이후 각각의 토픽에 포함된 특허들이 어떠한 기술적 특징 및 구성을 가지고 있으며, 이 중 중심적 역할을 수행하고 있는 기술 분야가 무엇인지 확인하기 위해, Gephi(0.10)를 활용하여 각 토픽에 포함된 특허 CPC에 대한 의미연결망 분석을 실시하였다.

특정 CPC가 연결망 내부에서 갖는 영향력을 확인하기 위한 지표로는 연결(Degree), 매개(Betweenness), 아이겐벡터(Eigenvector) 중심성을 활용하였다.

이는 연결중심성을 통해서만 특정 노드(Node)가 다른 노드들과 직접적으로 연결된 정도를 확인할 수 있으며, 매개중심성을 통해서만 어떤 노드가 노드들 간의 간접적 인용관계를 형성하는지, 즉 매개적 역할을 수행하고 있는지 확인할 수 있기 때문이다. 마지막으로 아이겐벡터 중심성의 경우 이를 통해 노드들 간 단순 연결 관계가 아닌 그와 연결된 노드들이 갖는 중심성 값들의 합에 의해 산출되어, 어떤 노드가 연결망 내부에서 실제로 영향력이 높은 노드와 연결되어 있는지 확인할 수 있기 때문이다[30, 36].

4. 분석 결과

4.1 토픽 모델링

분석 대상 특허 3,314건의 국문 초록에 대한 토픽 모델링을 실시한 결과, 이에 수록된 내용은 <표 1>과 같은 단어들로 구성된 5가지 토픽으로 분류되었다. 분류된 토픽의 토픽명을 명명하기 위한 절차는 다음과 같이 수행되었다. 첫째, 4인 이상의 과학기술 정책, 산업정책, 기술경영 분야 박사 인력을 중심으로 중심 단어들과 특허를 확인하여 개별적으로 토픽명을 명명하였다. 둘째, 각 분야 전문가에 의해 명명된 토픽명을 공유하여 상호 확인하고, 토론을 통해 수정·보완하는 절차를 거쳤다. 최종적으로 인천 지역 산업 전문가의 검토를 통해 각 토픽의 명칭을 항암 치료 및 생체 반응 연구, 식품 가공 및 저장 기술, 피부 건강, 방사선 치료, 호흡기 보호 및 감염 예방으로 명명하였다.

4.2 의미연결망분석

토픽1은 전체 3,411개의 특허 중 가장 많은 특허인 798개(24.1%)가 포함된 토픽이다. ‘항암치료 및 생체 반응 연구’로 명명되었다. 의미연결망분석 결과 토픽1에 포함된 특허들의 CPC 연결망 내부에서 연결 및 아이겐벡터중심성이 가장 높게 나타난 CPC는 ‘항종양제(A61P35/00)’였다. 그 다음으로는 ‘음식의 구성요소 및 공정(A23V2002/00)’의 연결 및 아이겐벡터중심성이 높게 나타났으나, 1순위에 비해 중심성 수치가 확연히 낮게 나타났다. 특히 연결중심성의 경우 3, 4순위인 ‘20개 이상의 아미노산을 갖는 펩타이드, 가스트린 등(C07K14/005)’ 및 ‘운동기능증정(A61B5/1124)’과 같은 값을 갖는다. 매개중심성의 경우 ‘컴퓨터 보조진단(G16H50/20)’과 ‘유전물질 변이에 의한 질병 측정 또는 시험 공정(C12Q1/6883)’으로 질병 진단과 관련된 기술이 1, 2순위로 나타났다.

〈표 2〉 토픽1. CPC 연결망 중심성

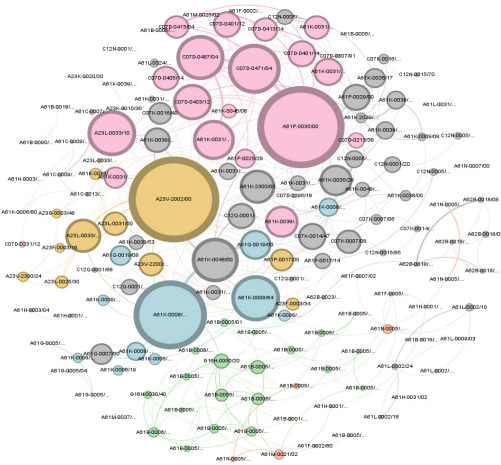
순위	중심성(수치)		
	연결	매개	아이겐벡터
1	A61P35/00 (0.269)	G16H50/20 (0.235)	A61P35/00 (1.000)
2	A23V2002/00 (0.209)	C12Q1/6883 (0.194)	A23V2002/00 (0.794)
3	C07K14/0005 (0.209)	A23V2002/00 (0.133)	C07D471/04 (0.662)
4	A61B5/1124 (0.209)	A61P35/00 (0.132)	A61K31/519 (0.626)
5	A61Q19/08 (0.194)	A61N5/0616 (0.123)	C07D487/04 (0.549)

이를 통해 토픽1에 포함된 특허들의 경우 ‘항종양제’가 가장 핵심적인 기술영역이며, 질병 진단과 관

〈표 1〉 토픽모델링 결과

토픽	토픽명	주요단어(Beta)	특허 수(%)
1	항암 치료 및 생체 반응 연구	단백질(0.020), 증식(0.009), 유전자(0.007), 면역(0.006), 생체(0.006), 항체(0.006), 치아(0.005), 구강(0.005), 효소(0.005), 항암(0.004)	798 (24.1)
2	식품 가공 및 저장 기술	식품(0.011), 성형(0.007), 식감(0.007), 수분(0.006), 냉각(0.006), 영양(0.006), 설탕(0.006), 냉동(0.005), 포장(0.005), 비타민(0.005)	622 (18.8)
3	피부 건강	염증(0.011), 항산화(0.009), 보습(0.009), 정제(0.008), 노화(0.006), 미백(0.006), 용액(0.006), 독성(0.006), 식물(0.006), 계면활성제(0.006)	717 (21.6)
4	방사선 치료	광원(0.008), 탄성(0.007), 마사지(0.006), 패드(0.005), 초음파(0.005), 충전(0.005), 휴대(0.005), 과장(0.005), 튜브(0.005), 헤드(0.004)	650 (19.6)
5	호흡기 보호 및 감염 예방	바이러스(0.011), 감염(0.01), 하우징(0.008), 향균(0.007), 자외선(0.007), 먼지(0.007), 호흡(0.007), 흡입(0.006), 세균(0.006), 필터(0.006)	527 (15.9)
전체			3,411 (100)

런된 ‘컴퓨터 보조진단’, ‘유전물질 변이에 의한 질병 측정 또는 시험 공정’이 다른 기술들을 연결하는 역할을 수행하고 있음을 알 수 있다. 즉 토픽1에 포함된 특허들의 경우 특정 기술의 영향력이 독보적으로 높고, 이와 별개의 기술이 다른 기술들 간의 관계를 연결하는 특징을 가진 연결망으로 구성되어 있다고 할 수 있다.



〈그림 3〉 토픽1. CPC 연결망 시각화

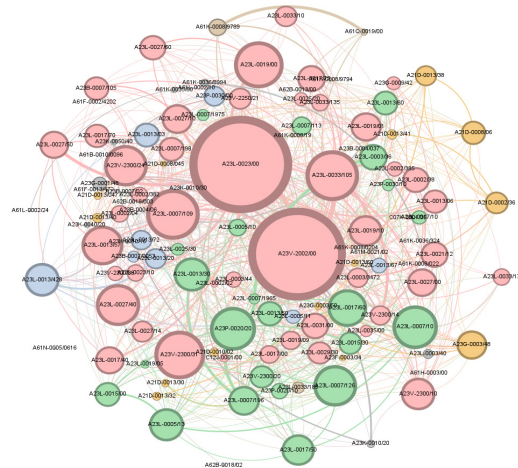
‘식품 가공 및 저장 기술’로 명명된 토픽 2의 경우 전체 특허 중 622개(18.8%)의 특허가 포함된 토픽이다. 의미연결망분석 결과 토픽2의 연결망 내부에서는 ‘음식의 구성요소 및 공정(A23V2002/00)’이 갖는 연결, 매개중심성이 가장 높게 나타났다. 또한 해당 CPC는 아이젠벡터중심성 또한 2순위로 높게 나타났고, 그 수치는 1순위와 크게 차이가 나지 않았다. 다음으로 ‘스프, 소스의 제조 또는 처리(A23L23/00)’가 연결중심성 2순위, 아이젠벡터중심성 1순위로 높게 나타났다.

이를 통해 토픽2에 포함된 특허들의 경우 ‘음식의 구성요소 및 공정’ 및 ‘스프, 소스의 제조 또는 처리’의 2가지가 가장 핵심적인 기술 영역이라 할 수 있다. 특히 ‘음식의 구성요소 및 공정’ 기술의 경우 다른 기술들을 연결하는 역할 또한 하는 것으로 나타나 중요성이 매우 높다고 할 수 있다. 즉 토픽2에 포

함된 특허들의 경우 특정 2가지 기술이 연결망 내부에서 절대적인 영향력을 가지고 있음을 확인할 수 있었다.

〈표 3〉 토픽2. CPC 연결망 중심성

순위	중심성(수치)		
	연결	매개	아이젠벡터
1	A23V2002/00 (0.252)	A23V2002/00 (0.272)	A23L23/00 (1.000)
2	A23L23/00 (0.235)	A61N5/0616 (0.131)	A23V2002/00 (0.953)
3	A23L19/00 (0.183)	A61Q19/00 (0.127)	A23L7/109 (0.521)
4	A23L27/40 (0.148)	A61K8/9789 (0.119)	A23L33/105 (0.521)
5	A23L33/105 (0.148)	A61L2/24 (0.088)	A23V2300/31 (0.497)



〈그림 4〉 토픽2. CPC 연결망 시각화

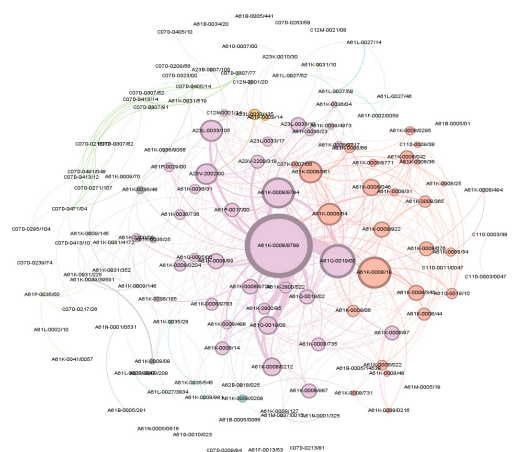
‘피부 건강’으로 명명된 토픽3은 전체 특허 중 717개(21.6%)의 특허가 포함되어, 토픽1 다음으로 가장 많은 특허가 포함된 토픽이다. 의미연결망분석 결과 토픽3의 연결망 내부에서는 ‘화장품 또는 유사 개인 위생용품의 썬덕익식물 제제(A61K8/9789)’의 연결, 매개 및 아이젠벡터 중심성이 모두 1순위로 높게 나타났다. 특히 ‘스킨 케어제(A61Q19/00)’가 연결 및 아이젠벡터중심성에서, ‘3개 이상의 이종원자 고리를 함유한 헥테로고리 화합물(C07D413/14)’가 매개

중심성에서 2순위로 나타났으나, 1순위가 갖는 중심성 수치와 큰 차이가 있음을 확인할 수 있다.

이를 통해 토픽3에 포함된 특허들의 경우 ‘화장품 또는 유사 개인위생용품의 썬떡잎식물 제제’가 가장 핵심적인 기술이며, 연결망 내부에서 독보적인 영향력을 가지고 있음을 확인할 수 있었다. 이는 어성초, 작약, 당나무 등 썬떡잎식물로 분류되는 식물들이 화장품 제제로서 다양하게 활용되고 있어, 이를 포괄하는 CPC가 다양한 기술들과 연결되고 다른 기술 간의 관계를 매개하는 등의 특징으로 인해 연결망 내부에서의 영향력이 강하게 나타난 것이라 판단할 수 있을 것이다.

〈표 4〉 토픽3. CPC 연결망 중심성

순위	중심성(수치)		
	연결	매개	아이젠벡터
1	A61K8/9789 (0.335)	A61K8/9789 (0.163)	A61K8/9789 (1.00)
2	A61Q19/00 (0.236)	C07D413/14 (0.086)	A61Q19/00 (0.520)
3	A61K8/9794 (0.193)	A61Q19/00 (0.076)	A61K8/19 (0.491)
4	A61K8/19 (0.149)	A23V2002/00 (0.067)	A61K8/9794 (0.477)
5	A23V2002/00 (0.137)	A61L27/52 (0.059)	A61K8/64 (0.348)

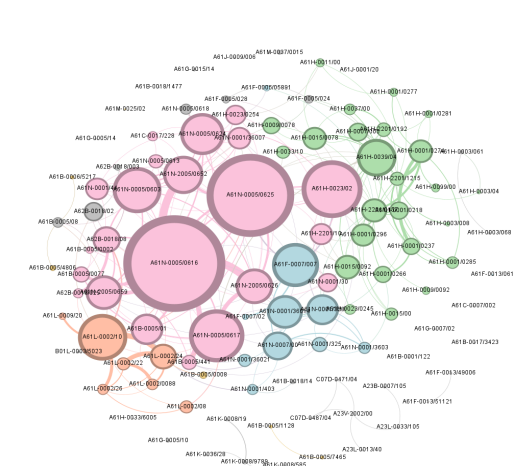


〈그림 5〉 토픽3. CPC 연결망 시각화

토픽4의 경우 ‘방사선 치료’로 명명되었고 해당 토픽에는 전체 특허 중 650개(19.6%)의 특허가 포함되었다. 의미연결망분석 결과 ‘그늘리는 것 외 피부 치료를 위한 방사선 치료(A61N5/0616)’의 연결, 매개, 아이젠벡터 중심성이 모두 1순위로 높게 나타났다. 연결, 아이젠벡터 중심성의 2순위는 ‘몸을 따뜻하게 하는 방사선 치료(A61N5/0625)’로 나타났고, 매개 중심성의 경우 ‘방사선 치료 등을 포함한 휴대가능한 물리치료 장치(A61H2201/0157)’가 2순위로 나타났다. 다만 매개 중심성의 경우 그 수치가 다른 토픽들의 의미연결망분석 과정에서 나타난 수치보다 상대적으로 낮게 나타나, 해당 토픽 내에서는 1, 2순위로 할지라도 실제 영향력은 낮을 가능성이 높을 것으로 판단된다.

〈표 5〉 토픽4. CPC 연결망 중심성

순위	중심성(수치)		
	연결	매개	아이젠벡터
1	A61N5/0616 (0.250)	A61N5/0616 (0.086)	A61N5/0616 (1.000)
2	A61N5/0625 (0.238)	A61H2201/0157 (0.077)	A61N5/0625 (0.855)
3	A61H2201/0157 (0.188)	A61N5/0625 (0.076)	A61H23/02 (0.594)
4	A61H23/02 (0.188)	A61H23/02 (0.066)	A61N5/0617 (0.543)
5	A61N2005/0652 (0.150)	A61H2201/1215 (0.064)	A61L2/10 (0.477)



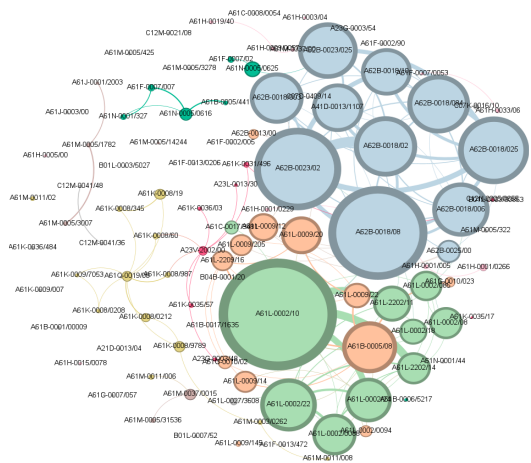
〈그림 6〉 토픽4. CPC 연결망 시각화

이를 통해 토픽4에 포함된 특허들의 경우 방사선 치료와 관련된 2가지 기술인 ‘그을리는 것 외 피부치료를 위한 방사선 치료’와 ‘몸을 따뜻하게 하는 방사선치료’가 핵심기술이라 할 수 있고, 기술들 간의 연결을 수행하는 특정 기술의 영향력은 매우 미미함을 확인할 수 있었다. 즉 토픽4의 특허들에 포함된 CPC 연결망의 경우 특정 2가지 기술이 다른 기술들과 여러 관계를 맺고 영향력이 높은 기술들과 근접하고 있으며, 특정 기술에 의해 기술들 간의 관계가 연결되기 보다는 각각의 기술들이 서로 직접적으로 연결된 특징을 가지고 있을 것이라 판단할 수 있다.

‘호흡기 보호 및 감염 예방’으로 명명된 토픽5는 전체 특허들 중 가장 적은 527개(15.9%)의 특허로 구성된 토픽이다. 의미연결망분석 결과 ‘가스마스크 또는 가스헬멧의 구성부품(A62B18/08)’이 연결중심성에서는 1순위, 아이젠벡터중심성에서는 2순위로 나타났다. ‘그을리는 것 외 피부 치료를 위한 방사선 치료(A61N5/0616)’의 매개중심성이 1순위로 가장 높게 나타났으며, 이와 근소한 차이로 2순위로 나타난 ‘자외선 조사를 통한 소독, 살균 장치(A61L2/10)’는 아이젠벡터중심성에서는 1순위로 높은 영향력을 가지고 있음이 확인되었다. ‘인공호흡기용 호흡보호 필터(A62B23/02)’의 경우 연결중심성에서는 2순위, 아이젠벡터중심성에서는 3순위로 나타났다.

〈표 6〉 토픽5. CPC 연결망 중심성

순위	중심성(수치)		
	연결	매개	아이젠벡터
1	A62B18/08 (0.171)	A61N5/0616 (0.154)	A61L2/10 (1.000)
2	A62B23/02 (0.143)	A61L2/10 (0.152)	A62B18/08 (0.839)
3	A61L2202/14 (0.129)	A61L2/22 (0.092)	A62B23/02 (0.739)
4	A61H5/00 (0.129)	A61K8/0212 (0.087)	A62B18/025 (0.601)
5	B04B1/20 (0.129)	A61M11/008 (0.081)	A62B18/02 (0.536)



〈그림 7〉 토픽5. CPC 연결망 시각화

이를 통해 토픽5에 포함된 특허들의 경우 ‘가스마스크 또는 가스헬멧의 구성부품’, ‘인공호흡기용 호흡보호 필터’, ‘그을리는 것 이외의 피부 치료를 위한 방사선 치료’, ‘자외선 조사를 통한 소독, 살균’의 4가지 기술이 연결망 내부에서 갖는 영향력이 높음을 확인할 수 있다. 매개중심성 측면에서는 토픽4의 핵심기술로 확인된 ‘그을리는 것 이외의 피부 치료를 위한 방사선 치료’ 및 ‘자외선 조사를 통한 소독, 살균 장치’ 기술이 비슷한 수치에서 1, 2순위의 영향력을 가진 것으로 확인되었다. 이는 ‘자외선 조사를 통한 소독, 살균 장치’와 유사하게 ‘그을리는 것 외 피부 치료를 위한 방사선 치료’의 기술 범위에도 음식물, 의약품 등의 세균이나 유해균 제거와 관련된 내용이 포함되어 있기에, 소독이나 살균 등을 목적으로 하는 기술들이 연결망 내부에서 다른 기술들 간의 관계를 연결하는 특징을 가지고 있는 것이라 판단된다.

5. 결론 및 시사점

외부 환경 변화에 따른 산업 구조 개편과 지역의 혁신 성장동력 창출을 위한 지역 정책의 필요성이 지속해서 강조되고 있다. 본 연구는 지역의 특색에 적합한 유망 기술 선정을 위해 특허 데이터를 중심

으로 토픽모델링과 의미연결망분석을 함께 활용한 복합적인 방법론을 제안하고, 인천 바이오 산업을 중심으로 분석을 수행하였다.

분석 결과, 인천 바이오 산업의 핵심 유망 기술로 항암 치료 및 생체 반응 연구, 식품 가공 및 저장 기술, 피부 건강, 방사선 치료, 호흡기 보호 및 감염 예방 등 항암, 바이오 시밀러, 식품 분야가 도출되었다. 인천은 인하대학교, 가천대학교 등 주요 대학에서 항암, 방사선 치료, 감염 예방 등 시민 건강과 관련된 연구를 활발히 진행하고 있다. 항암 치료 기술은 글로벌 바이오 시장의 성장세와 맞물려 전 세계적으로 높은 수요를 가지고 있는 분야이며, 인천 송도 지역에 삼성바이오로직스, 셀트리온 등 국내 대표적인 바이오 시밀러 기업들이 위치하고 있어 이와 관련된 유사 분야 특허들이 지속해서 출원되고 있다. 이러한 환경적 요소들로 인하여 항종양제(A61P35/00)와 같은 기술들이 다른 기술들에 비해 중심성이 높게 나온 것으로 추정되며, 항암, 바이오시밀러 등이 인천 바이오 산업의 유망 기술로 도출된 것으로 판단된다. 또한, 인천 산업단지 내 화장품 및 식품 가공 제조 기업들이 전국 평균보다 높은 밀도로 분포하고 있다. 식품 가공 및 저장 기술과 피부 건강이 유망 기술로 도출된 것은 인천이 주력 산업과 밀접한 연관이 있기 때문이다.

기존 선행 연구와 비교해보면 김동관, 이정영[3], 박소현 외[9], 유광민, 김동관[13]의 연구에서 도출된 바이오 산업 핵심 IPC인 A61K, A61L은 각각 본 연구에서 도출된 피부 건강 분야와 호흡기 보호 및 감염 예방 분야에서도 높은 중심성을 가지며 핵심 기술 분야로 나타났다. 특히 선행 연구와 비교하여 본 연구에서는 CPC 8자리까지 분석을 통해 세부적인 중심 기술 분야를 도출하였다. 피부 건강의 A61K 분야는 화장품 또는 유사 개인위생용품의 쌍떡잎식물 제제가 핵심적인 기술로 볼 수 있다. 지역이 강점을 보유한 기술로 바이오 관련 지역 산업 및 과학기술 정책을 수립함에 있어 반드시 고려되어야 하는 기술 중 하나로 판단된다. 마찬가지로 호흡기 보호 및 감염 예방의 A61L 분야는 자외선 조사를 통한 소독,

살균 장치가 아이젠벡터중심성 기준 높은 영향력을 보이는 기술 분야로 도출되었다. 아이젠벡터중심성은 연결된 CPC의 영향력을 함께 고려하여 산출되는 지표로 관련 분야 영향력이 지속해서 증가함을 시사한다. 마찬가지로 지역 기술 전략 방향성에 고려되어야 하는 분야로 볼 수 있다.

본 연구 결과를 바탕으로 인천 바이오 산업 발전을 위한 정책적 시사점은 다음과 같다. 첫째, 전략적 바이오 R&D 투자 강화가 필요하다. 본 연구에서 도출된 특허 네트워크를 통해 항암 치료, 질병 진단, 피부 건강, 방사선 치료, 감염 예방과 같은 핵심 기술이 바이오 산업에서 중요한 역할을 하고 있음을 확인할 수 있다. 이러한 핵심 기술 분야는 인천 바이오 산업의 경쟁력을 결정짓는 요소로 작용할 가능성이 높으며, 이를 더욱 발전시키기 위해 이들 기술을 중심으로 집중적인 R&D 투자가 이루어져야 한다. 둘째, 산업 간 융합을 통한 비즈니스 모델 창출이 필요하다. 본 연구에서 분석된 특허 네트워크는 바이오 의약품, 진단 기술, 식품 기술, 피부 건강 기술 간의 상호작용 가능성을 시사하며, 이를 기반으로 한 새로운 비즈니스 모델을 창출할 수 있는 기회가 존재함을 보여준다. 이를 위해서는 이러한 융합 기술이 원활하게 발전할 수 있도록 지원하는 정책 방안이 필요하다. 셋째, 바이오 산업 정책을 수립함에 있어 과학적으로 입증된 정보를 바탕으로 하는 데이터 기반 정책을 추진해야 한다. 본 연구에서 특허 데이터를 활용하여 인천의 바이오 핵심 기술을 분석하고 실질적인 정책 방향을 제시할 수 있었던 것처럼, 바이오 산업 정책에서도 데이터 기반 분석을 통한 정책 제언을 더욱 확대해 나가야 한다. 이를 통해 정책의 효과성을 최대화하고 실패 가능성을 최소화할 수 있을 것이다.

본 연구는 국내 표준인 한국표준산업분류를 활용하여 바이오 산업의 범위를 지정하고 지역의 주요 핵심 기술을 도출함으로써 포괄적인 데이터를 사용한 과학적 정책 제언에 기여한다. 본 연구에서 도출된 결과는 지역 산업 및 과학기술정책의 기초 자료로 활용될 수 있으며, 투자 방향 및 기술 로드맵 수립

의 증거로 활용될 것으로 기대된다. 또한, 기존 선행 연구의 한계를 극복하기 위해 토픽모델링, CPC 네트워크 분석, 그리고 전문가의 정성적 검증을 수행하는 복합적 방법론을 제안하며 지역 유망 기술 도출 연구의 방법론적 측면에서도 기여할 것으로 기대한다.

본 연구의 실무적, 학술적 기여에도 불구하고 연구 과정에서 제약된 부분을 보였기에 이는 향후 연구 방향으로 제안하고자 한다. 분석에 사용된 특허 데이터는 출원 후 1년 6개월 이상 공개되는 제도의 특성상 2022년 특허 중 미공개된 건이 많다는 점과 인천에서 특허를 출원했으나 본사가 다른 지역에 있는 경우 분석 대상에서 제외된다는 점으로 인해 실제 데이터보다 적게 산출될 수 있다는 것이다. 또한, IPC-KSIC 연계를 통해 산업과 특허를 연결했지만, 정확히 일치하지 않는 연계 구조로 인해 면밀한 분석이 어렵다는 문제가 있다. 이는 다수의 관련 연구에서도 제안하고 있는 한계점으로 우리는 관련 전문가의 정성적 판단을 통해 한계점을 보완하고자 하였다. 후속 연구에서는 다수의 전문가를 활용할 수 있는 델파이 기법 등을 활용한 복합적인 방법론을 제안한다면 분석 결과의 신뢰도를 더욱 제고할 수 있을 것이다. 그럼에도 불구하고, 국내 표준인 한국표준산업분류를 활용하여 바이오 산업의 범위를 지정하고 포괄적인 데이터를 분석하여 지역의 주요 핵심 기술 도출을 통해 과학적 정책 제언에 기여함으로써 가치가 있다고 판단된다.

참 고 문 헌

- [1] 광송비, 김건(2023), “토픽모델링과 의미연결망 분석을 통한 기계분야 특허 기술이전 동향 분석,” 기계기술정책 학술지, 2(2), 43-69.
- [2] 김동관, 박재우(2024), “텍스트 마이닝을 이용한 인천 바이오산업 연구개발 현황분석,” 산업혁신 연구, 40(2), 288-302.
- [3] 김동관, 이정영(2022), “특허정보를 활용한 인천 지역 기술혁신활동 분석에 관한 연구,” 기술경영, 7(4), 149-172.
- [4] 김명진, 강지민(2020), “지역과학기술정책 실행 방향 정립에 관한 연구: 경기도 광역자치단체 사업분석을 중심으로,” 국토지리학회지 제, 54(3), 197-212.
- [5] 김서현, 조성호, 신민수(2023), “메타버스 연구동향 및 대중인식 비교분석: 토픽모델링을 활용하여,” 한국경영과학회지, 48(3), 41-53.
- [6] 김태우, 서창교(2023), “저자동시인용분석에 의한 서비스 로봇 분야의 지적구조 분석,” 경영과학, 40(2), 15-37.
- [7] 김현우, 심위, 권오진, 노경란(2017), “특허 분석을 통한 지역맞춤형 미래유망산업 발굴 및 도출에 관한 연구: 부산 지역 사례를 중심으로,” 한국전자통신학회 논문지, 12(1), 129-138.
- [8] 남승주, 최솔샘, 김준환, 김진기(2020), “공향산업 동향분석을 위한 텍스트 애널리틱스 모델에 관한 연구,” 경영과학, 37(1), 61-74.
- [9] 박소현, 윤정원, 김민수, 구양미(2024), “지역산업 경로창출의 시공간적 특성 탐색 방법론: 인천 바이오의약산업을 사례로,” 도시연구, (25), 309-342.
- [10] 박장훈, 옥영석(2019), “특허지표를 활용한 지역 주력산업 유망기술 분석에 관한 연구: 창원시를 중심으로,” Journal of the Korea Institute of Information & Communication Engineering, 13(11).
- [11] 신서영, 이범준(2021), “코로나 19 확산에 따른 외식에 대한 소비자 인식 분석: 토픽모델링 및 네트워크 분석의 활용,” 호텔경영학연구, 30(8), 71-90.
- [12] 오준병, 허원창(2020), “인천지역의 지식재산 현황과 지역혁신 정책에 대한 시사점,” 경제연구, 38(4), 191-214.
- [13] 유광민, 김동관(2023), “특허정보를 활용한 인천 전략산업 융합기술 분석,” 혁신클러스터연구, 13(3), 90-117.
- [14] 윤호열, 김동관(2021), “2-모드 네트워크 분석을

- 활용한 지역혁신클러스터 탐색에 관한 연구: 지역혁신기관 연구개발네트워크를 중심으로,” 기술혁신학회지, 24(1), 17-40.
- [15] 윤호열, 김민호, 최상옥, 박근우(2020), “NTIS 국가연구개발사업 정보를 활용한 지역 과학기술 연구 동향 분석-충남 지역 NTIS 데이터를 중심으로,” 기술경영, 5(4), 21-44.
- [16] 이종호, 장후은(2020), “국가주도형에서 지역주도형으로: 일본의 지역과학기술정책 패러다임 전환과 그 특성 고찰,” 한국지역지리학회지, 26(4), 409-423.
- [17] 충남테크노파크(2018), “충남 지역 유망기술 도출 및 투자로드맵 수립 연구 보고서,” 충남테크노파크, 충남.
- [18] 한장협, 김채복(2018), “특허정보를 활용한 경북 지역 ICT 개발동향 분석 및 유망기술 도출에 관한 연구,” 기술혁신학회지, 21(1), 236-264.
- [19] 허동숙(2020), “EU 지역혁신정책의 동향 및 사례 연구: 스마트 전문화 전략 (RIS3) 을 중심으로,” EU 연구, 56, 293-333.
- [20] 허원창(2021), “특허 포트폴리오의 어휘 다양성과 특허성과의 관계,” 한국경영과학회지, 46(1), 1-13.
- [21] An, H. J. and Ahn, S. J.(2016), “Emerging technologies—beyond the chasm: Assessing technological forecasting and its implication for innovation management in Korea,” Technological Forecasting and Social Change, 102, 132-142.
- [22] Bengisu, M., and Nekhili, R.(2006), “Forecasting emerging technologies with the aid of science and technology databases,” Technological Forecasting and Social Change, 73(7), 835-844.
- [23] Blei, D. M.(2012), “Probabilistic topic models,” Communications of the ACM, 55(4), 77-84.
- [24] Butz, P., and Tauscher, B.(2002), “Emerging technologies: chemical aspects,” Food research international, 35(2-3), 279-284.
- [25] De Reyck, B., and Leus, R.(2008), “R&D project scheduling when activities may fail,” IIE transactions, 40(4), 367-384.
- [26] Degroote, B., and Held, P.(2018), “Analysis of the patent documentation coverage of the CPC in comparison with the IPC with a focus on Asian documentation,” World Patent Information, 54, S78-S84.
- [27] Ernst, H.(2003), “Patent information for strategic technology management,” World patent information, 25(3), 233-242.
- [28] Grimaldi, M., Cricelli, L., Di Giovanni, M., and Rogo, F.(2015), “The patent portfolio value analysis: A new framework to leverage patent information for strategic technology planning,” Technological forecasting and social change, 94, 286-302.
- [29] Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., and Zhao, L.(2019), “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey,” Multimedia tools and applications, 78, 15169-15211.
- [30] Jiang, K., Barnett, G. A., and Taylor, L. D.(2016), “Dynamics of culture frames in international news coverage: A semantic network analysis.”
- [31] Kwon, K., Jun, S., Lee, Y. J., Choi, S., and Lee, C.(2022), “Logistics technology forecasting framework using patent analysis for technology roadmap,” Sustainability, 14(9), 5430.
- [32] Leydesdorff, L., Kogler, D. F., and Yan, B.(2017), “Mapping patent classifications: portfolio and statistical analysis, and the comparison of strengths and weaknesses,” Scientometrics, 112, 1573-1591.
- [33] Lyu, F., Wang, M., and Choi, J.(2024), “Detection of Interaction-based Knowledge for Reclassification of Service Robots: Big Data Analytic Perspective,” Journal of The Korean Operations

- Research and Management Science Society, 49(1), 35-55.
- [34] Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., and Adam, S.(2021), "Applying LDA topic modeling in communication research: Toward a valid and reliable methodology," In Computational methods for communication science (pp. 13-38). Routledge.
- [35] O'callaghan, D., Greene, D., Carthy, J., and Cunningham, P.(2015), "An analysis of the coherence of descriptors in topic modeling," Expert Systems with Applications, 42(13), 5645-5657.
- [36] Saqr, M., Elmoazen, R., Tedre, M., López-Pernas, S., and Hirsto, L.(2022), "How well centrality measures capture student achievement in computer-supported collaborative learning?: A systematic review and meta-analysis," Educational Research Review, 35, 100437.
- [37] Yang, W.(2024), "Evidence-based social science: Why, what, and future implications," Humanities and Social Sciences Communications, 11(1), 1-11.
- [38] Zhou, Y., Dong, F., Liu, Y., Li, Z., Du, J., and Zhang, L.(2020), "Forecasting emerging technologies using data augmentation and deep learning," Scientometrics, 123, 1-29.