

기술테마 분석을 통한 기술기회발굴 연구*

—3D 프린팅 기술 사례를 중심으로—

서원철**

- I. 서론
- II. 관련 연구
 1. 특허분석을 통한 기술기회발굴
 2. 토픽모델링
 3. 연관규칙마이닝
- III. 기술테마 분석을 통한 기술기회발굴 방안
 1. 특정분야 관련 특허데이터 수집
 2. 기술테마 생성
 3. 잠재적 기술기회발굴
- IV. 사례 분석
 1. 3D 프린팅 기술 관련 특허데이터 수집
 2. 기술적 함의를 포괄하는 기술테마 생성
 3. 내부기술테마를 통해 접근 가능한 잠재적 기술기회발굴
- V. 논의 사항
 1. 잠재적 기술테마의 트렌드 분석
 2. 특허와 기술테마 간 할당관계 정제에 의한 영향
- VI. 결론
- VII. 부록
 1. 특허-토픽 간 관계를 활용한 apriori 알고리즘 기반 연관규칙 생성과정

* 이 논문은 부경대학교 자율창의학술연구비(2019년)에 의하여 연구되었음.

** 부경대학교 산업 및 데이터공학과 산업데이터공학융합전공 부교수.

초록

기술기회발굴은 기술 트렌드의 변화를 미래 관점에서 파악하고 이에 대응할 수 있는 기술적 아이디어의 도출을 가능하게 한다. 이에 기술기회발굴을 위한 체계적 프로세스를 정립하기 위한 연구가 활발하게 수행되어 왔다. 하지만 이들은 제품과 기술 간 고정된 형태의 시나리오만을 가정하거나, 보유 기술을 기술기회와 매칭하는 기준으로만 활용하는 등 기술기회발굴 과정에서 기술개발 주체인 기업의 적합성을 충분하게 고려하지 못하는 한계를 지니고 있다. 이에 본 연구는 기술적 함의에 대한 구체적 내용을 기술하는 기술적 속성의 집합으로 기술테마를 정의하고 기술테마 간 연관관계에 대한 분석을 통해 기술기회를 발굴해내기 위한 방안을 제시한다. 또한 발굴된 기술기회의 타당성을 확인하기 위하여 해당 기술테마 관련 기술개발 활동의 트렌드에 대한 논의를 수행하고, 본 연구의 방안에 대한 실제적 활용성을 탐색할 수 있도록 최근 제품 생산방식에 큰 변화를 유발하고 있는 3D 프린팅 분야를 대상으로 사례 분석을 시행한다. 본 연구는 국내 기업들의 취약한 R&D 기획역량을 증진시키고 지속 가능한 성장을 위한 신성장동력원의 조기 발굴 및 선점을 지원함으로써 궁극적으로 차세대 R&D 기획 시스템으로서의 역할을 수행할 수 있을 것으로 기대된다.

주제어

기술기회발굴, 특허분석, 토픽모델링, 연관규칙마이닝, 3D 프린팅

I. 서론

글로벌 기술경쟁의 격화에 따른 기술·제품수명주기의 단축 및 기업 경쟁 환경의 변동성 증대는 기업들로 하여금 새로운 기술기회를 조기에 발굴하고 이를 빠르게 사업화하는 역량을 갖추도록 요구한다.¹⁾ 기술변화에 대한 대응력을 증진하고 미래 잠재적 가치를 지닌 기술기회를 선점하는 것은 기업의 지속가능한 성장을 위해 필수적이다.²⁾ 따라서 기업의 잠재적 기술기회를 조기에 발굴할 수 있도록 지원하는 것은 해당 기업의 R&D 기획역량을 강화하고 R&D 결과물의 사업화 성과를 제고하는 데 있어 매우 중요하다.

기술기회는 기술 트렌드의 변화를 미래 관점에서 예측하는 포괄적 활동의 산출물로서 정의된다.³⁾ 정량적 데이터 분석을 기반으로 하는 기술기회발굴 연구는 최근 정보화 기술의 발전에 따라 활발하게 진행되고 있으며, 주로 기술개발을 통한 성과물인 특허정보를 중심으로 이루어지고 있다.⁴⁾⁵⁾ 특허는 기술신뢰성이 높은 발명의 결과물로서,⁶⁾ 기술혁신과 연구개발의 동향에 대한 정량적 대리변수로 널리 활용되고 있으며,⁷⁾⁸⁾⁹⁾ 이러한 특허정보의 분석

1) 박영진 외 2인, “보유특허 기반의 기술기회탐색을 위한 특허추천방법 — 3D 프린팅 산업을 중심으로”, 『지식재산연구』, 제10권 제1호(2015), 169-200면.

2) Abernathy, William & Clark, Kim, “Innovation: mapping the winds of creative destruction”, *Research Policy*, Vol.14 No.1(1985), pp.3-22.

3) Yoon, Byungun & Park, Yongtae, “A systematic approach for identifying technology opportunities: Keyword-based morphology analysis”, *Technological Forecasting and Social Change*, Vol.72 No.2(2005), pp.145-160.

4) Lee, Changyong et al., “Navigating a product landscape for technology opportunity analysis: A word2vec approach using an integrated patent-product database”, *Technovation*, Vol.96(2020), pp.102-140.

5) 정재민 외 2인, “비즈니스 기회 발굴을 위한 문제 — 해결방법 기반의 특허분석 방법”, 『지식재산연구』, 제15권 제2호(2020), 187-222면.

6) Yoon, Janghyeok & Kim, Kwangsoo, “Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks”, *Scientometrics*, Vol.88 No.1(2011), pp.213-228.

7) Ko, Namuk et al., “Analyzing interdisciplinarity of technology fusion using knowledge flows of patents”, *Expert Systems with Applications*, Vol.41 No.4(2014),

을 통해 자동화된 기술기회발굴 방법론에 대한 연구가 다양한 관점에서 수행되고 있다. 특허정보 기반의 기술기회발굴 관련 연구들은 주로 특정 기술 분야에 대한 기술 트렌드를 파악하기 위한 연구,¹⁰⁾¹¹⁾ 네트워크 분석기법을 활용하여 기술지식 파급효과를 파악하고 기회를 예측하는 연구,¹²⁾¹³⁾ 다변량 자료 분석기법을 활용하여 특허 맵을 구축하고 맵 상에서 공백영역, 아웃라이어, 군집 등에 대한 파악을 기반으로 공백기술 및 특이기술 관련 기회를 탐색하는 연구¹⁴⁾¹⁵⁾ 등 다양한 범위와 목적 아래에서 수행되어 왔다. 하지만 이들은 주로 기술분류코드나 특허 간 인용관계 등 정형화된 서지정보만을 활용하고 있어, 개별 특허가 내포하고 있는 기술적 내용을 활용하지 못한다는 한계를 지닌다. 또한 발굴된 기술기회가 기술개발 주체와 어느 정도의 적합성을 지니는지에 대해서도 고려하지 못한다. 유망한 기술 아이টে를 발굴해내더라도 기술개발 주체의 역량과 적합성이 결여된 경우, 발굴된 기회의

pp.1955-1963.

- 8) Seo, Wonchul et al., "Product opportunity identification based on internal capabilities using text mining and association rule mining", *Technological Forecasting and Social Change*, Vol.105(2016), pp.94-104.
- 9) Park, Youngjin & Yoon, Janghyeok, "Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering", *Technological Forecasting and Social Change*, Vol.118(2017), pp.170-183.
- 10) Cozzens, Susan et al., "Emerging technologies: quantitative identification and measurement", *Technology Analysis & Strategic Management*, Vol.22 No.3(2010), pp.361-376.
- 11) Trappey, Charles V. et al., "Using patent data for technology forecasting: China RFID patent analysis", *Advanced Engineering Informatics*, Vol.25 No.1(2011), pp.53-64.
- 12) Lee, Hakyeon et al., "An ANP-based technology network for identification of core technologies: A case of telecommunication technologies", *Expert Systems with Applications*, Vol.36 No.1(2009), pp.894-908.
- 13) Kim, Chulhyun et al., "Identifying core technologies based on technological cross-impacts: An association rule mining (ARM) and analytic network process (ANP) approach", *Expert Systems with Applications*, Vol.38 No.10(2011), pp.12559-12564.
- 14) Lee, Sungjoo et al., "An approach to discovering new technology opportunities: Keyword-based patent map approach", *Technovation*, Vol.29 No.6-7(2009), pp.481-497.
- 15) Yoon, Janghyeok & Kim, Kwangsoo, "Detecting signals of new technological opportunities using semantic patent analysis and outlier detection", *Scientometrics*, Vol.90. No.2(2012), pp.445-461.

적용 가능성은 매우 낮아질 수밖에 없다.¹⁶⁾ 이에 기술개발 주체인 기업과의 적합성에 대한 고려를 바탕으로 기술기회를 탐색하기 위한 연구들이 일부 진행되었다. 하지만 이들은 제품과 기술 간 고정된 형태의 시나리오만을 가정하거나,¹⁷⁾ 내부적으로 보유하고 있는 특허기술을 역량으로 일반화하고 이를 기술기회와 매칭하는 기준으로만 활용하는¹⁸⁾ 등 실질적인 보유역량을 활용한 기술기회 탐색 과정을 지원하지 못한다. 기술기회발굴 과정에 특허문서가 내포하고 있는 기술적 의미를 반영할 수 있도록 텍스트마이닝을 활용하는 연구¹⁹⁾²⁰⁾²¹⁾ 또한 수행되었지만, 이들 역시 기술그룹 단위에서 기술적 속성을 일반화한 기능정보만을 활용하고 있어 세부 기술단위의 개별적 속성을 반영하지 못하는 한계를 지니고 있다. 이러한 기술기회의 구체성 문제를 해결하기 위하여 시장에서의 활용성에 집중하는 제품기회를 탐색하기 위한 연구도 진행되었다.²²⁾²³⁾ 이들은 특허문서가 기술적 요소에 대한 정의뿐만

16) Choi, Jaewoong et al., “Technology opportunity discovery under the dynamic change of focus technology fields: Application of sequential pattern mining to patent classifications”, *Technological Forecasting and Social Change*, Vol.148(2019), pp.119737.

17) Yoon, Janghyeok et al., “Technology opportunity discovery (TOD) from existing technologies and products: A function-based TOD framework”, *Technological Forecasting and Social Change*, Vol.100(2015), pp.153-167.

18) 김창석 · 신준석, “기업 기술역량 기반 기술기회탐색에 관한 연구”, 『대한산업공학회 추계학술대회논문집』, 대한산업공학회, 2013, 1249-1258면.

19) Choi, Sungchul et al., “SAO network analysis of patents for technology trends identification: a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells”, *Scientometrics*, Vol.88 No.3(2011), pp.863-883.

20) Lee, Changyong et al., “Novelty-focused patent mapping for technology opportunity analysis”, *Technological Forecasting and Social Change*, Vol.90(2015), pp.355-365.

21) Feng, Lijie et al., “Discovering technology opportunity by keyword-based patent analysis: A hybrid approach of morphology analysis and USIT”, *Sustainability*, Vol.12 No.1(2020), pp.136.

22) Yoon, Janghyeok et al., “Identifying product opportunities using collaborative filtering-based patent analysis”, *Computers & Industrial Engineering*, Vol.107(2017), pp.376-387.

23) Jeong, Byeongki et al., “Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis”,

아니라 기술적 요소를 활용하여 구현되는 제품정보를 함께 포괄한다는 점에 초점을 맞춘다. 특허기술이 주장하고자 하는 기술적 요소를 바탕으로 실현 가능한 제품들을 발굴해낼 수 있게 되면, 유사한 기술적 요소의 범위 내에 존재하는 다양한 다른 제품들과의 연계를 설정할 수 있기 때문이다. 하지만 제품기회가 시장에서 실현되기 위해서는 해당 제품을 구현하기 위한 기술에 대한 정의가 이루어져야만 하므로 다시 기술기회에 대한 탐색을 요구하게 되는 한계를 지니게 된다.

이와 같은 기존 연구들의 한계를 극복할 수 있도록 본 연구는 기업과의 적합성을 고려하여 잠재적 기술기회를 발굴하는 방안을 제시한다. 본 연구는 보다 구체성이 높은 형태로 기술기회를 나타낼 수 있도록 기술테마 개념을 활용한다. 기술테마는 기술적 함의에 대한 구체적 내용을 기술하는 기술적 속성의 집합으로 정의되며 주로 발명의 사상을 포괄하는 특허데이터에 토픽 모델링을 적용하여 추출한 토픽으로 정의된다.²⁴⁾ 여기서 기술적 함의란 기술적 통찰을 실행할 수 있도록 하는 구체화된 기술내용의 실체를 의미한다. 이러한 기술테마에 대한 분석을 통해 본 연구는 최종적으로 기업과의 적합성이 높은 잠재적 기술기회를 발굴해낸다. 우선, 특정 분야의 특허데이터를 수집하고 토픽모델링을 적용함으로써 기술테마를 생성한다. 기술테마는 기업의 내부적 역량 및 외부적 기술기회의 가시적 표현을 위한 기본 단위로서 활용된다. 따라서 전자는 내부기술테마의 의미를 지니고 후자는 외부기술테마의 의미를 가진다. 특허-토픽 관계에 기반하여 내부기술테마와 외부기술테마 간 함의관계를 나타내는 연관규칙을 생성한다. 여기서 함의관계란 연관규칙 내 선행 항목과 관련한 역량의 확보가 결과 항목과 관련한 기술적 속성으로 나아갈 수 있는 직접적 기반이 됨을 의미한다. 내·외부기술테마 간 함의관계에 대한 탐색을 통해 최종적으로 새로운 기술기회를 발굴한다. 본

International Journal of Information Management, Vol.48(2019), pp.280-290.

24) Kim, Jusung et al., "Identifying potential technology themes based on internal capabilities using topic modeling and association rule mining", *ICIC Express Letters, Part B: Applications*, Vol.7 No.10(2016), pp.2267-2273.

연구는 국내 기업들의 취약한 R&D 기획역량을 증진시키고 지속 가능한 성장을 위한 신성장동력원의 조기 발굴 및 선점을 지원함으로써 궁극적으로 차세대 R&D 기획 시스템으로서의 역할을 수행할 수 있을 것으로 기대된다.

II. 관련 연구

1. 특허분석을 통한 기술기회발굴

기술기회발굴 프로세스를 정립하기 위한 목적으로 다양한 종류의 데이터를 활용하는 연구들이 진행되어 왔다. 이들 중, 오랜 시간동안 대량의 데이터가 축적되어 왔을 뿐만 아니라 기술에 대한 신뢰할 수 있는 기술혁신 및 연구개발 활동의 정량적 대리변수로서 인식되는 특허데이터가 기술기회발굴 프로세스 구축을 위한 목적으로 가장 널리 활용된다.²⁵⁾ 일반적으로 특허데이터 기반의 기술기회발굴 방안은 특허맵을 활용하는 접근법과 형태학(morphological) 분석 접근법으로 구분된다.²⁶⁾

특허맵은 특허문서를 벡터화한 후, 클러스터링 또는 다차원척도법 및 주성분분석 등의 차원축소 기법을 적용하여 특허들을 2차원 평면상에 매핑시키는 것을 의미한다.²⁷⁾²⁸⁾²⁹⁾³⁰⁾ 특허맵상 비어 있는 공간을 공백기술로서 정의하고 그 공백과 관련한 특허들을 탐색함으로써 해당 공백기술의 의미를 추론하거나 아웃라이어에 해당하는 특허들을 중심으로 특이기술 또는 새롭

25) Lee, Changyong et al., "Monitoring trends of technological changes based on the dynamic patent lattice: A modified formal concept analysis approach", *Technological Forecasting and Social Change*, Vol.78 No.4(2011), pp.690-702.

26) Lee, Changyong et al., *op. cit.*, pp.102-140.

27) Kim, Young Gil et al., "Visualization of patent analysis for emerging technology", *Expert systems with applications*, Vol.34 No.3(2008), pp.1804-1812.

28) Lee, Sungjoo et al., *op. cit.*, pp.481-497.

29) Yoon, Janghyeok et al., *op. cit.*, pp.445-461.

30) Park, Hyunseok et al., "Identifying patent infringement using SAO based semantic technological similarities", *Scientometrics*, Vol.90 No.2(2012), pp.515-529.

게 부각되는 기술적 속성을 추정하는 등 특허맵을 활용하여 기술기회를 발굴하기 위한 시도가 다양하게 수행되어 왔다. 또한 기술-제품 간 관계를 설정한 후 기존 기술에 대한 변화를 통해 새로운 기술 및 제품으로의 연계를 생성하거나, 기존 기술을 활용하여 적용 가능한 다른 분야의 잠재적 제품을 발굴하는 등의 기술기회발굴 경로를 개발하기 위한 연구도 진행되었다.³¹⁾

형태학 분석은 분석대상 또는 개념을 복수의 차원요소로 구분하고 각 차원에 속하는 개념 수준의 형태요소를 나열한 후 이를 재조합함으로써 새로운 개념을 파생해낸다.³²⁾ 일반적으로 하나의 시스템은 서로 다른 다양한 방법에 의해 여러 서브시스템들로 분해될 수 있고, 형태학 분석은 이들의 가능한 조합들에 대한 탐색을 체계적인 관점에서 수행할 수 있도록 한다. 하나의 큰 개념을 체계적으로 구분, 분석, 재조합하는 탐색적 접근을 통해 새로운 개념을 창출해내는 특징으로 인해 형태학 분석 또한 기술예측 및 기술기회발굴을 위한 목적으로 널리 활용되어 오고 있다.³³⁾³⁴⁾ 형태학 분석은 예상하지 못했던 획기적인 아이디어의 발견을 이끌어낼 수 있는 장점이 있는 반면, 분석의 핵심 과정이라고 할 수 있는 차원 및 요소의 식별에 있어 전문가의 개입이 상당부분 필요한 단점도 존재한다.

기술기회발굴을 위한 체계적 프로세스를 정립하는 데 있어 기존 연구들은 상당히 합리적인 성과를 창출해낼 수 있음을 입증하였다. 그럼에도 불구하고 서론에서 살펴본 것과 같이 해결해야 하는 한계 또한 여전히 많은 실정이다. 이러한 한계를 극복하고 기술개발 주체인 기업과의 적합성을 고려할 수

31) 박현석 외 4인, “기업의 보유 기술 및 제품에 기반한 기술기회발굴”, 『대한산업공학회지』, 제40권 제5호(2014), 442-450면.

32) 서용운·이학연, “이종사업간 지식의 형태학적 결합을 통한 신사업 기회 탐색”, 『대한산업공학회 추계학술대회논문집』, 대한산업공학회, 2013, 1135-1141면.

33) Yoon, Byungun & Park, Yongtae, “Development of new technology forecasting algorithm: Hybrid approach for morphology analysis and conjoint analysis of patent information”, *IEEE Transactions on Engineering Management*, Vol.54 No.3(2007), pp.588-599.

34) Yoon, Byungun et al., “Exploring technological opportunities by linking technology and products: Application of morphology analysis and text mining”, *Technological Forecasting and Social Change*, Vol.86(2014), pp.287-303.

있도록 본 연구는 기업의 내부기술테마를 기반으로 접근 가능한 잠재적 기술기회를 외부기술테마로부터 발굴해내는 방안을 제시한다. <표 1>은 기술적 내용에 기반하여 기술 및 제품기회를 발굴하는 다양한 기존 연구들의 주요내용 및 본 연구와의 차별성을 정리하여 나타낸다.

<표 1> 기술적 내용에 기반한 기술 및 제품기회발굴 관련 기존 연구 및 본 연구 비교

목적	주요 내용	본 연구와의 차별점	참조문헌
제품추천 시스템	<ul style="list-style-type: none"> 기업의 제품 포트폴리오를 바탕으로 새로운 제품 추천 특허출원인-제품 간 관계에 협업 필터링 적용을 통한 제품 추천 제품토픽에 대한 감성분석 기반 제품 추천 	<ul style="list-style-type: none"> 제품기회의 실현을 위한 구체적인 기술적 속성에 대한 정의 부재 본 연구는 기술테마를 활용하여 구체화된 기술내용 중심의 기술기회발굴 	35)36)
기능중심 기술기회 발굴	<ul style="list-style-type: none"> 기능을 정의하고 이들의 집합을 통해 기술 및 제품 정의 기능 간 유사도 측정을 통해 기술 및 제품 간 연계 직접적 연계관계를 활용한 새로운 기술 및 제품기회발굴 	<ul style="list-style-type: none"> 기능 간 유사도 수준에 따른 기술-제품 간 고정된 형태의 시나리오 가정 본 연구는 다양한 기술테마 간 방향성 있는 직접적 합의 관계에 기반하여 새로운 기술기회발굴 	37)38)
유망특허 기술발굴	<ul style="list-style-type: none"> 기능 네트워크 또는 특허맵을 활용한 유망기술발굴 신규성 중심의 특허맵을 활용한 구체적 유망기술 제시 기술 키워드 중심의 기술기회발굴 	<ul style="list-style-type: none"> 기술개발 주체의 내부적 역량에 대한 고려 없이 특정 분야의 유망기술 제시 본 연구는 기술개발 주체의 내부 역량을 통해 접근 가능한 잠재적 기술기회발굴 	39)40)41)

2. 토픽모델링

토픽모델링은 문서들의 집합으로부터 숨겨진 주제를 발견하기 위한 확률

35) Yoon, Janghyeok et al., *op. cit.*, pp.376-387.

36) Jeong, Byeongki et al., *op. cit.*, pp.280-290.

37) Yoon, Janghyeok et al., *op. cit.*, pp.153-167.

38) 박현석 외 4인, 앞의 글(주 31), 442-450면.

39) Choi, Sungchul et al., *op. cit.*, pp.863-883.

40) Lee, Changyong et al., *op. cit.*, pp.355-365.

41) Feng, Lijie et al., *op. cit.*, p.136.

모델로서 문서 내 단어의 출현빈도와 연관성을 고려하여 유사한 의미를 가진 단어들을 클러스터링하는 형태로 토픽을 추론한다.⁴²⁾⁴³⁾ 가장 대표적인 토픽모델링 기법인 LDA(Latent Dirichlet Allocation)는 디리클레 분포에 기반한 확률적 방안으로서 각 문서가 디리클레 분포를 따르는 여러 토픽들의 확률적 집합임을 가정하고 해당 문서에 포함된 단어들이 어떤 토픽과 연관되어 있는지를 결정한다.⁴⁴⁾ 그리고 문서 내 단어들이 각 토픽에 포함될 확률을 산출한 후, 이를 바탕으로 토픽을 구성하는 단어들의 집합을 추출해낸다. 각 토픽에 대한 단어들의 기여 수준은 단어들의 확률분포로서 표현되기 때문에, 추출된 토픽의 의미를 보다 명확하게 해석할 수 있는 장점을 지닌다.

숨겨진 흥미로운 주제들에 대한 탐색이 가능하다는 특징으로 인해 특허문서에 LDA를 적용함으로써 다양한 기술 트렌드를 분석하는 연구들이 활발하게 수행되어 오고 있다. 대표적으로 LDA 기반의 네트워크 분석을 통한 산업 안전분야 기술융합 트렌드 분석,⁴⁵⁾ 기업의 지식 프로파일 생성 및 다이내믹스 분석,⁴⁶⁾ 기술 트렌드 식별을 위한 특허 개발 맵 생성 방법론⁴⁷⁾ 등이 있다. LDA 실행을 통해 문서와 토픽 간 확률분포를 나타내는 문서-토픽 매트릭스와 토픽과 키워드 간 확률적 관계를 의미하는 토픽-키워드 매트릭스가 생성

42) Jelodar, Hamed et al., "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey", *Multimedia Tools and Applications*, Vol.78 No.11(2019), pp.15169-15211.

43) 박상현 외 3인, "동적토픽모델링과 네트워크 분석 기반의 기술융합 동향분석 - ICT-Agritech 융합 사례 연구", 『대한산업공학회지』, 제46권 제3호(2020), 211-221면.

44) 오승현 외 2인, "특허의 토픽 모델링을 활용한 증강현실 기술 모니터링", 『대한산업공학회지』, 제43권 제3호(2017), 213-228면.

45) Song, Bomi & Suh, Yongyoon, "Identifying convergence fields and technologies for industrial safety: LDA-based network analysis", *Technological Forecasting and Social Change*, Vol.138(2019), pp.115-126.

46) Suominen, Arho et al., "Firms' knowledge profiles: Mapping patent data with unsupervised learning", *Technological Forecasting and Social Change*, Vol.115(2017), pp.131-142.

47) Kim, Mujin et al., "Generating patent development maps for technology monitoring using semantic patent-topic analysis", *Computers & Industrial Engineering*, Vol.98 (2016), pp.289-299.

된다. 본 연구는 특허데이터에 LDA를 적용하여 기술적 함의를 포괄하는 다양한 토픽을 추출하고 이를 기술테마로 정의한다. 따라서 문서-토픽 매트릭스를 통해 기술테마와 특허문서 간 확률적 연관관계를 파악하고 이를 기반으로 기술테마를 특허들에 할당할 수 있게 된다. 이는 특허에 내포된 발명의 기술적 사상을 체계화된 기술테마의 형태로 나타낼 수 있도록 하기 때문에 기업의 내부적 역량 및 기술기회의 가시적 표현을 위한 기본 단위로서 기술테마가 활용된다. 또한, 토픽과 키워드 간 확률관계는 각 기술테마가 지니는 의미의 정성적 추론을 위한 기반으로 활용된다.

3. 연관규칙마이닝

연관규칙마이닝은 대규모 데이터베이스에 존재하는 다양한 항목들 간 잠재된 연관관계를 도출해내는 데이터마이닝 기법이다.⁴⁸⁾ 초기에는 고객들의 물건구매 행동특성을 밝히기 위한 목적으로 활용되었으며, 현재는 데이터베이스로부터 다양한 항목들 간 의미있는 관계를 도출하기 위하여 널리 활용되고 있다.⁴⁹⁾ 기본적으로 트랜잭션 내에 존재하는 다양한 항목들 간에는 복합적 연관성이 내포되어 있을 것이라는 가정을 기반으로 한다. 따라서 내포된 연관규칙을 추출함으로써 트랜잭션을 유발하는 패턴 및 트렌드에 대한 규명이 가능해진다.⁵⁰⁾ 이러한 특징으로 인해 특허데이터에 연관규칙마이닝을 적용하여 기술트렌드 및 기술적 파급효과를 분석하기 위한 연구가 활발하게 수행되었다.⁵¹⁾⁵²⁾

48) Shih, Meng-Jung et al., "Discovering competitive intelligence by mining changes in patent trends", *Expert Systems with Applications*, Vol.37 No.4(2010), pp.2882-2890.

49) Gilbert, Andrew et al., "Action recognition using mined hierarchical compound features", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.33 No.5(2010), pp.883-897.

50) 정으뜸 외 2인, "기술적 파급효과 도출을 통한 기술융합 분석 연구 — 나노바이오 융합 기술 사례를 중심으로", 『지식재산연구』, 제10권 제4호(2015), 255-290면.

51) Altuntas, Serkan et al., "Analysis of patent documents with weighted association rules", *Technological Forecasting and Social Change*, Vol.92(2015), pp.249-262.

연관규칙은 $A \rightarrow B$ 의 형태로 표현되며, 선행항목 A 와 결과항목 B 사이에 함의관계가 존재함을 의미한다. 생성된 연관규칙의 특성을 살펴보기 위하여 일반적으로 지지도(support), 신뢰도(confidence), 향상도(lift)와 같은 3가지 기준이 활용된다.⁵³⁾ 지지도는 $P(A \cap B)$ 로 산출되며, 규칙의 지지도가 높다는 것은 해당 규칙에 포함된 선행 및 결과항목들이 다른 규칙들에서도 상당히 많이 함께 포함되고 있음을 의미한다. 신뢰도는 $P(B|A)$ 로 산출되며, 규칙의 신뢰도가 높다는 것은 해당 규칙에 포함된 선행 및 결과항목 간 연관관계의 정도가 매우 강함을 의미한다. 향상도는 $\frac{P(BA)}{P(B)}$ 로 산출되며, 1보다 큰 값을 가지면 양의 연관성, 1보다 작은 값을 가지면 음의 연관성이 있다고 판단하고 1의 값을 가질 경우에는 선행 및 결과항목의 출현이 독립적이라고 판단한다. 의미 있는 연관규칙의 생성을 위해서는 충분한 지지도 및 신뢰도 값을 가질 수 있도록 최소값을 사전에 설정하고 해당 값 이상의 규칙만이 생성될 수 있도록 해야 한다. 가장 대표적인 연관규칙 생성 알고리즘인 Apriori는 모든 데이터셋에서 최소 지지도 이상을 가지는 빈발 항목집합들을 찾고 이를 활용하여 연관규칙을 생성한다.⁵⁴⁾ 기본 아이디어는 특정 항목집합이 빈발하지 않으면 해당 항목집합을 포함하는 모든 항목집합은 빈발할 수 없다는 것이다. 이는 연관규칙 생성과정의 효율성을 증진시키는 요인이 된다.

III. 기술테마 분석을 통한 기술기회발굴 방안

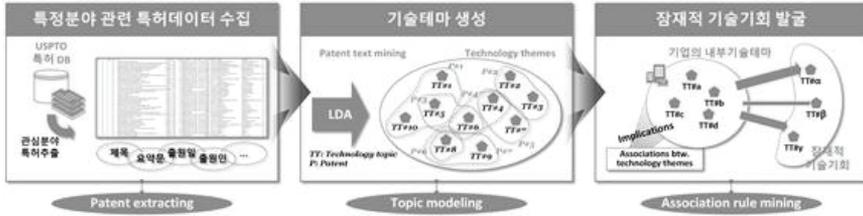
기업의 내부기술테마를 기반으로 접근 가능한 잠재적 기술기회를 외부기

52) Choi, Sungchul et al., "Analyzing technological spillover effects between technology classes: The case of Korea technology finance corporation", *IEEE Access*, Vol.6(2018), pp.3573-3584.

53) Agrawal, Rakesh et al., "Mining association rules between sets of items in large databases", In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Association for Computing Machinery*, 1993, pp.207-216.

54) Agrawal, Rakesh et al., *op. cit.*, pp.207-216.

〈그림 1〉 기술테마 분석을 통한 기술기회발굴 연구 프레임워크



술테마로부터 발굴하기 위한 본 연구는 〈그림 1〉과 같이 1) 특정 기술분야와 관련된 특허데이터 수집, 2) 특허데이터에 LDA를 적용하여 역량 및 기술기회의 표현을 위한 기본 단위로서의 기술테마 생성, 3) 연관규칙마이닝을 적용하여 기술테마 간 연관규칙을 생성하고 연관규칙에 기반하여 기업과의 적합성을 지닌 잠재적 기술기회발굴의 3단계로 이루어진다. 추가적으로 본 연구를 통해 발굴된 기술기회의 타당성을 살펴보기 위한 논의를 수행한다.

1. 특정분야 관련 특허데이터 수집

기술테마 분석을 통해 기술기회를 발굴하기 위한 본 연구는 특허분석을 기반으로 하기 때문에 첫 번째 단계에서는 특허데이터의 수집이 이루어진다. 발굴된 기술기회가 보다 명확한 의미를 지닐 수 있도록 하기 위해서는 특정 기간의 모든 특허를 활용하기보다 특정 기술분야와 관련된 특허만을 활용하는 것이 합리적이다. 특허데이터는 해당 기술분야를 포괄할 수 있는 검색식에 기반하여 수집된다. 따라서 기술기회발굴을 위한 특정 기술분야를 선정하고 해당 분야의 여러 키워드를 활용하여 특허 추출을 위한 검색식을 작성한다. 미국은 글로벌 기업들의 대표적 경쟁장소로서 미국에서의 안정적인 비즈니스 활동 및 시장 선점을 위하여 많은 기업들은 USPTO(United States Patent and Trademark Office)에 특허를 출원하기 위한 활동을 활발하게 수행한다. 특히 유럽, 일본, 중국, 한국에 비해 USPTO에 출원되는 특허는 자국민보다 해외 출원인의 비중이 상당히 높은 특징이 있다. 이는 다양한 관점의 기술기회를 발굴해낼 수 있는 가장 적합한 특허 데이터베이스가 USPTO 데

이터베이스임을 의미한다.

기업의 내부기술테마를 정의하고 새로운 기술기회를 표현하기 위해서는 단순한 특허 서지정보가 아닌 텍스트 정보가 반드시 요구된다. 따라서 본 단계에서는 USPTO에 출원 및 등록된 특정 기술분야 관련 특허들에 대해 텍스트 정보를 중심으로 수집한다.

2. 기술테마 생성

본 단계에서는 수집된 특허데이터에 LDA를 적용하여 기술적 속성을 지닌 토픽을 추출하고 이를 기술테마로 정의한다. 이를 위하여 우선 수집한 특허의 텍스트 정보를 활용하여 코퍼스(corpus)를 구축한다. 특허문서 내 다양한 텍스트 요소 중, 본 연구에서는 제목과 요약문을 활용한다. 발명의 내용 및 발명 실시를 위한 구체적 내용은 너무 방대한 양의 텍스트를 지니고 있어 기술적 특성에 대한 효과적 추출이 어렵고, 배경기술 및 청구항 등은 기술관련 핵심 내용이 아닌 주변 내용들을 다수 포함하고 있는 반면, 제목과 요약문은 기술적 속성에 대한 핵심 내용이 주를 이루고 있어 기술적 속성을 포괄하는 토픽 추출에 있어 가장 적합하기 때문이다.⁵⁵⁾ 불용어(stopword) 리스트 구성을 통한 의미 없는 단어 제거, 어간 추출(stemming)을 통한 중복 의미를 지닌 단어 통합 등의 전처리 과정을 거쳐 구축된 코퍼스를 정제한다. 그리고 정제된 코퍼스를 활용하여 문서-단어 매트릭스를 생성한다. 기본적으로 문서-단어 매트릭스는 매우 큰 차원을 지닐 수밖에 없는데 이는 출현 빈도수가 매우 작은 단어들까지도 포함하여 매트릭스가 구성되기 때문이다. 이렇게 미미한 관계까지 포괄하여 토픽을 추출하는 경우, 기술적 의미가 모호한 토픽들이 생성될 우려가 있다. 따라서 빈도수가 충분한 단어들을 사전에 지정하고 이들만이 포함되도록 문서-단어 매트릭스를 정제한 후, LDA를 적용하여 토픽들을 추출함으로써 충분한 수준의 기술적 함의가 포함될 수 있도록 한다.

55) 최성철·서원철, “기술평가를 위한 평가 참조정보 생성 방안에 관한 연구 - 기술적 속성 유사성 관점에서”, 『지식재산연구』, 제14권 제2호(2019), 193-226면.

LDA 적용을 위해서는 추출할 토픽의 개수를 사전에 결정해야 한다. 토픽의 개수는 LDA 적용을 통해 생성된 토픽모델의 품질에 영향을 미친다. Perplexity는 이러한 품질의 측정을 위하여 일반적으로 활용되는 기준들 중 하나로서 추출할 토픽의 개수에 대한 가시적 가이드를 제공한다.⁵⁶⁾ Perplexity는 다음 식을 통해 측정된다.

$$perplexity = \exp \left\{ - \frac{\sum_{d=1}^M \log p(w_d)}{\sum_{d=1}^M N_d} \right\} \quad (1)$$

여기서, M 은 전체 문서의 수, N_d 는 하나의 문서 d 에 포함된 단어들의 수, $p(w_d)$ 는 문서 d 에 포함된 단어들의 확률을 의미하며 이는 다음 식으로서 정의된다.⁵⁷⁾

$$p(w_d) = \sum_{t=1}^T p(t_d) p(w_t) \quad (2)$$

여기서, T 는 전체 토픽의 수, $p(t_d)$ 는 문서 d 에 포함된 토픽들의 확률, $p(w_t)$ 는 토픽 t 에 포함된 단어들의 확률을 의미한다. 전체 문서 내에 포함된 단어들의 총 개수와 단어들이 포함되는 확률의 로그 합에 대한 비율을 이용하는 것으로서, $p(w_d)$ 가 큰 값을 가질수록 즉, perplexity가 작은 값을 가질수록 더 나은 품질수준을 의미하기 때문에 본 연구에서는 토픽의 개수를 달리 하면서 perplexity를 측정하고 이 값의 변화정도를 관찰하여 변화 정도가 약해지는 지점을 찾아 해당 지점에 대응되는 토픽의 개수를 최종적으로 추출할 토픽의 개수로 설정한다.

56) Han, Xiaotong et al., "R&D trend analysis based on patent mining: An integrated use of patent applications and invalidation data", *Technological Forecasting and Social Change*, Vol.167(2021), pp.120691.

57) Blei, David & Lafferty, John, "Correlated topic models", *Advances in Neural Information Processing Systems*, Vol.18(2006), pp.147-154.

특허데이터에 대한 LDA의 적용은 특허문서와 토픽 간 확률적 연관관계를 나타내는 문서-토픽 매트릭스를 생성한다. 본 연구는 기술개발 주체인 기업의 내부 역량을 바탕으로 접근 가능한 기술기회를 발굴하고자 하며, 이를 위해서는 토픽을 이용하여 생성되는 기술테마를 기업과 연계시키는 것이 필요하다. 기업이 보유하고 있는 기술적 역량은 해당 기업의 특허 포트폴리오로부터 추정할 수 있기 때문에 특허-기술테마 간 할당관계를 정의하게 되면 최종적으로 기업과 기술테마의 연계를 이루어낼 수 있다. 기술테마는 기본적으로 토픽에 의해 정의되므로 특허문서와 토픽 간 확률적 연관관계를 나타내는 문서-토픽 매트릭스를 활용함으로써 특허를 기술테마에 할당할 수 있다. 하지만 특허와 토픽 간 확률 값 역시 상당히 편차가 클 수밖에 없기 때문에 이 매트릭스에 존재하는 모든 관계를 기반으로 특허에 기술테마를 할당하게 되면 너무 미미한 관계까지도 포괄해버리는 문제가 있다. 즉, 하나의 기업에 거의 모든 기술테마가 내부 역량으로서 포함되게 되고 해당 기업과 무관하여 새롭다고 여겨질 수 있는 기술테마로의 접근을 제안하는 연관규칙이 거의 생성되지 못하는 문제가 발생할 수 있는 것이다. 이에 본 연구에서는 적절한 기준값을 설정하고 해당 값 이상의 확률관계를 지니는 관계만을 기반으로 특허와 기술테마 간 할당관계를 정의한다. 이러한 할당관계는 특허에 내포된 발명의 기술적 사상을 체계화된 기술테마의 형태로서 나타낼 수 있도록 한다.

3. 잠재적 기술기회발굴

본 단계에서는 연관규칙마이닝을 활용하여 기술테마 간 연관규칙을 생성한다. $A \rightarrow B$ 의 형태로 표현되는 연관규칙은 선행항목 A 를 포함하는 트랜잭션이 결과항목 B 를 포괄할 수 있을 것이라는 함의관계를 의미한다. 따라서 선행 및 결과항목이 모두 기술테마로 이루어져 있는 연관규칙의 생성은 기술적 속성에 대한 구체적 의미를 지니고 있는 기술테마 간 직접적 영향관계를 도출해낼 수 있도록 한다. 연관규칙에서 선행항목은 기술지식의 제공을

통해 새로운 기술지식의 직접적 창출을 이끌어내는 역할을 수행하고, 결과 항목은 외부의 기술지식에 대한 적극적 활용을 통해 실행 가능한 잠재적 기술기회로서의 역할을 수행한다. 본 연구는 기술테마 간 연관관계를 생성한 후, 선행항목의 내부기술테마를 역량으로 지니는 기업들에 대해 해당 역량을 기반으로 접근 가능한 잠재적 기술기회를 결과항목의 외부기술테마에 기반하여 발굴한다. 이는 각 기업의 역량에 적합한 기술기회의 발굴을 가능하게 한다.

의미 있는 연관규칙의 생성을 위하여 3가지 기준인 지지도, 신뢰도 및 향상도를 활용한다. 우선 생성될 연관규칙의 유용성과 확신성에 대한 의미를 지니는 지지도 및 신뢰도가 충분한 규칙들이 생성될 수 있도록 최소 지지도 및 최소 신뢰도 값을 사전에 설정하고 이들보다 큰 경우에만 의미있는 연관규칙으로 판단한다. 따라서 본 연구의 실제적 활용을 위해서는 생성될 연관규칙의 지지도 및 신뢰도에 대한 임계치로서의 최소값 기준을 사전에 수립하는 것이 필요하다. 연관규칙을 구성하는 선행 및 결과항목 간 양의 상관성을 보장할 수 있도록 1보다 큰 향상도 값을 가지는 규칙만을 의미있는 연관규칙으로 판단한다. 향상도가 1이면 선행 및 결과항목 간 상관성이 존재하지 않음을, 1보다 작으면 음의 상관성이 존재함을 나타낸다. 특허와 토픽 간 관계를 바탕으로 기술테마 간 연관규칙을 생성하기 위해서는 특허문서를 트랜잭션으로, 각 특허문서와 밀접한 관계를 지니고 있는 기술테마를 항목집합으로 가정하는 것이 필요하다. 이러한 가정하에서 apriori 알고리즘을 적용하여 기술테마 간 연관규칙을 생성하는 구체적인 과정은 부록 1에 제시되어 있다.

기술테마 간 연관관계를 나타내는 연관규칙은 선행 및 결과항목으로 표현되는 내·외부기술테마로 구성된다. 연관규칙은 선행 내부기술테마에 대한 역량을 바탕으로 결과 외부기술테마로의 접근에 대한 잠재적 가능성이 있음을 함축한다. 하나의 기술테마는 여러 특허들에 할당되어 있고, 기업은 자신만의 특허 포트폴리오를 지니고 있기 때문에 기업과 기술테마 간 관계 또한 유추할 수 있다. 만약 한 연관규칙의 선행 기술테마가 특정 기업의 특허 포트폴리오에 포함되어 있는 특허에 할당되어 있다면, 해당 기술테마가 이 기업의 역량에 포함된 기술적 속성을 나타낸다고 볼 수 있게 된다. 따라서 기

업은 생성된 연관규칙에 대한 탐색을 통해 내부기술테마와 새로운 기술기회로서의 외부기술테마 간 관계를 파악할 수 있고 이를 활용하여 접근 가능한 잠재적 기술기회에 대한 기술적 아이디어를 발굴해낼 수 있다.

IV. 사례 분석

1. 3D 프린팅 기술 관련 특허데이터 수집

본 연구는 기본적으로 특정 기술분야와 관련된 최근의 특허데이터를 필요로 한다. 본 연구의 기술기회발굴 과정을 도식화하기 위하여 3D 프린팅을 사례 분석 기술분야로서 선정하였다. 3D 프린팅은 3차원 설계데이터를 기반으로 재료를 쌓아올리는 방식으로 3차원 물체를 형상화하는 기술로서 제품 생산을 위해 일반적으로 요구되는 별도의 금형을 필요로 하지 않고, 생산 가능한 제품의 범위가 넓다는 장점으로 인해 오랜 기간동안 이루어져 왔던 전통적 제품 생산방식에 큰 변화를 유발하고 있다.⁵⁸⁾ 특히 폴리머, 금속, 종이 등 다양한 종류의 재료를 사용할 수 있을 뿐만 아니라 적용 가능한 산업 범위 또한 매우 광범위하여 관련 기술개발이 매우 활발하게 이루어지고 있다.⁵⁹⁾ 따라서 이러한 3D 프린팅 기술분야를 대상으로 사례 분석을 시행하여 본 연구에서 제시하는 기술기회발굴 방안에 대한 실제적 활용성을 탐색하고자 한다.

3D 프린팅 관련 특허데이터를 수집하기 위해서는 적절한 검색식의 작성이 요구된다. 검색식의 작성이 적절하지 못할 경우 관련성이 떨어지는 데이터가 함께 포함되게 되고 이를 제거하는 데 많은 시간과 노력이 소요될 수밖에 없다. 따라서 보다 정교한 특허 검색식 작성을 위하여 본 연구는 EPO (European Patent Office)의 3D 프린팅 관련 보고서⁶⁰⁾에 제시된 검색식⁶¹⁾을

58) 지선구·한덕원, “3D 프린팅과 특허발명의 법적 보호”, 『지식재산연구』, 제11권 제1호(2016), 37-70면.

59) 김용기, 『3D 프린팅 시장 및 기술 동향』, 과학기술일자리진흥원, 2019, 1-3면.

이용하고자 한다. 이 보고서는 EPO에 등록된 특허데이터를 기반으로 3D 프린팅 기술 관련 최근의 트렌드 및 미래 방향에 대한 분석을 실행한다. 해당 보고서의 특허 검색식에 기반하여 최종적으로 2010년부터 2019년까지 USPTO에 출원된 3D 프린팅 기술 관련 특허 8,936건을 수집하였다.

2. 기술적 함의를 포괄하는 기술테마 생성

본 단계에서는 수집된 특허의 제목과 요약문에 LDA를 적용하여 3D 프린

60) Ceulemans, Judy et al., *Patents and additive manufacturing: Trends in 3D printing technologies*, European Patent Office, 2020, pp.1-81.

61) 해당 보고서에 제시된 검색식을 Wips On의 형식으로 변경하여 3D 프린팅 관련 특허 데이터를 수집하였고 실제 활용된 검색식은 다음과 같다.

((Rapid NEAR prototype*)) (((additive* "layer wise*" "free form?") NEAR (manufacturing manufactured fabricat*)) NOT (additive? AND 2d AND "manufactured by")) NOT (additive? AND 2d AND manufacturer?) (("3D" "three dimension*" "3 dimension*" "three D") AND 2d AND print*) NOT (fig* NEAR "3D") (3OG AND fused AND deposition* AND model*) (3UG AND fused AND filament* AND (deposit* print*)) (((((((((((((((laser AND 2d AND (manufacturing manufactured fabricat*)) NOT (laser AND 2d AND microscope*)) NOT (laser AND 2d AND corporation*)) NOT (laser AND 2d AND ablati*)) NOT (laser AND 2d AND machin*)) NOT (laser AND 2d AND cutting)) NOT (laser AND 2d AND manufactured_by)) NOT (laser AND 2d AND manufacturer?)) NOT (Laser AND 2d AND weld*)) NOT (laser AND 2d AND cut*)) NOT (laser* AND 2d AND mark*)) NOT (laser AND 2d AND trimm*)) NOT (laser AND 2d AND chips)) NOT (laser AND 2d AND diffrac*)) NOT (laser AND 2d AND scan*) ("electron beam" AND 2d AND (melt* fus*)) ("Stereo lithograph" OR "mi?ro stereolith") ("Free form fabri") (3OG AND direct AND digital* AND manufact*) (3OG AND additive* AND layer* AND manufact*) ("vat polymeris" OR "vat photo polymer") ("drop on demand") (3OG AND laminat* AND object* AND manufact*) (desktop NEAR manufact*) (4OG AND laser AND engineered AND net AND shap*) Robocasting (binder NEAR jet+) (4OG AND powder AND bed AND (fusi* melt* sinter*)) (3OG AND "plaster base?" AND print*) (3OG AND laser AND metal AND form*) (3OG AND direct AND ink AND writ*) (3OG AND direct AND light AND process*) (3OG AND two AND photon AND (lithograph* polymeri*)) (4OG AND continu* AND light AND interface AND produc*) (3OG AND direct AND energy AND deposit*) "Multi material" NEAR jet* (3OG AND shaped AND metal AND deposit*) (4OG AND direct AND metal AND deposit*) (selectiv* NEAR sinter*)) AND (@AD)=20100101(<=20191231)

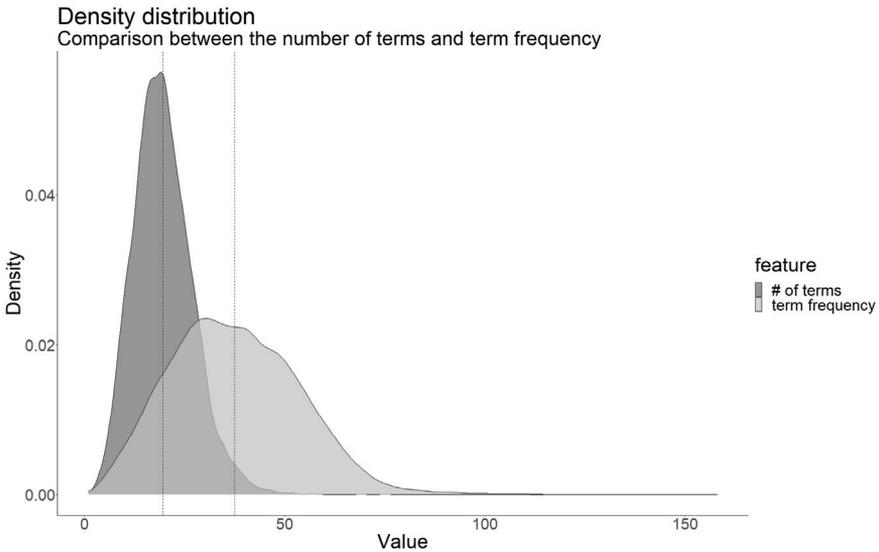
팅 관련 기술테마를 생성한다. 특히 텍스트 정보를 활용하여 코퍼스를 구축하고 불용어 처리, 중복 의미 단어 통합 등의 전처리 과정을 거쳐 정제한 후, 이를 활용하여 총 10,188개의 단어가 포함된 문서-단어 매트릭스를 생성하였다. 하지만 이는 매우 희소(sparse)한 매트릭스이기 때문에 기술적 의미가 명확한 토픽을 생성하기 위해서는 차원을 축소할 필요가 있다. 이를 위하여 단어의 최소 출현빈도를 결정하고 이보다 큰 빈도를 지닌 단어들만을 남기도록 한다. 출현빈도에 따라 단어들을 정렬하고 출현빈도의 누적 비율을 산출한 결과, 80% 비율을 나타내는 단어 수는 760여개, 90% 비율을 나타내는 단어 수는 1,500여 개로 나타났다. 따라서 본 연구에서는 단어의 최소 출현빈도 기준을 85%로 하여 1,000여 개의 단어들만을 선택하고 최종적으로 이들만이 포함될 수 있도록 문서-단어 매트릭스에 대한 정제를 실행하였다. 물론 전체 단어의 출현빈도 기준 85% 수준의 관계를 남긴다 하더라도 15%의 출현빈도가 정제됨으로 인하여 문서-단어 간 관계가 약화될 우려가 있다. 따라서 <표 2> 및 <그림 2>와 같이 정제 후 문서별 단어 수 및 출현빈도 관점의 분포와 기술통계량을 살펴보도록 한다. 단어 수 및 출현빈도 모두 최소값이 0이 아닌 결과를 나타내는 것으로 보아 하나의 단어도 포함하지 않는 문서는 발생하지 않았고, 분포 역시 각각의 평균을 중심으로 대칭적인 모습을 보이고 있어 매우 소수의 단어들만을 포함하는 문서도 흔하게 나타나지 않았음을 확인할 수 있다. 그러므로 출현빈도 기준 85% 수준으로 문서-단어 매트릭스에 대한 정제를 실행하더라도 문서-단어 간 관계가 불명확해지는 문제는 발생하지 않는다고 판단할 수 있다.

추출할 토픽의 개수를 결정하기 위해 perplexity를 측정한다. 토픽의 개수를 100개까지 1개씩 늘려가면서 perplexity 값을 측정한 결과 최대인 100개 일 때 최소값을 보이는 것으로 나타났다. Perplexity는 토픽의 개수가 늘어남에 따라 줄어드는 경향을 나타내는 것이 일반적이다. 따라서 최소값만을 기준으로 추출 토픽의 개수를 100개로 결정하는 것은 의미가 모호하거나 다른 토픽과의 중복성이 높은 토픽을 추출하게 될 문제가 있다. 본 연구에서는 단순히 최소 perplexity 값만을 이용하지 않고, 토픽의 개수를 1개씩 늘려감에

〈표 2〉 문서-단어 매트릭스 정제 후, 문서별 단어 수 및 출현빈도 기술통계량

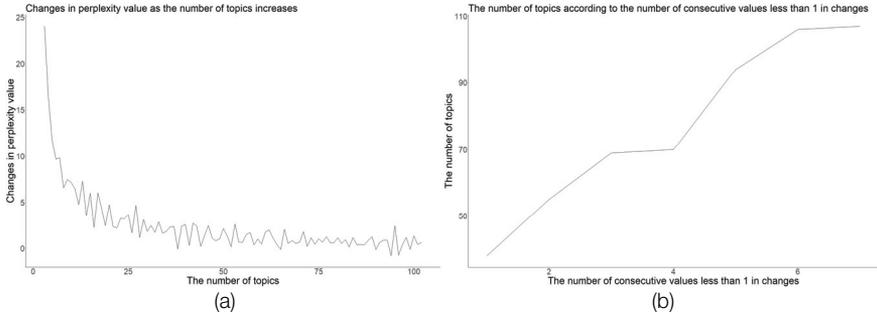
구분	최소값	Q1	Q2	평균	Q3	최대값	표준편차
단어 수	1	14	19	19.5875	24	75	7.3586
단어 출현빈도	1	26	36	37.4690	48	158	16.0620

〈그림 2〉 문서-단어 매트릭스 정제 후, 문서별 단어 수 및 출현빈도 분포



따라 얻을 수 있는 perplexity 값의 감소량을 살펴본다. 〈그림 3(a)〉는 perplexity 값의 감소량을 도식화하여 나타낸 것으로 대략 20개까지는 큰 폭으로 감소하다가 이후로는 5 이하의 범위 내에서 변동성을 보인다. 20개의 토픽은 기술기회발굴 분석을 하기에 매우 부족하기 때문에 안정적 수준의 감소량 변화가 나타나는 부분을 기준으로 활용할 수 있도록 연속해서 감소량이 1이하의 값을 나타내는 경우에 따른 토픽 개수를 측정한다. 〈그림 3(b)〉는 이를 나타낸 것으로 처음으로 1이하의 값을 보일 때의 토픽 개수는 38개, 처음으로 2번 연속 1이하의 값을 보일 때의 토픽 개수는 55개이다. 처음으로 3번, 4번, 5번 연속 1이하의 값을 보일 때의 토픽 개수는 각각 69, 70, 94개로

〈그림 3〉 (a) 토픽 개수의 증가에 따른 perplexity 값 감소량, (b) 연속해서 10이하의 감소량을 나타내는 개수에 따른 토픽 개수



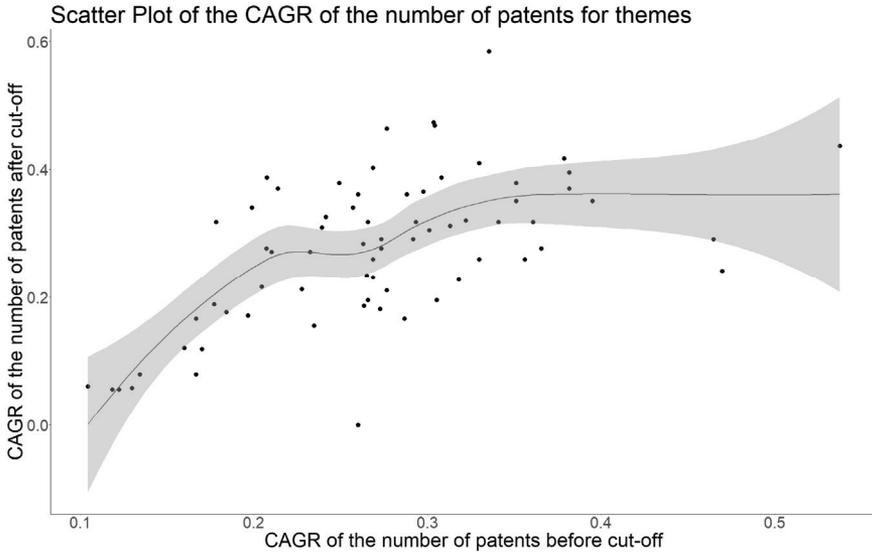
나타난다. 이러한 토픽 개수의 변화량을 살펴볼 때 가장 큰 폭의 변화를 나타내는 부분이 4번과 5번 사이이다. 따라서 4번일 때의 토픽 개수인 70개 이후로는 어느 정도 토픽 개수 변화에 따른 perplexity 값의 변화가 안정화된다고 판단할 수 있다. 이를 바탕으로 본 연구에서는 최종적으로 추출할 토픽의 개수를 70개로 설정한다.

LDA의 적용을 통해 8,936개의 특허문서와 기술적 속성을 지닌 토픽으로 정의된 70개의 기술테마 간 확률적 연관관계가 생성된다. 각 기술테마에 명확한 기술적 의미를 부여하기 위해서는 이러한 확률적 연관관계를 가중치로 활용하여 각 기술테마를 특허에 할당하는 과정이 요구된다. 하지만 가중치 역시 상당히 큰 편차를 보일 수밖에 없기 때문에 이들 모두를 활용하는 경우 매우 미미한 관계까지도 포괄해버리는 문제가 있다. 이에 각 기술테마의 특허할당을 위한 적절한 기준의 마련이 필요하다. 이러한 기준 수립에 있어 기본적인 원칙은 기술테마의 기술적 의미가 명확하게 나타날 수 있도록 하고, 기술테마 간 중복성을 최대한 줄일 수 있도록 한 특허에 할당되는 기술테마의 수가 1개를 크게 초과하지 않도록 하는 것이다. 본 연구에서는 특허문서로부터 토픽을 추출하고 이를 이용하여 기술테마를 정의한다. 하나의 특허문서는 기본적으로 하나의 기술적 속성을 포괄하기 때문에 그러한 속성을 기술테마의 형태로서 활용할 수 있도록 한 특허에 할당되는 기술테마의 수가 1개 수준으로 유지되도록 한다. 이는 특허 및 해당 특허를 보유한 기업과 기

술테마 간 관계를 보다 명확하게 하는 장점을 지닌다. 특허와 기술테마 간 확률적 연관관계 수치의 Q2, 평균, Q3을 산출하고 이들 각각을 기준으로 컷오프(cut-off)한 후, 남은 관계에 기반하여 기술테마를 할당한 결과 특허 1개당 평균적으로 할당된 기술테마의 수는 각각 1.49개, 1.16개, 0.75개로 나타났다. 따라서 본 연구에서는 최종적으로 1과 가장 가까운 결과를 보인 평균을 컷오프 기준으로 활용한다. 해당 기준에 기반하여 기술테마를 할당한 결과, 단 1개의 기술테마만이 할당된 특허가 약 79%를 차지한 반면, 3개 이상의 기술테마가 할당된 특허는 1.7%에 불과한 것으로 나타났다. 물론 이러한 관계를 이용하여 연관규칙을 생성하게 되면, 79%에 해당하는 특허-기술테마 간 관계는 유용한 정보를 제공해 주지 못하는 한계가 있다. 그럼에도 불구하고 본 연구에서는 특허와 기술테마 간 불명확한 관계에 의해 무의미한 연관규칙이 생성되는 것을 방지하기 위하여 약 21%에 해당하는 보다 명확한 특허-기술테마 간 관계만을 활용하여 연관규칙을 생성하고자 한다.

컷오프를 통해 일부 데이터만을 활용하는 본 연구의 접근이 편향되지 않고 전체 데이터를 대변할 수 있는지에 대한 탐색을 수행할 필요가 있다. 이를 살펴볼 수 있는 적절한 방법 중 하나는 기술테마와 할당관계를 맺고 있는 특허의 수가 컷오프 전 대비 후에 어떻게 변화하였는지를 살펴보는 것이다. 기본적으로 본 연구의 사례 분석 기술분야인 3D 프린팅 관련 특허는 2010년부터 2019년에 월등히 많은 양의 출원이 이루어졌기 때문에 70개의 기술테마와 할당관계를 맺고 있는 특허 또한 시간의 흐름에 따라 증가하는 양상을 보일 것이다. 이러한 증가 양상이 컷오프 이후에도 변함없이 일관되게 나타난다면 컷오프를 기반으로 하는 본 연구의 접근이 편향되지 않음을 보여준다고 할 수 있다. <그림 4>는 기술테마와 할당관계를 맺고 있는 특허의 수에 대한 2010년부터 2019년까지의 연평균 증가율을 컷오프 이전과 이후로 나누어 산점도로 나타낸 것이다. 전반적으로 양의 상관관계를 보이는 것으로 확인된다. 피어슨 상관계수 값은 0.5805이며, p-value는 0.001 이하로 유의수준 0.1%에서도 유의한 상관성을 나타낸다. 물론 0.5805라는 값이 절대적으로 아주 강한 양의 선형관계를 나타내는 것은 아니지만, 컷오프 전과 후의

〈그림 4〉 컷오프 전 및 후의 기술테마별 할당 특허 수 변화 양상 비교



특허 수에 대한 연평균 증가율을 비교하는 것이기 때문에 컷오프 전의 변화하는 양상을 컷오프 이후에도 충분히 편향되지 않게 반영하고 있음을 나타낸다고 할 수 있다.

3. 내부기술테마를 통해 접근 가능한 잠재적 기술기회발굴

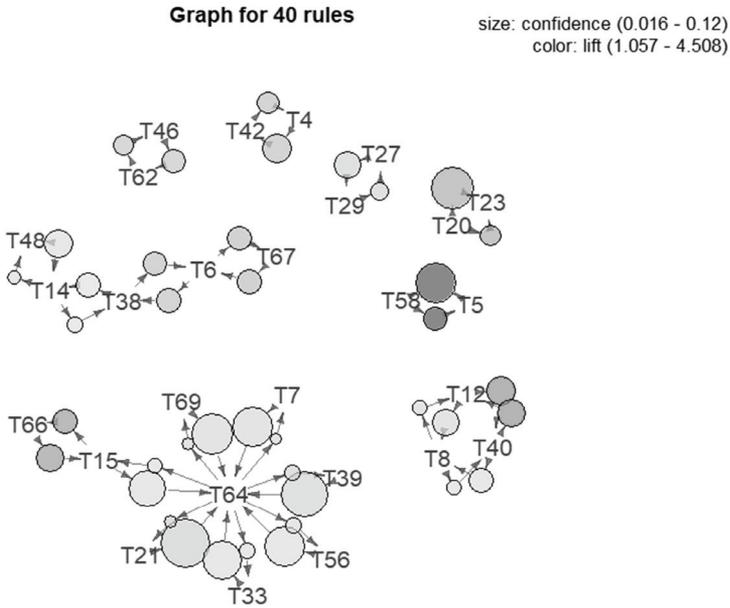
Apriori 알고리즘을 이용하여 기술테마 간 연관규칙을 생성한다. Apriori 알고리즘의 적용을 위해서는 지지도 및 신뢰도에 대한 최소값을 사전에 지정하는 것이 필요하다. 지정된 최소 지지도 값 이상을 가지는 항목집합들을 빈발항목집합으로 정의하고 이를 결합 또는 분해하는 형태로서 연관규칙을 생성하기 때문이다. 최소 지지도 값을 너무 높게 설정하면, 기술테마 간의 의미있는 영향관계를 밝히기에 부족한 수의 규칙이 생성될 것이고, 반대로 너무 낮게 설정하면, 미미한 수준의 영향관계를 지닌 규칙들이 다수 생성될 것이다. 따라서 기술테마 간 영향관계를 바탕으로 기업과의 적합성을 고려한

잠재적 기술기회를 발굴해내기 위해 충분한 수준의 규칙들을 생성하면서도 미미한 관계를 지닌 규칙들을 배제할 수 있도록 최소 지지도 값 기준으로 0.1%가 활용되었다.

신뢰도 값은 연관규칙을 구성하는 선행 및 결과 기술테마 간 직접적인 함의관계의 수준을 의미하며, 최소 신뢰도 값에 대한 변화는 미미한 함의관계를 그대로 활용할 것인지 아니면 제거할 것인지에 대한 선택의 문제가 된다. 본 연구에서는 생성된 기술테마 간 연관규칙의 신뢰도 값을 이용하여 잠재적 기술기회의 수준을 정량화하기 때문에 미미한 직접적 함의관계를 지닌 연관규칙의 존재 유무 자체는 결과에 큰 영향을 미치지 않게 된다. 따라서 최소 신뢰도 값 기준은 최소 지지도 값과 동일한 수준으로 설정하였다. 특허와 기술테마 간 관계에 기반하여 연관규칙이 생성되며, 대다수의 특허에 1개의 기술테마만이 할당되어 있기 때문에 지지도 및 신뢰도 값의 절대적 수준은 낮을 수밖에 없다. 따라서 최소 지지도 및 최소 신뢰도 값 또한 절대적 기준으로 매우 낮은 값으로 설정하였다. 향상도는 연관규칙에 존재하는 기술테마 간 상관성을 나타내며, 오직 양의 상관성을 지닌 규칙만이 생성될 수 있도록 향상도 값은 1보다 커야한다는 제약조건 또한 추가로 설정하였다. 이러한 제약조건하에서 특허와 기술테마 간 관계에 apriori 알고리즘을 적용하여 최종적으로 40개의 연관규칙을 생성하였다. <그림 5>는 생성된 40개의 기술테마 간 연관규칙을 도식화하여 보여준다. 원의 크기는 신뢰도를 나타내고, 색의 진함 정도는 향상도를 의미한다.

본 연구는 기업과의 적합성을 고려하여 잠재적 기술기회를 발굴하는 방안을 제시하고자 한다. 본 방안의 실행 및 적용 과정을 살펴보기 위해서는 하나의 사례기업을 선정하고 해당 기업의 관점에서 실행 가능한 잠재적 기술기회를 발굴해보는 것이 필요하다. HP, Xerox 등 전통적 프린터 기업들은 포화상태인 전통 프린터 시장을 벗어나 신규 비즈니스 창출을 위한 기회로서 3D 프린팅을 적극 활용하고 있다. 2010년대 초중반부터 3D 프린팅 관련 특허를 확보하기 위한 노력을 활발하게 기울여 왔다. HP는 이미 자체기술로 3D 프린터 제품을 개발 및 판매하고 있으며, 컬러 3D 프린터 및 금속 3D 프

〈그림 5〉 생성된 기술테마 간 연관규칙



린터 등 새로운 제품 개발을 통해 시장 점유율을 확대해나가고 있다. Xerox 역시 3D 프린팅 분야 진출을 위한 로드맵을 공개하고 액상 메탈젯 3D 프린터 제조회사인 Vader Systems를 합병하는 등 시장 진출을 위한 활동을 적극적으로 시행하고 있다.⁶²⁾ 본 연구는 기존의 역량을 바탕으로 새롭게 접근 가능한 기술기회를 탐색해내고자 하며, Xerox는 3D 프린팅 시장에서는 신생 기업의 위치이지만 관련 특허의 확보에 상당한 노력을 기울여 왔기 때문에 새로운 기술기회의 창출을 위한 충분한 역량을 보유하고 있는 기업이다. 따라서 Xerox는 3D 프린팅 관련 잠재적 기술기회의 발굴을 위한 과정을 묘사하는 데 있어 매우 적합한 사례기업이라 여겨진다.

생성된 연관규칙은 선행 기술테마와 관련한 역량을 이용하여 접근할 수

62) Elliott, Amy, *Advancing liquid metal jet printing*, Oak Ridge National Laboratory, 2019, pp.1-12.

있는 결과 기술테마 관련 기술기회 간 함의관계를 보여준다. 이러한 관계를 바탕으로 사례기업인 Xerox의 새로운 잠재적 기술기회를 발굴해내기 위해서는 Xerox의 내부 역량으로 판단되는 요소가 선행 기술테마에 존재하는 반면, 결과 기술테마는 이와 무관해야만 한다. 총 70개의 기술테마 중, Xerox가 출원한 특허와 밀접한 관련을 지니고 있는 것은 25개이며 이들이 Xerox의 내부 역량을 나타내는 기술테마라고 할 수 있다. 따라서 생성된 40개의 기술테마 간 연관규칙에서 Xerox의 내부 역량 기술테마를 선행 기술테마로서 지니고 그 외의 기술테마를 결과 기술테마로서 지니는 규칙만을 추출한다. 이를 통해 Xerox의 내부 역량을 기반으로 접근 가능한 잠재적 기술기회를 결과 기술테마의 관점에서 제시할 수 있게 된다. 이러한 형태의 접근은 기업과의 적합성을 고려하지 않고 일반론적인 관점에서 도출하는 공허한 기술기회가 아닌, 기업의 내부 역량을 통해 접근 가능한 실질적 기술기회를 기술테마의 형태로 제시한다는 점에서 해당 기업에게 적합한 기술기회의 발굴을 가능하게 하는 장점을 지닌다.

추출된 모든 연관규칙을 바탕으로 Xerox는 새로운 기술기회에 대한 탐색을 실행할 수 있다. 하지만 이들에 대한 우선순위를 판단할 수 있는 기준을 제시해주지 못한다면 발굴된 기술기회의 활용성을 저해하게 될 것이다. 연관규칙의 특성을 살펴보기 위하여 일반적으로 지지도, 신뢰도, 향상도의 3가지 기준이 활용된다. 지지도는 규칙 자체에 대한 의미를 나타내기보다는 규칙을 구성하는 선행 및 결과 기술테마들의 단순 출현빈도에 대한 비중만을 표현하는 것으로서 그들 간 연관성의 수준을 포괄하지 못한다. 향상도는 규칙을 구성하는 기술테마들 간 양의 연관성 및 음의 연관성을 구별하여 나타내지만 방향성이 없어 $A \rightarrow B$ 와 $B \rightarrow A$ 의 값이 언제나 동일할 뿐만 아니라 가질 수 있는 값의 범위에 제한이 없다.⁶³⁾ 따라서 본 연구에서는 선행 기술테마를 통한 결과 기술테마로의 접근 형태를 지니는 기술기회에 대한 잠재적 수준을 정량화하기 위한 수단으로 연관규칙의 신뢰도를 이용한다.

63) 박희창, "The development of relative interestingness measure for comparing with degrees of association", 『한국데이터정보과학회지』, 제19권 제4호(2008), 1269-1279면.

〈표 3〉은 이러한 조건에 부합하는 연관규칙 중 신뢰도 기준 상위 5개를 보여준다. 최종적으로 T4, T6, T20, T62와 관련한 내부 역량을 활용하여 T42, T67, T38, T23, T46과 관련한 기술적 요소들을 잠재적 기술기회로서 활용할 수 있음을 나타낸다.

〈표 3〉 Xerox의 내부기술테마 및 외부기술테마 간 연관규칙(신뢰도기준 상위 5개)

선행 내부기술테마	결과 외부기술테마	지지도	신뢰도	향상도
T4	T42	0.0015	0.0640	2.0264
T6	T67	0.0012	0.0517	2.3178
T6	T38	0.0012	0.0517	2.0947
T20	T23	0.0016	0.0438	2.8107
T62	T46	0.0011	0.0410	1.9540

〈표 4〉는 Xerox의 내부 역량으로서의 의미를 지니는 내부기술테마와 이를 통해 접근 가능한 기술기회로서의 의미를 지니는 외부기술테마에 기술적 의미를 부여한 결과이다. LDA의 적용을 통해 생성된 기술테마는 그 의미가 자동적으로 정의되지 않기 때문에 해당 기술테마와 관련된 키워드를 활용하여 개별적으로 의미를 파악하기 위한 과정이 실행되어야 한다. 이를 위하여 관련 전문가와의 인터뷰를 수행하였고 내부 및 외부기술테마에 대한 기술적 의미를 생성한 후 〈표 4〉와 같이 정리하였다. 본 연구의 사례 분석 기술분야는 3D 프린팅이기 때문에 스마트제조 관련 박사학위 소지자 3명과 3D 프린터 판매 및 교육업에 종사하는 박사학위 소지자 1명 및 학사학위 소지자 1명 등 총 5명의 관련 전문가와 인터뷰를 수행하였다. 각 기술테마와 연관된 기술 키워드 및 관련 특허들을 제시한 후, 각 테마가 지니는 기술적 의미에 대한 의견을 수집하고 해당 결과를 종합함으로써 최종적으로 기술테마에 대한 기술적 의미부여가 이루어졌다.

T4는 3D 프린팅 제품의 표면작업을 포괄하는 후공정 관련 기술테마이고, T42는 3D 프린팅을 위한 복합재료 관련 기술테마이다. 3D 프린팅 후공정은 기본적으로 재료의 특성에 대한 높은 이해를 지니고 있어야 가능한 영역이

다. Xerox는 PARC(Palo Alto Research Center)를 통한 지속적 기술혁신 활동을 수행하면서 고분자 플라스틱 등 복합재료에 대한 풍부한 역량을 이미 보유하고 있다. 따라서 이러한 내부 역량을 기반으로 하여 3D 프린팅 복합재료를 개발하기 위한 활동은 충분한 가능성을 지닌 기술기회로서 판단할 수 있다.

T6은 3D 프린터 노즐 등 부품의 정밀한 조정과 관련한 기술테마이고, T67과 T38은 각각 자유도가 높은 다중노즐 및 3D 프린팅 인공관절과 관련한 기술테마이다. 다양한 재료의 안정적 출력을 위하여 싱글노즐 및 듀얼노즐을 활용하는 기술이 다각도로 활용되어 왔다. 하지만 이들은 주로 서포트와 출력물의 복합출력이나 다중색깔의 출력 등의 목적을 지니고 있었던 반면,⁶⁴⁾ 다중노즐은 서로 다른 재료를 빠르고 연속적으로 출력할 수 있도록 하기 때문에 3D 복합재료 부품의 정밀한 출력을 위해 필수적으로 요구되는 기술로서 인식된다. Xerox의 노즐에 대한 정밀한 조정역량 및 인장강도 향상을 위한 다중노즐의 활용역량은 복합재료 부품의 출력을 위해 요구되는 다중노즐 관련 기술기회를 실현할 수 있는 충분한 기반이 될 것이다. 3D 프린팅을 위한 대표적 생체적합 소재로는 금속, 세라믹, 고분자 등이 있으며,⁶⁵⁾ Xerox는 폴리설폰 등 고분자 및 액상메탈에 대한 충분한 역량을 보유하고 있어 3D 프린팅 인공관절 역시 충분히 합리적인 기술기회로 판단할 수 있다.

T20은 3D 데이터 처리 관련 기술테마이고, T23은 3D 프린팅 제품들의 효율적 결합을 위한 위치 조절 관련 기술테마이다. 3D 프린팅을 통해 생산된 제품들은 다른 제품을 구성하는 하나의 요소가 될 수 있기 때문에 이들을 결합하기 위한 작업이 필수적이다. 금속이나 플라스틱을 포함하는 모든 재료에 있어 3D 프린팅 제품들은 접착, 기계적 결합 및 용접이라는 전통적인 3가지 프로세스를 이용하여 결합되며, 이러한 결합 공정은 3D 프린터가 아주 복잡하고 거대한 제품까지도 하나의 완성품으로 생산해내는 능력을 갖추기 전

64) 김태영·이용구, “듀얼 노즐 FDM 프린터에서 노즐 간의 간섭을 최소화하는 모델의 빌드 방향 최적화를 위한 방법”, 『한국 CDE 학회 논문집』, 제22권 제1호(2017), 37-43면.

65) 서미란, 『국내 외 3D프린팅 활용사례와 시사점』, 정보통신산업진흥원, 2019, 1-15면.

〈표 4〉 Xerox의 내부 역량 및 기술기회로서의 의미를 지니는 내·외부기술테마

기술테마 구분	기술테마		높은 관련도를 지닌 키워드(일부)
내부 기술테마 (내부 역량)	T4	Surface finishing for objects	surface, finish, heat, temperature, smooth, material, component, coat, layer, particle, remove, treatment, shape, pressure, rough
	T6	Object arrangement	arrange, angle, face, path, align, edge, line, parallel, segment, point, space, offset, position, horizontal, vector
	T20	3D data processing	data, processor, transmit, visual, compute, file, control, digital, communicate, code, interface, program, network, parameter, software
	T62	Chamber for metal additive manufacturing	chamber, wall, space, flow, tube, inlet, nozzle, assembly, conduit, outer, inner, fluid, outlet, cavity, pipe
외부 기술테마 (기술 기회)	T23	Precision locating and assembly	assembly, precision, mount, device, machine, tool, configure, feature, move, position, point, predetermine, place, holder, rotate
	T38	3D printed prosthetic joint	prosthetic, bone, joint, replace, body, implant, patient-specific, surgical, orthopedic, anchor, shape, instrument, anatomy, treatment, align
	T42	Advanced composite material	material, powder, composite, particle, titanium, aluminum, alloy, mixture, binder, combine, ceramic, feedstock, copper, silicon, sinter
	T46	Heat exchanger or transfer	heat, thermal, exchange, transfer, apparatus, device, chamber, fluid, flow, vapor, channel, manage, energy, pressure, pipe
	T67	Adaptive nozzle	nozzle, adaptive, configurable, flow, fluid, filter, channel, control, conduit, device, pump, diffuse, sensor, dispose, dispense,

까지는 반드시 이루어져야 하는 후공정 작업이다.⁶⁶⁾ 따라서 부품 결합을 위해 요구되는 정밀한 위치 통제 및 조절 관련 데이터 처리 역량은 결합 공정을 보다 효과적으로 실행하기 위한 기술적 속성으로서 가치를 지닌다.

T62는 온도조절, 밀폐 등의 역할을 통해 열간 등방압 가압을 발생시켜 단조부품과 유사한 품질을 보장할 수 있도록 하는 금속 3D 프린팅을 위한 챔버 관련 기술테마이고, T46은 열교환기 관련 기술테마이다. 발전소, 냉동·

66) Loctite, *3D printing bonding guide*, Henkel, 2020, pp.1-12.

공조기기 등에 있어 에너지 효율에 큰 영향을 미치는 핵심 부품인 열교환기는 금속 3D 프린팅을 활용하여 생산 가능한 대표적인 제품으로 경량화, 열전도율 향상 및 비용절감 등의 효과를 창출할 수 있어 3D 프린팅 기반 제조 혁신 기술의 적용 가능성이 매우 높다.⁶⁷⁾ 따라서 이는 Xerox의 금속 3D 프린팅에 대한 대표적 응용분야로서의 가치를 지니는 것으로 판단할 수 있다.

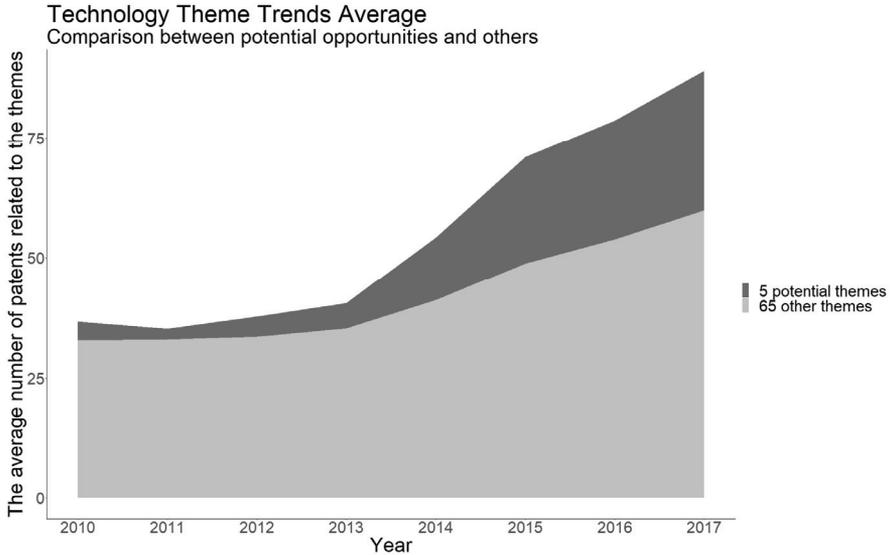
V. 논의 사항

1. 잠재적 기술테마의 트렌드 분석

기술테마 간 연관규칙에서 제시하는 모든 외부기술테마가 항상 사례기업에게 기술기회로서의 잠재적 가치를 제공해준다고 단언할 수는 없다. 그러한 외부기술테마의 최근 트렌드를 살펴보고 관련한 기술개발 활동이 활발한지에 대한 확인이 추가적으로 이루어져야 한다. 모든 외부기술테마에 대해서 그들이 도입, 성장, 정체, 쇠퇴와 같은 수명주기상 어느 위치에 존재하는지를 구체적으로 탐색하는 것은 본 연구의 범위를 벗어나는 것일지라도 기술개발 활동에 대한 트렌드가 증가양상을 보이는지에 대한 확인 정도는 요구되는 것이다. 이를 살펴볼 수 있는 가장 객관적 기준은 각 기술테마와 관련한 기술개발 활동의 수준을 기술혁신과 연구개발의 동향에 대한 대리변수로 널리 활용되는 특허를 이용하여 정량적 관점에서 측정하는 것이다. 물론 3D 프린팅에 대한 기술적 관심이 급증해오고 있기 때문에 절대적 기준으로 이를 살펴보게 되면 당연하게 급격한 증가추세를 보일 것이다. 따라서 <표 3> 및 <표 4>에서 Xerox의 새로운 기술기회로서 제시된 외부기술테마와 그 외 나머지 기술테마 간 트렌드를 비교한다.

67) 나덕주, 『3D 프린팅 기술의 현황과 미래 전망』, 한국과학기술정보연구원, 2019, 1-10면.

〈그림 6〉 Xerox의 잠재적 기술테마 및 나머지 기술테마 관련 특허 트렌드 비교



본 연구에서 활용한 데이터는 2010년부터 2019년까지 USPTO에 출원 및 등록된 특허인데, 2018년 및 2019년의 특허는 아직 등록 또는 공개가 이루어지지 않은 특허가 많기 때문에 2010년부터 2017년까지의 특허 데이터를 활용하여 트렌드를 파악한다. 〈그림 6〉은 Xerox의 내부 역량을 통해 접근 가능한 잠재적 기술기회로서 제시된 5개의 외부기술테마(T23, T38, T42, T46, T67)와 나머지 65개의 기술테마에 대해 그들 각각과 관련된 특허의 수를 비교하여 나타낸다. 5개의 잠재적 기술테마와 65개의 나머지에 대해 특허의 절대적 양을 비교하는 것은 무의미하기 때문에 이들의 평균을 산출하여 기술테마 1개당 평균 특허의 개수를 산출하여 시각화한다. 이러한 과정을 통해 제시된 〈그림 6〉을 살펴보면 잠재적 기술테마 관련 기술개발 활동의 증가양상이 월등하게 뚜렷함을 알 수 있다. 잠재적 기술테마 1개당 평균 특허의 개수는 2010년 3.8개, 2017년 29.0개인 반면, 나머지 기술테마 1개당 평균 특허의 개수는 2010년 32.82개, 2017년 59.94개로, 연평균 증가율이 각각 33.69%, 8.99%로 큰 차이를 보인다. 이는 잠재적 기술테마에 대한 관심도가 초기에

극도로 낮아 여전히 높은 신규성을 지니고 있을 뿐만 아니라 이를 기술적으로 활용하기 위한 기술개발 활동이 최근에는 급격하게 증가하고 있기에 미래 발전 가능성이 상당히 높은 유망한 기술기회가 될 수 있음을 확인할 수 있다.

2. 특허와 기술테마 간 할당관계 정제에 의한 영향

본 연구는 기업의 내부 역량을 바탕으로 접근 가능한 기술기회를 발굴하고자 하며, 기업의 역량은 기업의 특허 포트폴리오를 이용하여 표현할 수 있기 때문에 특허와 기술테마 간 할당관계를 생성함으로써 기업의 내부 역량을 기술테마 관점에서 정의할 수 있다. LDA의 적용은 특허문서와 토픽 간 확률적 연관관계를 나타내는 문서-토픽 매트릭스를 생성한다. 기술테마는 토픽을 기반으로 정의되기 때문에 이러한 연관관계를 활용하여 특허를 기술테마에 할당할 수 있다. 하지만 특허와 토픽 간 확률 값 역시 상당히 편차가 커서 이 매트릭스에 존재하는 모든 관계를 기반으로 특허에 기술테마를 할당하게 되면 너무 미미한 관계까지도 포괄해버리는 문제가 있다. 즉, 하나의 기업에 거의 모든 기술테마가 내부 역량으로서 포함되게 되고 해당 기업과 무관하여 새롭다고 여겨질 수 있는 기술테마로의 접근을 제안하는 연관규칙이 거의 생성되지 못하는 문제가 발생할 수 있는 것이다. 이에 본 연구에서는 적절한 기준값을 설정하고 해당 값 이상의 확률관계를 지니는 관계만을 기반으로 특허와 기술테마 간 할당관계를 정의하였다. 기준값을 설정하기 위한 원칙은 기술테마의 기술적 의미가 명확하게 나타날 수 있도록 하고, 기술테마 간 중복성을 최대한 줄일 수 있도록 한 특허에 할당되는 기술테마의 수가 1개를 크게 초과하지 않도록 하는 것이다. 하나의 특허문서는 기본적으로 하나의 기술적 속성을 포괄하기 때문에 그러한 속성을 기술테마의 형태로서 활용할 수 있도록, 한 특허에 할당되는 기술테마의 수가 1개 수준으로 유지되도록 하는 것이 합리적이기 때문이다. 이는 특허 및 해당 특허를 보유한 기업과 기술테마 간 관계를 보다 명확하게 하는 장점이 있다. 이러한 형태의 컷오프 기반의 할당관계 정의는 특허와 기술테마 간 전체의 연관관

계 중 일부만을 선택적으로 활용함을 의미한다. 즉, 각 연관관계 중, 평균의 컷오프 기준에 따라 이에 미치지 못하는 부분은 노이즈로 판단하는 반면, 이를 넘어서는 부분은 본 연구에서 활용 가능한 의미있는 관계로서 판단하는 것이다. 이러한 판단은 분명 본 연구의 분석결과에 직접적 영향을 미칠 수 있다. 본 연구에서 특허는 하나의 기술적 사상에 대한 명확한 설명을 기술하고 있는 문서로서 출원인의 개별 기술적 역량을 모델링하는 단위로서 활용된다. 따라서 출원인 기업이 복수 개의 특허로 구성된 특허 포트폴리오를 구성하고 있는 경우, 해당 기업이 각 특허로 표현되는 개별 기술들에 대한 복합적 역량을 보유하고 있다고 판단한다. 이러한 판단은 하나의 특허가 하나의 개별 기술에 대한 명확한 설명을 포괄할 수 있음을 가정함으로써 가능하게 된다. 따라서 본 연구에서는 평균적으로 하나의 특허가 하나의 개별 기술만을 포괄할 수 있도록 즉, 한 특허에 할당되는 기술테마의 수가 1개 수준으로 유지되도록 하는 원칙을 활용하게 되었다.

이러한 원칙에 의해 최종적으로 약 21%에 해당하는 특허-기술테마 간 할당관계만이 연관규칙 생성에 활용되게 되었다. 만약 하나의 특허에 할당되는 기술테마의 수를 크게 늘리게 되면, 현재보다 더 많은 할당관계를 활용할 수 있는 장점은 있을 것이다. 하지만 그 반대급부로 명확하지 않고 상대적으로 모호한 할당관계가 연관규칙 생성에 활용되게 되고, 이는 기술적 함의관계가 강하지 않은 연관규칙을 생성해내는 문제점을 유발할 것이다. 본 연구에서는 기술테마 간 연관규칙에 대한 해석적 접근을 통해 잠재적 기술기회를 도출하게 되는데, 함의관계가 강하지 않은 연관규칙의 경우, 잠재적 기술 기회로서의 가치를 판단하기 어려운 문제를 유발하게 된다. 이는 본 연구에서 제안하는 방법론의 활용성을 저하시키는 요인이 될 것이다. 만약 하나의 특허에 할당되는 기술테마의 수를 지금보다도 더 줄이게 되면, 연관규칙 생성에 활용할 할당관계의 의미적 명확성이 높아지는 장점은 있을 것이다. 하지만 그 반대급부로 연관규칙 생성에 활용할 복합적 관계의 절대적 수가 줄어들기 때문에 분석을 시행할 충분한 수의 연관규칙을 확보하지 못하게 될 우려가 있다. 게다가 연관규칙을 구성하는 선행 및 결과 기술테마 간 직접적

인 함의관계 수준의 절대적 값이 매우 낮아지는 결과를 초래할 수 있다. 따라서 본 연구에서는 잠재적 기술기회발굴을 위해 충분한 수의 연관규칙을 생성하면서도 그들의 기술적 함의관계를 명확하게 해석할 수 있도록 하는 원칙에 따라 특허와 기술테마 간 할당관계를 정의하게 되었다.

VI. 결론

기술기회의 발굴은 기술 트렌드의 변화를 미래 관점에서 파악하고 이에 대응하기 위한 기술적 아이디어를 도출해낼 수 있도록 하기 때문에 그 중요성이 매우 높다. 이에 본 연구는 기술적 함의에 대한 구체적 내용을 기술하는 기술적 속성의 집합으로 기술테마를 정의하고 기술테마 간 연관관계에 대한 분석을 통해 기술기회를 발굴해내기 위한 방안을 제시하였다. 그리고 발굴된 기술기회의 타당성을 확인하기 위하여 해당 기술테마 관련 기술개발 활동의 트렌드에 대한 논의를 수행하였다. 뿐만 아니라 본 연구에서 제시하는 방안의 실제적 활용성을 탐색하기 위하여 최근 제품 생산방식에 큰 변화를 유발하고 있는 3D 프린팅 분야를 대상으로 사례 분석도 시행하였다. 본 연구는 국내 기업들의 취약한 R&D 기획역량을 증진시키고 지속 가능한 성장을 위한 신성장동력원의 조기 발굴 및 선점을 지원함으로써 궁극적으로 차세대 R&D 기획 시스템으로서의 역할을 수행할 수 있을 것으로 기대된다.

이러한 기여점에도 불구하고 본 연구는 여러 한계점을 지니고 있기에 다양한 추후 연구들이 수행될 필요가 있다. 우선, 본 연구는 특정 기술분야에 대한 특허 데이터를 수집한 후, 이를 활용하여 기술개발 주체인 기업의 내부 역량에 해당하는 내부기술테마와 잠재적 기술기회로서 논의될 수 있는 외부 기술테마를 모두 생성하였다. 이러한 접근은 해당 기술분야 내에서의 기술진보만을 이루어낼 수 있는 한계를 지닌다. 즉, 해당 분야와 기술적 속성이 완전히 다른 분야의 내부 역량을 활용하여 접근 가능한 기술기회의 발굴은 원천적으로 불가능하다. 따라서 기술분야의 범위를 유동적으로 설정할 수

있는 방안의 접목이 요구된다. 두 번째로, 본 연구는 완전히 새로운 기술기회의 발굴이 불가능하다는 한계를 지닌다. 신규성에 대한 탐색을 통해 도입 또는 성장기와 같은 수명주기상 위치에 대한 추정 은 가능하지만, 기존에 존재하지 않았던 새로운 기회를 창출해내지는 못한다. 특허분석을 통한 기술기회발굴 방안의 하나인 형태학 분석은 여러 서브시스템들을 분석 및 조합하는 탐색적 접근을 통해 새로운 개념을 창출해내는 특징이 있다. 따라서 이러한 형태학 분석을 본 연구의 방안과 융합적으로 활용함으로써 새로운 기회의 창출이 불가능한 한계를 극복할 수 있을 것이다. 마지막으로, 본 연구에서는 특허 데이터에 LDA를 적용하여 기술적 함의를 포괄하는 토픽을 추출하고 이를 기술테마로 정의하였다. 기술테마에 기술적 의미를 부여하는 과정은 토픽과 키워드 간 확률적 관계에 기반하기 때문에 키워드가 매우 중요한 역할을 차지하게 된다. 본 연구는 기술기회발굴 분석의 프로토타입 관점에서 단순 키워드 중심의 분석을 시행하였다. 기술적 의미에 대한 명확성을 제고함으로써 발굴된 기술기회의 합리적 타당성을 증진하기 위해서는 이를 복합어 단위로 확장하기 위한 후속연구가 수행되어야 한다.

VII. 부 록

1. 특허-토픽 간 관계를 활용한 apriori 알고리즘 기반 연관규칙 생성과정

특허와 토픽 간 관계를 활용하여 연관규칙을 생성하기 위해서는 하나의 특허문서를 개별 트랜잭션으로 가정하고, 해당 특허문서와 연관성을 지니는 기술테마를 트랜잭션 내에 존재하는 항목집합으로 가정하는 것이 필요하다. 예를 들어, 4개의 특허가 각각 (기술테마A, 기술테마B, 기술테마C), (기술테마B, 기술테마D), (기술테마A, 기술테마C), (기술테마C, 기술테마D)와 관련되어 있다고 가정하자. 이러한 경우, 항목집합 4개와 트랜잭션 4개로 구성된 데이터셋을 구성할 수 있고 이로부터 연관규칙을 생성할 수 있다. 첫

번째 단계는 최소 지지도 값에 대한 지정을 통해 빈발항목집합을 구성하는 것이다. 만약 최소 지지도가 0.5라고 한다면, 해당 조건을 만족하는 집합을 개수를 늘려가면서 탐색한다. 우선 기술테마 A, B, C, D 각각 하나씩의 관점에서 보면, A, B, D 는 2번, C 는 3번씩 트랜잭션에 포함되어 있다. 따라서 이들 각각의 지지도는 다음과 같다.

$$P(A) = P(B) = P(D) = \frac{2}{4}, P(C) = \frac{3}{4} \quad (A1)$$

모두 최소 지지도 이상이기 때문에 모두를 포함하여 2개씩 묶음을 판단한다. 모든 조합에 대한 지지도를 산출하면 다음과 같다.

$$P(A \cap B) = \frac{1}{4}, P(A \cap C) = \frac{2}{4}, P(A \cap D) = 0, P(B \cap C) = \frac{1}{4}, P(B \cap D) = \frac{1}{4}, P(C \cap D) = \frac{1}{4} \quad (A2)$$

이 중, 최소 지지도 이상을 나타내는 것은 (A, C) 밖에 없다. 그러므로 우리가 구성할 수 있는 빈발항목집합은 (A, C) 가 유일하고 이로 부터 생성할 수 있는 규칙은 기술테마 A 와 C 사이의 연관관계를 나타내는 $A \rightarrow C$ 와 $C \rightarrow A$ 둘 뿐이다. 규칙을 생성하고 난 후, 지지도, 신뢰도 및 향상도에 대한 측정이 이루어진다. 우선, 지지도의 경우 두 규칙 모두 $P(A \cap C) = P(C \cap A)$ 로서 식 (A2)에서 볼 수 있는 것처럼 $\frac{2}{4}$ 이다. 두 규칙의 신뢰도 및 향상도는 다음 식을 통해 산출할 수 있다.

$$P(C|A) = \frac{P(A \cap C)}{P(A)} = \frac{2/4}{2/4} = 1, \quad \frac{P(CA)}{P(C)} = \frac{1}{3/4} = \frac{4}{3} \quad (A3)$$

$$P(A|C) = \frac{P(A \cap C)}{P(C)} = \frac{2/4}{3/4} = \frac{2}{3}, \quad \frac{P(AC)}{P(A)} = \frac{2/3}{2/4} = \frac{4}{3} \quad (A4)$$

즉, 규칙 $A \rightarrow C$ 의 신뢰도는 1, 향상도는 $\frac{4}{3}$ 이고, 규칙 $C \rightarrow A$ 의 신뢰도는 $\frac{2}{3}$, 향상도는 $\frac{4}{3}$ 이다. 만약 최소 신뢰도가 0.8이라고 한다면, 최종적으로 생성되는 규칙은 오직 $A \rightarrow C$ 뿐이다.

참고문헌

〈단행본(국내 및 동양)〉

- 김용기, 『3D 프린팅 시장 및 기술 동향』, 과학기술일자리진흥원, 2019.
나덕주, 『3D 프린팅 기술의 현황과 미래 전망』, 한국과학기술정보연구원, 2019.
서미란, 『국내 외 3D프린팅 활용사례와 시사점』, 정보통신산업진흥원, 2019.

〈단행본(서양)〉

- Ceulemans, Judy et al., *Patents and additive manufacturing: Trends in 3D printing technologies*, European Patent Office, 2020.
Elliott, Amy, *Advancing liquid metal jet printing*, Oak Ridge National Laboratory, 2019.
Loctite, *3D printing bonding guide*, Henkel, 2020.

〈학술지(국내 및 동양)〉

- 김태영 · 이용구, “듀얼 노즐 FDM 프린터에서 노즐 간의 간섭을 최소화하는 모델의 빌드 방향 최적화를 위한 방법”, 『한국 CDE 학회 논문집』, 제22권 제1호(2017).
박상현 외 3인, “동적토포모델링과 네트워크 분석 기반의 기술융합 동향분석 — ICT-Agritech 융합 사례 연구”, 『대한산업공학회지』, 제46권 제3호(2020).
박영진 외 2인, “보유특허 기반의 기술기회탐색을 위한 특허추천방법 — 3D 프린팅 산업을 중심으로”, 『지식재산연구』, 제10권 제1호(2015).
박현석 외 4인, “기업의 보유 기술 및 제품에 기반한 기술기회발굴”, 『대한산업공학회지』, 제40권 제5호(2014).
박희창, “The development of relative interestingness measure for comparing with degrees of association”, 『한국데이터정보과학회지』, 제19권 제4호(2008).
오승현 외 2인, “특허의 토포 모델링을 활용한 증강현실 기술 모니터링”, 『대한산업공학회지』, 제43권 제3호(2017).
정으뜸 외 2인, “기술적 파급효과 도출을 통한 기술융합 분석 연구 — 나노바이오 융합 기술 사례를 중심으로”, 『지식재산연구』, 제10권 제4호(2015).
정재민 외 2인, “비즈니스 기회 발굴을 위한 문제-해결방법 기반의 특허분석 방법”, 『지식재산연구』, 제15권 제2호(2020).
지선구 · 한덕원, “3D 프린팅과 특허발명의 법적 보호”, 『지식재산연구』, 제11권 제1호(2016).

최성철·서원철, “기술평가를 위한 평가 참조정보 생성 방안에 관한 연구 — 기술적 속성 유사성 관점에서”, 『지식재산연구』, 제14권 제2호(2019).

〈학술지(서양)〉

- Abernathy, William & Clark, Kim, “Innovation: mapping the winds of creative destruction”, *Research Policy*, Vol.14 No.1(1985).
- Altuntas, Serkan et al., “Analysis of patent documents with weighted association rules”, *Technological Forecasting and Social Change*, Vol.92(2015).
- Blei, David & Lafferty, John, “Correlated topic models”, *Advances in Neural Information Processing Systems*, Vol.18(2006).
- Choi, Jaewoong et al., “Technology opportunity discovery under the dynamic change of focus technology fields: Application of sequential pattern mining to patent classifications”, *Technological Forecasting and Social Change*, Vol.148 (2019).
- Choi, Sungchul et al., “SAO network analysis of patents for technology trends identification: a case study of polymer electrolyte membrane technology in proton exchange membrane fuel cells”, *Scientometrics*, Vol.88 No.3(2011).
- Choi, Sungchul et al., “Analyzing technological spillover effects between technology classes: The case of Korea technology finance corporation”, *IEEE Access*, Vol.6(2018).
- Cozzens, Susan et al. “Emerging technologies: quantitative identification and measurement.” *Technology Analysis & Strategic Management*, Vol.22 No.3(2010).
- Feng, Lijie et al., “Discovering technology opportunity by keyword-based patent analysis: A hybrid approach of morphology analysis and USIT”, *Sustainability*, Vol.12 No.1(2020).
- Gilbert, Andrew et al., “Action recognition using mined hierarchical compound features”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.33 No.5(2010).
- Han, Xiaotong et al., “R&D trend analysis based on patent mining: An integrated use of patent applications and invalidation data”, *Technological Forecasting and Social Change*, Vol.167(2021).
- Jelodar, Hamed et al., “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey”, *Multimedia Tools and Applications*, Vol.78 No.11(2019).

- Jeong, Byeongki et al., "Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis", *International Journal of Information Management*, Vol.48(2019).
- Kim, Chulhyun et al., "Identifying core technologies based on technological cross-impacts: An association rule mining (ARM) and analytic network process (ANP) approach", *Expert Systems with Applications*, Vol.38 No.10(2011).
- Kim, Jusung et al., "Identifying potential technology themes based on internal capabilities using topic modeling and association rule mining", *ICIC Express Letters, Part B: Applications*, Vol.7 No.10(2016).
- Kim, Mujin et al., "Generating patent development maps for technology monitoring using semantic patent-topic analysis", *Computers & Industrial Engineering*, Vol.98(2016).
- Kim, Young Gil et al., "Visualization of patent analysis for emerging technology", *Expert systems with applications*, Vol.34 No.3(2008).
- Ko, Namuk et al., "Analyzing interdisciplinarity of technology fusion using knowledge flows of patents", *Expert Systems with Applications*, Vol.41 No.4(2014).
- Lee, Changyong et al., "Monitoring trends of technological changes based on the dynamic patent lattice: A modified formal concept analysis approach", *Technological Forecasting and Social Change*, Vol.78 No.4(2011).
- Lee, Changyong et al., "Novelty-focused patent mapping for technology opportunity analysis", *Technological Forecasting and Social Change*, Vol.90 (2015).
- Lee, Changyong et al., "Navigating a product landscape for technology opportunity analysis: A word2vec approach using an integrated patent-product database", *Technovation*, Vol.96(2020).
- Lee, Hakyoon et al., "An ANP-based technology network for identification of core technologies: A case of telecommunication technologies", *Expert Systems with Applications*, Vol.36 No.1(2009).
- Lee, Sungjoo et al., "An approach to discovering new technology opportunities: Keyword-based patent map approach", *Technovation*, Vol.29 No.6-7(2009).
- Park, Hyunseok et al., "Identifying patent infringement using SAO based semantic technological similarities", *Scientometrics*, Vol.90 No.2(2012).

- Park, Youngjin & Yoon, Janghyeok, “Application technology opportunity discovery from technology portfolios: Use of patent classification and collaborative filtering”, *Technological Forecasting and Social Change*, Vol.118(2017).
- Seo, Wonchul et al., “Product opportunity identification based on internal capabilities using text mining and association rule mining”, *Technological Forecasting and Social Change*, Vol.105(2016).
- Shih, Meng-Jung et al., “Discovering competitive intelligence by mining changes in patent trends”, *Expert Systems with Applications*, Vol.37 No.4(2010).
- Song, Bomi & Suh, Yongyoon, “Identifying convergence fields and technologies for industrial safety: LDA-based network analysis”, *Technological Forecasting and Social Change*, Vol.138(2019).
- Suominen, Arho et al., “Firms’ knowledge profiles: Mapping patent data with unsupervised learning”, *Technological Forecasting and Social Change*, Vol.115 (2017).
- Trappey, Charles V. et al., “Using patent data for technology forecasting: China RFID patent analysis”, *Advanced Engineering Informatics*, Vol.25 No.1(2011).
- Yoon, Byungun & Park, Yongtae, “A systematic approach for identifying technology opportunities: Keyword-based morphology analysis”, *Technological Forecasting and Social Change*, Vol.72 No.2(2005).
- Yoon, Byungun & Park, Yongtae, “Development of new technology forecasting algorithm: Hybrid approach for morphology analysis and conjoint analysis of patent information”, *IEEE Transactions on Engineering Management*, Vol.54 No.3(2007).
- Yoon, Byungun et al., “Exploring technological opportunities by linking technology and products: Application of morphology analysis and text mining”, *Technological Forecasting and Social Change*, Vol.86(2014).
- Yoon, Janghyeok & Kim, Kwangsoo, “Identifying rapidly evolving technological trends for R&D planning using SAO-based semantic patent networks”, *Scientometrics*, Vol.88 No.1(2011).
- Yoon, Janghyeok & Kim, Kwangsoo, “Detecting signals of new technological opportunities using semantic patent analysis and outlier detection”, *Scientometrics*, Vol.90. No.2(2012).
- Yoon, Janghyeok et al., “Technology opportunity discovery (TOD) from existing

technologies and products: A function-based TOD framework”, *Technological Forecasting and Social Change*, Vol.100(2015).

Yoon, Janghyeok et al., “Identifying product opportunities using collaborative filtering-based patent analysis”, *Computers & Industrial Engineering*, Vol.107 (2017).

〈기타 자료〉

김창석 · 신준석, “기업 기술역량 기반 기술기회탐색에 관한 연구”, 『대한산업공학회 추계학술대회논문집』, 대한산업공학회, 2013.

서용윤 · 이학연, “이종사업간 지식의 형태학적 결합을 통한 신사업 기회 탐색”, 『대한산업공학회 추계학술대회논문집』, 대한산업공학회, 2013.

Agrawal, Rakesh et al., “Mining association rules between sets of items in large databases”, In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, Association for Computing Machinery, 1993.

A Study on Technology Opportunity Discovery by Analyzing Patent-Based Technology Themes

Seo, Wonchul*

Technology opportunity discovery enables us to identify changes in technology trends and to derive new technological ideas from the future-oriented perspective. Researchers have been actively working on establishing a systematic procedure for discovering technology opportunities. However, they are limited in not sufficiently considering the suitability of the company that seeks to discover new technology opportunities, such as assuming only a fixed type of scenario between product and technology, or using only the technological capability for matching technology opportunities. This study aims to present a patent-based approach to discovering technology opportunities by analyzing the relationships between technology themes that describe specific technological implications as a set of technology attributes. In addition, a case study is conducted in the field of 3D printing, which is causing a substantial change in the product manufacturing system, so that the practical usability of the results of this study can be examined. This study is expected to ultimately play a role as a next-generation R&D planning system by enhancing the companies' R&D planning capabilities

* Associate Professor, Major in Industrial Data Science & Engineering, Department of Industrial and Data Engineering, Pukyong National University.

and supporting early discovery and preoccupation of new technological sources for sustainable growth.

Keyword

Technology opportunity discovery, Patent analysis, Topic modeling, Association rule mining, 3D printing