

국가핵심기술 관계망 구축을 통한 연관정보 분석연구: 디스플레이 기술을 중심으로

A Study on the Analysis of Related Information through the Establishment of the National Core Technology Network: Focused on Display Technology

박세희(Se Hee Pak)*, 윤원석(Won Seok Yoon)**, 장항배(Hang Bae Chang)***

초 록

경제 구조의 기술 의존성이 강해져 국가핵심기술의 중요성은 더욱 대두되고 있다. 하지만 기술 자체적 특성으로 인해 연관 범위가 추상적이고 국가핵심기술 고유의 특성상 정보공개가 제한적이기 때문에 보호대상이 될 기술의 범위를 정하는 것에 어려움이 있다. 이를 해결하기 위해 국가핵심기술과 연관성이 높은 중요 기술을 판별하는 데에 최적화 된 문헌 종류와 분석 기법을 제안하였다. 디스플레이 분야 국가핵심기술 키워드로 수집한 네 개 문헌종류(뉴스, 논문, 보고서, 특허) 데이터에 빅데이터 분석의 텍스트 마이닝 분석기법인 TF-IDF와 LDA 토픽 모델링을 적용하는 파일럿 테스트를 진행하였다. 그 결과로 특히 데이터에 LDA 토픽 모델링을 적용한 결과가 국가핵심기술과 연관성이 높은 중요기술을 추출하였다. OLED, 마이크로LED를 포함하여 디스플레이 전후방산업에 관련된 중요 기술을 판별 할 수 있었으며 이 결과를 관계망으로 시각화하여 국가핵심기술과 연관된 중요 기술의 범위를 명확히 하였다. 본 연구를 통해 기술이 가지는 연관범위의 모호성을 보다 명확히 하였으며, 국가핵심기술이 가지는 제한적인 정보공개 특성을 극복할 수 있다.

ABSTRACT

As the dependence of technology on the economic structure increases, the importance of National Core Technology is increasing. However, due to the nature of the technology itself, it is difficult to determine the scope of the technology to be protected because the scope of the relation is abstract and information disclosure is limited due to the nature of the National Core Technology. To solve this problem, we propose the most appropriate literature type and method of analysis to distinguish important technologies related to National Core Technology. We conducted a pilot test to apply TF-IDF, and LDA topic modeling, two techniques of text mining analysis for big data analysis, to four types of

이 논문은 2021년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임 (P0008703, 2021년 산업혁신인재성장지원사업).

* First Author, Ph.D. Course, Department of Security Convergence, Graduate School, Chung-Ang University (saeh95@cau.ac.kr)

** Co-Author, Master's Course, Department of Security Convergence, Graduate School, Chung-Ang University (i3629i@cau.ac.kr)

*** Corresponding Author, Professor, Department of Industrial Security, Chung-Ang University (hbchang@cau.ac.kr)

Received: 2020-12-30, Review completed: 2021-02-16, Accepted: 2021-05-20

literature (news, papers, reports, patents) collected with National Core Technology keywords in the field of Display industry. As a result, applying LDA theme modeling to patent data are highly relevant to National Core Technology. Important technologies related to the front and rear industries of displays, including OLEDs and microLEDs, were identified, and the results were visualized as networks to clarify the scope of important technologies associated with National Core Technology. Throughout this study, we have clarified the ambiguity of the scope of association of technologies and overcome the limited information disclosure characteristics of national core technologies.

키워드 : 국가핵심기술, 연관기술 범위, 관계망, TF-IDF, LDA 토픽모델링, 디스플레이 기술
National Core Technology, Scope of Related Technology, Concept Network,
TF-IDF(Term Frequency-Inverse Document Frequency), LDA Topic Modeling,
Display Technology

1. 서 론

제 4차 산업혁명과 팬데믹(Pandemic)의 영향으로 국가의 경제가 기술 중심으로 변화하였다. 전 세계 선진국들은 기술경쟁력 확보와 보호를 위하여 강한 노력을 하고 있는데, 우리나라에서는 2006년 최초 제정된 “산업기술의 유출방지 및 보호에 관한 법률”(이하 산업기술보호법)을 통해 산업기술을 정의하고 그 범위를 설정해 기술들을 보호하고 있다[1].

산업기술보호법 제2조의 2에 따른 산업기술의 범위 중 국가적으로 가장 중요한 기술이 바로 “국가핵심기술”이다. 국가핵심기술이란, 국내외 시장에서 차지하는 기술적·경제적 가치가 높거나 관련 산업의 성장잠재력이 높아 해외로 유출 될 경우에 국가의 안전보장 및 국민경제의 발전에 중대한 악영향을 줄 우려가 있는 기술이다. 이에 대한 구체적인 지정방법과 보호방안은 제9조 4항의 “국가핵심기술 지정 등에 관한 고시”(이하 고시)에서 규정하고 있다.

고시에 따라 2019년 7월, 총 12개 분야의 69개 기술이 국가핵심기술로 지정되어 약 1년 6

개월 간 유지되었으나 2021년 4월부터 시행을 계획으로 한 개정안이 2021년 1월에 발표되었다. 개정안에 따르면 지정된 지 10년이 경과하여 진입장벽이 낮아진 4개의 기술은 지정 해제되고 5개의 신기술이 추가로 지정되어 총 71개로 변경되었다[1, 35].

국가핵심기술로 지정된 신기술들은 반도체 분야에서 2개, 정보통신 분야에서 1개, 생명공학 분야에서 1개, 우주 분야에서 1개이다[1, 35].

이 기술들은 모두 선진국들이 국가차원의 지원을 아끼지 않으며 선점하려 노력하고 있는 차세대 핵심기술이다. 하지만 국가핵심기술은 심각한 유출 위협에 노출되어있으며 이로 인한 피해가 심각하다. 2020년 9월 산업통상자원부 등에서 발표한 바에 따르면 지난 6년간 산업기술의 해외 유출사건은 총 121건으로, 이 중 약 24%인 29건이 국가핵심기술에 해당한다. 국가핵심기술 보유기관 중 지난 3년 이내에 1회 이상 기술 유출 사고를 경험한 기관의 비율은 신규 기업을 제외하고 7.8%였다[7].

특히 디스플레이 분야는 2012~2017년 6년간 발생한 국가핵심기술 유출사건 중 45%로 가장

많은 비중을 차지하였으며, 2017년 7월에는 해당 시점에 한국이 세계 시장점유율 1위였던 OLED (Organic Light Emitting Diodes) 세정 기술이 경쟁국에 유출되었다[15]. 국가핵심기술 뿐만 아니라 산업 내 중요기술이 유출 되어 경쟁국 으로부터의 추격 위협이 강한 상황이다.

중요기술의 유출방지를 위해 국가핵심기술 로 지정하였음에도 불구하고 유출사건이 발생 하는 이유는 기술의 범위를 정하는데 한계가 있기 때문이다. 기술 자체의 특성상 주변 기술 과의 연관정도를 추정하는 데에 어려움이 있을 뿐 아니라, 산업기술보호법 제9조 2항에 따라 국가핵심기술의 범위를 “필요최소한의 범위” 로 규정하여 관리와 보호를 위한 범위를 파악 하기 어렵기 때문이다[1].

또한 동법 제9조의2에 따라 국가핵심기술에 대한 정보는 특수한 상황이 아닐 경우 공개하 지 않는다[1]. 이외에도 국가핵심기술 보유 기관은 인수합병, 기술의 해외수출 등 기술 성과 와 관련된 모든 활동에 강도 높은 규제가 적용 된다. 이러한 이유로 국가핵심기술 선정 범위를 넓힌다면 산업의 발전 속도를 저감할 우려 가 있기 때문에 신중한 절차를 밟아 제한적인 범위의 핵심기술을 국가핵심기술로 선정한다. 하지만 이번 개정안에서는 이러한 한계를 극복 하고자 국가핵심기술의 범위를 확대·조정하였 다고 밝혔기 때문에 국가핵심기술의 보호범위 를 설정하는 연구로 국가핵심기술이 가지는 연 관 범위를 더욱 구체화하여 제한적인 정보 공 개 특성을 극복하고 보호 대상의 범위를 확대 규명 할 필요가 있다.

본 연구에서는 디스플레이 분야 국가핵심기 술 키워드로 뉴스, 논문, 보고서, 특허의 4가지 문헌에서 데이터를 수집하여 텍스트 마이닝 분

석기법인 TF-IDF와 LDA 토픽 모델링을 적용 하는 파일럿 테스트를 진행하였다.

파일럿 테스트 결과 중 국가핵심기술과 연관 성이 높은 중요 기술을 가장 많이 도출해 낸 방법으로 기술 관계망을 구축하여 국가핵심기 술의 연관 기술 범위를 구체화 및 시각화 하였다.

2. 선행연구

2.1 디스플레이 분야 국가핵심기술 유출 사례 분석

디스플레이(Display)란 전자 기기로부터 출 력되는 전기신호를 화상 정보로 나타내어 다양 한 정보를 인간이 시각적으로 볼 수 있도록 화 면으로 구현해 주는 영상 표시장치를 의미한다. 디스플레이 산업이란, 디스플레이 디바이스 및 연관 소재·부품·장비를 생산하는 산업을 총 칭하며, LCD(Liquid Crystal Display), OLED (Organic Light Emitting Diode) 등 평판디스 플레이(Flat panel display(FPD))(이하 패널)가 주력이다[26, 32].

TFT-LCD(Thin film Transistor Liquid Crystal Display)와 AMOLED(Active Matrix Organic Light-Emitting Diode, 능동형 유기발 광다이오드)기술이 디스플레이 분야 국가핵심 기술로 지정되어있다[26, 35].

TFT-LCD는 매우 얇은 액정을 통하여 정보 를 표시하는 디지털 디스플레이로 박막 트랜지 스텐터 기술을 활용하여 화질을 향상시킨 LCD기 술의 변종이다. AMOLED는 백라이트에 의해 빛을 발하는 LCD와는 달리 자체에서 빛을 발 하는 디스플레이 기술이다[26, 32].

<Table 1> Share of Display Market by Country(Based on amount)

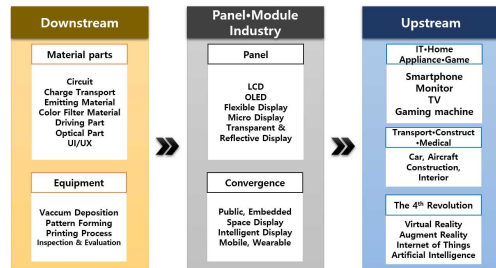
Country	2015	2016	2017	2018	2019
Republic of Korea	45.2	45.8	44.4	42.6	41.1
China	14.1	17.6	21.0	25.1	30.2
Taiwan	24.6	21.3	19.5	19.1	16.6
Japan	15.4	14.3	14.1	12.0	11.3
Etc.	0.6	0.9	0.9	1.2	0.9

우리나라는 2004년 이후 디스플레이 세계시장에서 점유율 1위를 지켰으나, 최근 5년간 중국과의 격차가 위협적으로 축소되고 있는 것을 확인할 수 있다. <Table 1>은 금액을 기준으로 국가별 디스플레이시장 점유율을 조사한 것이다. 이러한 현상의 이유로 추정되는 것은 디스플레이 산업 국가핵심기술을 대상으로 발생한 다수의 해외 유출사건이다.

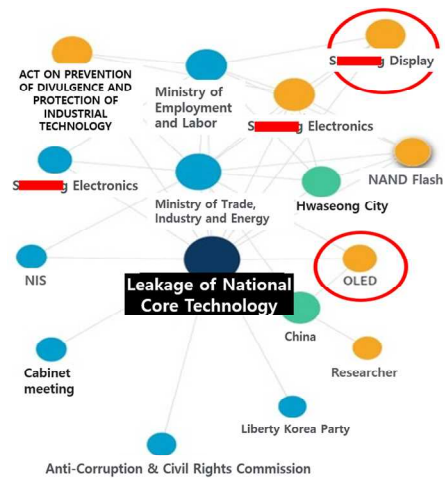
또한, <Figure 1>에서 뉴스토픽 분석 서비스인 빅카인즈에서 “국가핵심기술 유출”을 키워드로 최근 5년간(2016년 1월 1일~2021년 1월 1일)의 데이터를 분석 해본 결과 디스플레이 분야에 대한 키워드가 가장 많이 분석되었으며, OLED 기술 또한 분석된 키워드 중 상위에 위치한 것을 확인하였다.

OLED 기술에 대한 심각한 유출위협은 서론에서 언급한 국가핵심기술 범위의 모호성 문제로 인한 결과로 해석할 수 있다. 현행법에 따라 국가핵심기술 선정의 범위가 최소화되어 적용되고 있기 때문에 국가핵심기술을 지정해도 국가핵심기술과 연관성이 높은 해당 산업분야 내 중요 기술이 유출되어 산업의 경쟁력을 위협하고 있는 상황을 확인할 수 있다.

<Figure 2>는 디스플레이 산업의 전·후방 연관산업을 도식화 한 것이다. 디스플레이 산업



Source: Korea Trade Insurance Corporation[26].
<Figure 2> Downstream & Upstream of Display



<Figure 1> News Keyword Analysis

은 패널을 중심으로 전후방 연관효과가 큰 시스템 산업으로서 시장 선도를 위해 기술 경쟁력이 매우 중요한 산업이다. 후방 산업으로는 다양한 부품, 구동 재료산업이 연관되어있고, 전방 산업으로는 TV·휴대폰 등의 보편화 된 완제품으로 이루어졌다. 특히 전방산업의 전자기기들은 트렌드에 민감하여 시장 예측 및 선제적 기술 투자가 필요한 산업이다. 차량용 디스플레이, 스크린골프, 증강현실 및 가상현실 디스플레이, 게임분야 등에 사용되는 차세대 디스플레이에 대한 개발이 가속화되고 있어 디스플레이 산업의 중요성과 규모는 커질 전망이다[26].

2.2 문헌 별 관계망 구축 선행연구 분석

인터넷 환경이 확장되고 전자문서의 활용도가 높아짐에 따라 전자문서를 대량으로 수집한 비정형 데이터인 빅데이터에 대한 분석 및 연구가 활발히 진행되고 있다. 빅데이터 분석은 사회 널리 포진해 있는 데이터로부터 다양한 가치를 창출하는 과정으로 공공 서비스 개선, 민간 기업 생산성 향상등의 장점을 가지고 있다. 연구 분야에서는 텍스트/이미지/영상 등 다양한 종류의 빅데이터에 맞는 다양한 분석 방법을 적용하여 사회문제를 해결하고 있다.

그 중 텍스트 마이닝 분석은 넓은 학문 분야에서 다양한 주제를 대상으로 연구되고 있다. Matsuo and Ishizuka[34]는 통계적 분석 방법인 단어 동시출현 네트워크 분석(Co-Occurrence network Anlysis) 방법으로 키워드 추출 연구를 수행하였으며, 이성직[29]은 뉴스 데이터에 TF-IDF 분석 기법으로 주제 및 키워드를 추출 하였다.

현시점의 기술 동향을 파악하고 미래 유망 기술을 예측할 때에도 다양한 문헌을 수집하여 텍스트 마이닝 분석기법을 적용한다. 주로 사용되는 분석 기법은 키워드 분석과 네트워크 분석으로 나뉘며, 문헌별로 이를 적용한 연구를 분석해보았다. 먼저 국내 온라인 뉴스기사에 텍스트 마이닝 기법 중 토픽 모델링과 키워드 네트워크 분석을 사용한 노병준의 연구를 분석하였다. 이 연구에서는 토픽 모델링을 활용하여 뉴스로부터 대표 키워드를 추출하였고 키워드 네트워크 분석을 통해 동시출현 키워드 네트워크를 구성하였다. 이후 키워드 네트워크의 타임라인 변화를 분석하는 시계열 분석을 수행하였다. 이를 통해 구제역 발병에 따른 다양한 사회적 파급효과의 분석이 가능하였다[37].

국내 논문을 대상으로 키워드 네트워크와 토픽 모델링을 적용한 강지운의 연구에서는 2020년 10월까지 국내 학술등재지(KCI)에 수록된 관리자 코칭리더십 관련 논문 총 131편을 수집하였다. 수집된 논문의 초록을 대상으로 키워드 전처리 작업 후 총 1,142개의 키워드를 추출하였고, NetMiner Ver.4를 활용하여 키워드 빈도분석, 시기별 키워드 분석, 단어쌍의 동시출현 빈도분석을 수행하였다. 이를 통해서 국내 관리자 코칭리더십 연구 동향을 살펴볼 수 있었고, 향후 발전적 연구 수행 방향성을 모색할 수 있었다[17].

국내 보고서를 대상으로 빈도분석과 토픽 모델링, 연관성 분석을 사용한 임시영은, 공간정보 분야의 연구 동향을 파악을 위해 국가과학기술전자도서관(NDSL, National Discovery for Science Library)에서 데이터를 수집하였다. “공간정보”를 키워드로 1996년부터 2013년까지 발행된 논문 총 4,000개, 1994년부터 2013년까지 발간된 국가연구개발보고서 총 225개를 추출하였다[33].

국내 특허 데이터를 활용한 조은누리는 관련 기술의 동향을 지속적으로 분석하여 미흡한 부분을 파악해 5G 기반 산업 고도화를 도모하기 위해 특허 데이터에 소셜 네트워크 분석(SNA, Social Network Analysis) 기법을 활용하여 기술개발 동향 분석을 수행하였다. 소셜 네트워크 분석기법이란, 개별 정책 행위자 사이에 형성되어 있는 관계의 속성을 분석 대상으로 삼는 방법론으로, 이 분석기법의 목적은 구조나 네트워크 형태의 특징을 도출하고 관계성으로 체계의 특성을 설명하거나 체계를 구성하는 단위의 행위를 설명하는 것이다. 특히 특허 데이터의 초록을 대상으로 LDA 토픽 모델링을 수행하였을 때 5G 관련 핵심기술을 도출해 낼 수

있었다. 제조업 내 스마트공장의 도입 등으로 인해 통신장비의 진화가 빠르게 일어나고 있어 5G로 연결된 다양한 정보통신 기기들을 신속하고 종합적으로 관리할 수 있는 중요한 논거를 제시하였다[10].

2.3 국가핵심기술 관계망 구축을 위한 알고리즘 분석

비정형 데이터인 텍스트를 분석하여 문헌과 주제에 대해 핵심 내용을 뽑아내는 텍스트 마이닝(Text Mining)은, 데이터 마이닝(Data Mining)의 일종으로 정형/비정형의 텍스트 데이터들을 자연어 처리 방식과 문서처리 방법을 적용하여 정보를 추출 및 가공하는 기법이다. 텍스트의 기준은 일반 문서, 도서, 웹페이지, 블로그, 전자저널, 이메일, 특허까지 넓게 포함된다. 텍스트 마이닝 기법을 통하여 문서 요약, 문서 분류, 문서 군집, 특징 추출 등 수행이 가능하다.

본 연구에서는 키워드 분석 기법 중 하나인 TF-IDF와 네트워크 분석 기법 중 하나인 LDA 토픽 모델링을 활용해 다양항 문헌에 적용하는 파일럿 테스트를 수행하였다. 그리고 실루엣 지수로 연구 결과의 군집에 대한 신뢰성을 평가하였다.

2.3.1 TF-IDF

텍스트 마이닝의 분석 기법 중 키워드 분석에 포함되는 단어의 중요도 산정은 문서 내 등장 빈도, 문서 간 등장빈도, 총 등장 빈도 등 다양한 방법으로 측정이 가능하다. 그 중 TF-IDF(Term Frequency-Inverse Document Frequency)는 정보 검색과 텍스트 마이닝에서 활용하는 가중치 계산방법으로 어떤 문서에서의

특정 단어의 중요도를 측정할 수 있는 알고리즘으로 매우 널리 활용되고 있다. 아래 식에서 설명하듯이 단어 빈도와 역문서 빈도의 곱으로, 여러 문서에서 중요한 특정 단어를 추출한다[28, 39].

$$TF-IDF = tf(t,d) * \log \frac{D}{1+df(t)}$$

Source: Salton and Buckley[39].

2.3.2 LDA 토픽 모델링

토픽 모델링(Topic Modeling)은 언어 텍스트 집합을 가장 잘 표현하는 토픽(Topic) 또는 주제범주를 구분하는 기법이다[30]. 문서 집합을 구성하는 단어를 분석하여 주제를 추론하는 확률적 모형이며, 구조화되지 않은 방대한 문서 집단에서 잠재적인 주제를 찾아내고 대량 문서를 분류하는 기법으로 활발히 활용되고 있다.

본 연구에서는 토픽 모델링에서 가장 활발히 사용되는 LDA(Latent Dirichlet Allocation, 잠재 디리클레 할당) 토픽 모델링을 활용해 디스플레이 분야 국가핵심기술의 연관기술 추출 실험을 수행하였다. Griffiths and Steyvers의 연구에서 제시한 LDA 토픽 모델링의 파라미터 $\alpha = 50/\text{토픽 수}$, $\beta = 0.1$ 를 채택하였고[14], Chang, Edward Y.의 연구에서 밝혀진 대로 LDA 토픽 모델링의 추론기법 중 가장 대중적인 MCMC(Markov chain Monte Carlo)로 실험을 진행하였다[8].

2.3.3 실루엣 지수

문서를 군집화 할 때, 이상치(Outlier)로 인해 데이터의 분포가 왜곡될 가능성이 있다. 군집의 중심을 평균값으로 정하는 K-means 알고리즘을 활용하여 군집화 평가 척도로써 실루엣

(Silhouette) 지수를 계산 후 최적 군집 수를 결정한다. 실루엣 기법은 데이터들이 얼마나 잘 군집화(Clustering) 되었는지를 나타내는 것으로, -1과 1사이의 값을 가지고, 1에 가까울수록 데이터는 올바른 클러스터에 분류된 것이며, -1에 가까울수록 잘못된 클러스터에 분류되었음을 나타낸다[31].

Panichella의 연구에서는 군집화 평가 척도로 실루엣 지수를 활용하여 LDA 토픽 모델링의 최적의 토픽 수, α , β 등의 파라미터 수치를 추천하였다. 본 연구에서는 국가핵심기술의 관계망 구축 후에 그 신뢰성을 판별할 지표로 실루엣 지수를 사용하였다[38].

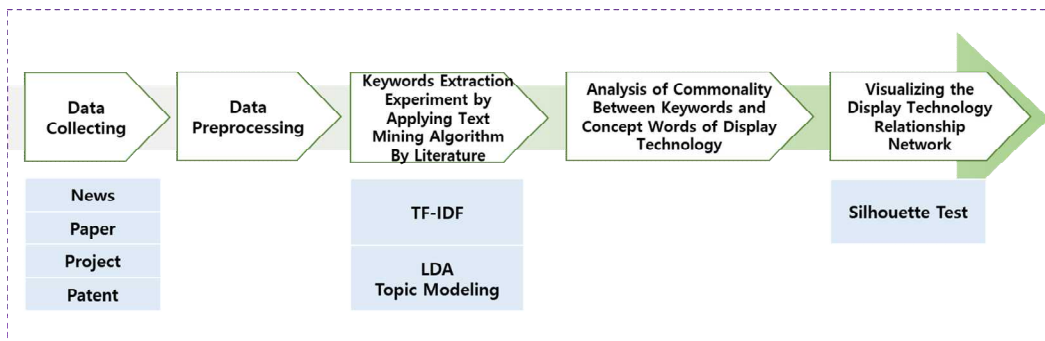
3. 연구방법론과 설계

선행연구를 분석하여 기술동향 분석에 적합한 문헌에서 데이터를 수집한 뒤 텍스트 마이닝 분석기법을 각각 적용하는 파일럿 테스트를 진행하여 국가핵심기술의 연관 중요 기술 분석을 수행하였다. 뉴스, 논문, 보고서, 특허의 4종류 문헌에 디스플레이 분야 국가핵심기술과 직접적인 연관이 있는 키워드인 “디스플레이”,

“TFT-LCD”, “AMOLED”의 단어로 데이터를 수집하였다. 우리 국가핵심기술에 대한 분석이므로 모두 국내 문헌을 수집하였다.

2019년 1월~2020년 12월까지 약 2년간 작성 및 발간된 자료를 수집하였으며, 특히 데이터는 출원일, 등록일을 기준으로 하였다. 또한 뉴스는 빅카인즈에서 중복치를 제거한 상태로 수집하였고, 논문은 한국학술인용지색인(KCI, Korea Citation Index), 보고서는 국가과학기술전자도서관(NDSL), 특허는 특허청 특허정보검색서비스(KIPRIS, Korea Intellectual Property Rights Information Service)에서 수집하여 각각 제목과 초록, 요약 분석에 사용하였다. 형태소 분석을 활용한 전처리를 포함한 모든 분석과정에 NetMiner를 활용하였다. 문헌별 특성을 파악하고자 전 처리는 일반명사(기술, 산업과 관계없는)를 대상으로 최소화하여 진행하였다. 한 음절 단어와 빈도 수 10 이하는 삭제하였다.

문헌들에 대해 TF-IDF와 LDA 토픽 모델링을 적용했고, 선행연구 분석 결과, LDA 토픽 모델링에 적용할 예측 알고리즘은 MCMC, 사용할 파라미터는 $\alpha = 50$ /토픽 수, $\beta = 0.1$ 를 채택하였다. 각 문헌에 대한 파일럿 테스트 이후 키워드(디스플레이, TFT-LCD, AMOLED)에 대한



〈Figure 3〉 Model of Research Methodology

〈Table 2〉 Amount of Collected Data

Document/Keyword	Display	TFT-LCD	AMOLED	Amount of each Document
News	1,000	156	397	1,553
Thesis	532	10	9	551
Project	248	2	7	257
Patent	1,500	1,023	413	2,936
Amount of Keyword	3,280	1,191	826	5,297

사전적 정의를 10개씩 수집하여 명사 추출 후 명사의 사용 빈도를 분석하여 개념 핵심어를 도출하였다. 이를 파일럿 테스트의 결과로 도출된 연관 기술들과 공통성 비교를 수행해 문헌과 분석 기법의 적합성을 검증하였다. 가장 높은 공통성을 보인 문서 종류-분석기법으로 관계망 구축을 수행하였으며, 관계망의 신뢰도 평가 척도로 실루엣 지수를 활용하였다.

〈Table 2〉는 문헌별 수집된 데이터의 데이터를 추출한 양을 정리한 것이다. 특히 데이터가 가장 많은 양이 수집되었고, 그 다음으로 뉴스 데이터, 논문, 보고서 순으로 데이터가 수집되었다. 이를 통해서 문헌별 작성 및 발간(출판)의 활성화 정도를 파악할 수 있다.

4. 연구결과

4.1 디스플레이 분야 국가핵심기술 개념 핵심어와 문헌 별 연관 기술 추출 실험 결과 간 공통성 분석

본 연구에서 지정한 키워드(디스플레이, TFT-LCD, AMOLED)에 대한 사전적 정의를 각각 10개씩 수집한 뒤 명사 단위로 분리 및 빈도 분석하여 국가핵심기술 개념 핵심어를 추출하였다. 〈Table 3〉의 “디스플레이(Display)”열에

있는 단어가 개념 핵심어로, 이와 문헌 별 실험 결과를 공통성 분석하였다. 단어 일치 여부에 따라 일치하지 않으면 X, 일치하면 O로 표시하였으며, O로 표시된 단어의 수에 따라 공통성 비율을 산정하였다.

〈Table 3〉에서 개념 핵심어와 문헌 별 실험 결과를 공통성 분석한 결과, 특허 데이터에 LDA 토픽 모델링을 적용한 결과와 개념 핵심어 간 공통성 분석 결과가 25%로 가장 높게 나타났다. 두 번째로 높은 공통성을 보여준 조합은 논문-TF-IDF로 10%의 공통성을 보였다.

〈Table 4〉는 TFT-LCD 키워드의 파일럿 테스트 결과와 개념 핵심어간 공통성을 분석한 결과이다. 가장 높은 공통성을 보인 문헌과 분석 알고리즘 기법의 조합은 특허-LDA 토픽 모델링으로, 30%의 공통성을 보였다. 두 번째로 높은 공통성을 가진 실험 결과는 보고서-TF-IDF(5%)와 보고서-LDA 토픽 모델링(5%)이다.

〈Table 5〉는 AMOLED 키워드의 파일럿 테스트 결과와 개념 핵심어의 공통성 분석 결과이다. 특허 데이터에 LDA 토픽 모델링을 적용한 파일럿 테스트 결과가 공통성 50%로 모든 테스트 결과 중 가장 높은 공통성을 보였다. 두 번째로 높은 공통성을 가진 실험 결과는 뉴스 데이터에-LDA 토픽 모델링 적용한 결과로 15%, 보고서 데이터에 LDA 토픽 모델링을 적용한 결과인 15%이다.

<Table 3> Analyze the Commonality Between the Concept Keywords and Experimental Results(Display)

No.	Display	News		Thesis		Project		Patent	
		TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling
1	Computer	X	X	X	X	X	X	X	X
2	Demonstration effect	X	X	X	X	X	X	X	X
3	Cathode Ray Tube	X	X	X	X	X	X	X	X
4	Data	X	X	X	X	X	X	X	O
5	Keyboard	X	X	X	X	X	X	X	X
6	Light	X	X	X	X	X	X	X	O
7	Image	X	X	X	X	X	X	X	O
8	Sound	X	X	X	X	X	X	X	X
9	Show window	X	X	X	X	X	X	X	X
10	Speed	X	X	O	X	X	X	X	X
11	Advertising	X	X	X	X	X	X	X	X
12	Color	X	X	O	X	X	X	X	X
13	Board	X	X	X	X	X	X	X	X
14	Display	X	X	X	X	X	X	X	O
15	Gestalt Psychology	X	X	X	X	X	X	X	X
16	Fluorescence	X	X	X	X	X	X	X	X
17	Panel	X	X	X	X	X	X	X	O
18	Tate	X	X	X	X	X	X	X	X
19	Attention	X	X	X	X	X	X	X	X
20	Modeling	X	X	X	X	X	X	X	X
Rate of Commonality		0%	0%	10%	0%	0%	0%	0%	25%

<Table 4> Analyze the Commonality Between the Concept Keywords and Experimental Results(TFT-LCD)

No.	Display	News		Thesis		Project		Patent	
		TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling
1	Transistor	X	X	X	X	X	X	X	X
2	Liquid Crystal	X	X	X	X	X	X	X	O
3	Glass	X	X	X	X	X	X	X	O
4	Pixel	X	X	X	X	X	X	X	X
5	Backlight	X	X	X	X	X	X	X	X
6	Substrate	X	X	X	X	O	O	X	O
7	Thin Film	X	X	X	X	X	X	X	O
8	Polarizer	X	X	X	X	X	X	X	X
9	Glass Substrate	X	X	X	X	X	X	X	O
10	Monitor	X	X	X	X	X	X	X	X
11	Light	X	X	X	X	X	X	X	X
12	Video	X	X	X	X	X	X	X	X
13	Display	X	X	X	X	X	X	X	O
14	Computer	X	X	X	X	X	X	X	X
15	Cathode Ray Tube	X	X	X	X	X	X	X	X
16	Thickness	X	X	X	X	X	X	X	X
17	Notebook	X	X	X	X	X	X	X	X
18	Resolution	X	X	X	X	X	X	X	X
19	Filter	X	X	X	X	X	X	X	X
20	Ultraviolet Ray	X	X	X	X	X	X	X	X
Rate of Commonality		0%	0%	0%	0%	5%	5%	0%	30%

<Table 5> Analyze the Commonality Between the Concept Keywords and Experimental Results(AMOLED)

No.	Display	News		Thesis		Project		Patent	
		TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling	TF-IDF	LDA Topic modeling
1	Active	X	X	X	X	X	X	X	O
2	Emitting	X	O	X	X	X	O	X	O
3	Light and Shade	X	X	X	X	X	X	X	X
4	Self-illumination	X	X	X	X	X	X	X	O
5	Unit	X	X	X	X	X	O	X	O
6	OLED	X	O	X	X	X	X	X	O
7	Light	X	X	X	X	X	X	X	X
8	Backlight	X	X	X	X	X	X	X	X
9	Glass	X	X	X	X	X	X	X	X
10	Pixel	X	X	O	X	X	X	X	O
11	Organic matter	X	X	X	X	X	O	X	X
12	Primary Color	X	X	X	X	X	X	X	X
13	Drive System	X	X	X	X	X	X	X	X
14	Fluorescence	X	X	X	X	X	X	X	O
15	Color	X	X	X	X	X	X	X	X
16	Afterimage	X	X	X	X	X	X	X	X
17	Liquid Crystal	X	X	X	X	X	X	X	X
18	Thickness	X	X	X	X	X	X	X	O
19	Diode	X	O	X	X	X	X	X	O
20	Curved	X	X	X	X	X	X	O	O
Rate of Commonality		0%	15%	5%	0%	0%	15%	5%	50%

<Table 6> Result of Analyze the Commonality

	TF-IDF			LDA Topic Modeling			
	Display	TFT-LCD	AMOLED	Display	TFT-LCD	AMOLED	
News	0%	0%	0%	0%	0%	15%	15%
Thesis	10%	0%	5%	0%	0%	0%	15%
Project	0%	5%	0%	0%	5%	15%	25%
Patent	0%	0%	5%	25%	30%	50%	110%
Total amount of Commonality Rate	10%	5%	10%	25%	35%	80%	-

<Table 6>은 위 공통성 결과를 정리한 것이다. 15% 이상(3개 단어 이상) 공통성을 보인 문헌 종류-분석기법은 뉴스-LDA(AMOLED), 보고서-LDA(AMOLED), 특허-LDA(디스플레이, TFT-LCD, AMOLED)의 5가지로 확인되었고, 문헌 별 공통성 비율 계의 순위는 ①특허, ②보고서, ③논문, ④뉴스로 나타났다. 이는 기술의 핵심어를 파악하기 가장 좋은 문헌

의 순위를 보여준다고 해석할 수 있다. 키워드 별 비율계의 순위는 AMOLED(90%), TFT-LCD(40%), 디스플레이(35%) 순으로 가장 핵심어가 많이 노출된 키워드의 순으로 볼 수 있다. 연구 결과를 통하여 디스플레이 분야 국가 핵심기술 관계망 구축으로는 특허 데이터와 LDA 토픽 모델링 기법이 가장 적합한 것으로 확인 되었다.

4.2 국가핵심기술 관계망 구축

특히 데이터에 LDA 토픽 모델링 기법을 적용한 방법이 국가핵심기술의 주요 연관 기술을 추출하는 데에 가장 적합하므로, 키워드 별로 관계망을 구축하였다. 토픽 명을 지정하는 방식은 토픽에 포함되는 단어들을 주제 일관성 순서대로 특허정보검색서비스(KIPRIS)에서 검색하여 상위 결과물의 기술 명으로 지정하였다. 특허정보검색서비스는, 국내외 지식재산권을 DB로 구축 및 제공하는 한국특허정보원의 서비스이다. 주제 일관성이란, 토픽에 모인 단어들의 주제(토픽 포함)일치 정도이다. 높을수록 해당 주제(토픽)에 대해 일관성이 높다는 것이다.

<Table 7>은 디스플레이 키워드의 토픽 별 하위 단어를 주제 일관성수치를 기준으로 나열한 것이며, 각 토픽을 설명하는 의미론적 결과로 해석된다. 액정표시(Liquid Crystal Display) 토픽은 패널(Panel), 출력(Print) 등 액정에 정보를 표현하는 단어로 구성되었고, 주제 일관성은 최대 0.072335~10위는 0.011535까지 분포

되어있다. 발광 소자(Emitting Diode) 토픽은 상기(Display Unit), 기판(Substrate), 소자(Diode), 프로브(Probe), 발광(Emitting) 등 디스플레이에 빛을 내기 위한 기술에서 발견되는 단어들이 주를 이뤘다. 주제 일관성은 0.218819~10위는 0.007431까지 분포되었다. 수지 조성물 및 헤드업(Resin Composition & Head up Display) 토픽에는 화면(Display), 대응(Confrontation), 조명(Light), 수지(Resin Composition), 조성물(Composition) 등의 단어로 이루어졌다. 주제 일관성은 0.01618~10위 0.010663으로 이루어졌다.

<Figure 4>는 디스플레이 키워드 토픽의 결과를 시각화 한 것으로, 세 개의 토픽 모두 “모듈(Module)”이라는 부품 단어에서 공통성을 가지며, 발광소자와 수지 조성물 및 헤드업 토픽의 연관 정도를 보았을 때 “필름(Film)”, “도파관(Waveguide)”, “유기(물)(Organic)”이라는 공통성이 있어 이 부분에 대한 연관된 보호가 필요할 것으로 분석된다. 실루엣 지수는 0.894로 1에 가까운 수치를 보여 군집의 신뢰성이 확보되었다.

<Table 7> Topic Result(Keyword: Display)

No.	Liquid Crystal Display	Coherence	Emitting Diode	Coherence	Resin Composition & Head up Display	Coherence
1	Panel	0.072355	Display Unit	0.218819	Display	0.01618
2	Print	0.02833	Substrate	0.03845	Confrontation	0.015609
3	Data	0.023585	Diode	0.0179	Light	0.013896
4	Image	0.020241	Probe	0.016645	Resin Composition	0.013613
5	Signal	0.019986	Emitting	0.015108	Composition	0.012885
6	Display Unit	0.018409	Connector	0.012984	Optics	0.012394
7	Port	0.01705	Curved Display	0.011065	Housing	0.012347
8	Touch	0.016938	Pixel	0.009101	User	0.011737
9	Frame	0.015831	Module	0.008052	Head up	0.011186
10	Infrared Ray	0.011535	Anode	0.007431	Status	0.010663

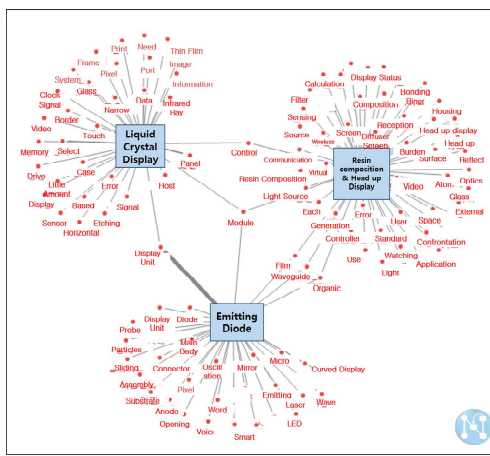
<Table 8> Topic Result(Keyword: TFT-LCD)

No.	OLED	Coherence	Micro LED	Coherence	Semiconductor Glass Substrate	Coherence
1	Display	0.095093	Display Unit	0.139538	Panel	0.093525
2	Panel	0.047684	Including	0.034141	Formation	0.040594
3	Device	0.042509	Information	0.027333	Glass Substrate	0.035253
4	Film	0.03111	Location	0.015418	Electrode	0.031536
5	Liquid Crystal	0.029746	Circuit	0.014558	Display Unit	0.02995
6	Emitting	0.014012	Micro	0.012841	Including	0.027427
7	Absorption	0.011987	Security	0.012266	Phase	0.02723
8	Thin Film	0.010193	Signal	0.011959	Semiconductor	0.022717
9	Unit	0.010038	Display	0.011673	Composition	0.020887
10	Flexible	0.009791	Block	0.010207	Layout	0.018164

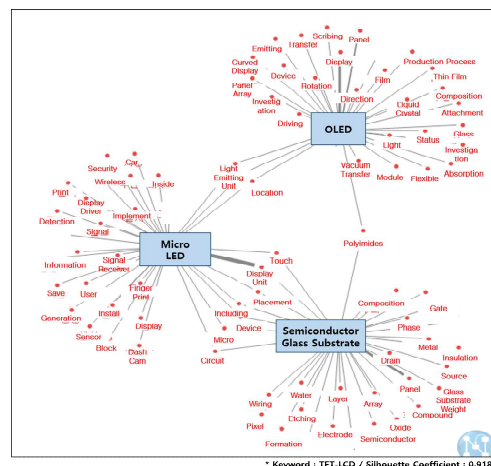
<Table 8>은 TFT-LCD 키워드의 LDA 토픽 모델링 분석 결과이다. 토픽명은 OLED와 마이크로 LED(Micro LED), 반도체 유리기판(Semiconductor Glass Substrate)으로 결정되었다. <Figure 2>에서 정리하였듯이 OLED와 마이크로 LED는 모두 디스플레이의 패널 산업에 해당되므로 두 기술 간의 연관성을 더욱 명확히 할 수 있다.

<Figure 5>는 토픽 모델링의 결과로 연관기술 관계망을 구축한 것이다. OLED 토픽과 마이크로

LED 토픽은 “위치(Location)”, “발광 유닛(Light Emitting Unit)” 단어에서 공통성이 있고 OLED와 반도체 유리기판 간에는 “폴리이미드(Polyimides)”라는 절연 특성과 높은 내열성을 가진 물질이 공통성으로 분석되었다. 마이크로 LED와 반도체 유리기판 토픽 간에 가장 많은 공통 기술(단어)들을 관찰할 수 있으며 그 단어 들로는 “상기(Display Unit)”, “터치(Touch)” 등이 있다. 실루엣 지수는 0.918로 1에 가까운 수치를 보여 군집의 신뢰성이 확보되었다.



<Figure 4> Related Technology Network(Keyword: Display)



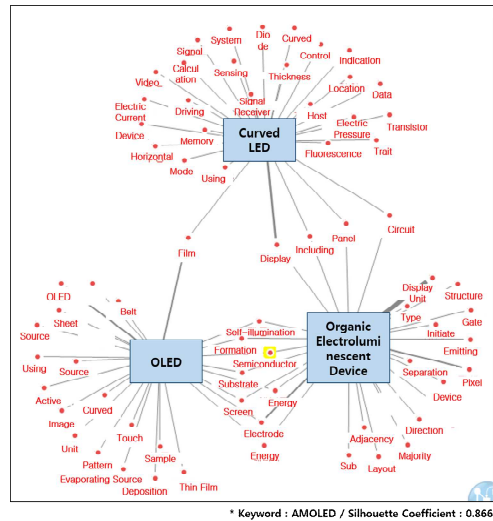
<Figure 5> Related Technology Network(Keyword: TFT-LCD)

<Table 9> Topic Result(Keyword: AMOLED)

No.	OLED	Coherence	Curved LED	Coherence	Organic Electroluminescent Device	Coherence
1	OLED	0.05975	Display	0.084821	Display Unit	0.146199
2	Film	0.050525	Including	0.042347	Pixel	0.055032
3	Formation	0.025085	Data	0.036123	Electrode	0.054644
4	Unit	0.024509	Device	0.027421	Substrate	0.026363
5	Substrate	0.024026	Panel	0.021542	Majority	0.025733
6	Pattern	0.022611	Film	0.019624	Emitting	0.021948
7	Thin Film	0.016527	Driving	0.01809	Including	0.020021
8	Electrode	0.014046	Curved	0.017617	Circuit	0.016669
9	Active	0.013688	Fluorescence	0.016867	Energy	0.016266
10	Touch	0.01194	Thickness	0.016351	Data	0.014197

<Table 9>는 AMOLED 키워드의 LDA 토픽 모델링 결과이다. 토픽 명은 특허정보검색 서비스에서 각 토픽의 하위단어를 검색하여 지정하였다. 그 결과 OLED, 곡면 LED(Curved LED), 유기발광소자(Organic Electroluminescent Device)로 토픽 명이 선정되었다. AMOLED 키워드에 OLED와 곡면LED가 토픽으로 선정된 것은 세 기술의 연관성을 보여주는 결과로 해석된다. 특히 OLED와 유기발광소자 토픽 간에 8개의 단어가 연결되어 깊은 연관성을 보이는 것은 기술간 특성을 잘 반영한 결과이다.

<Figure 6>에서 볼 수 있듯, OLED 토픽은 유기발광소자 토픽과 가장 높은 공통성을 보였으며 “에너지”, “스크린”, “진극”, “기관”, “반도체” 등의 단어에서 그 연결성을 확인하였다. 또한 유기발광소자 토픽은 곡면 LED 토픽과도 비교적 높은 연관성을 보였다. OLED 토픽과 곡면 LED 토픽은 “필름” 단어에서만 공통성을 보였다. 실루엣 지수는 0.866으로 1에 가깝기 때문에 군집의 신뢰성은 확보되었다.



<Figure 6> Related Technology Network(Keyword: AMOLED)

5. 결론 및 향후연구

국가 경제에 미치는 제조업의 영향이 커질수록 핵심기술의 관리방법과 효율적인 보안을 위한 창의적인 연구가 필요하다. 국가핵심기술의 특성 상 정보 공개가 제한적이고 설정범위가

좁기 때문에 관리 및 보호에 한계가 있어 유출 사고가 끊이지 않고 있다. 또한 기술 자체의 특성 상 그 연관범위를 산정하는 것에 명확한 기준이 부재하기 때문에 국가핵심기술이 가진 기술적 특성과 관련 법률로 인한 특성을 고려하여 연관범위를 산정하는 연구가 필요한 상황이다.

본 연구에서는 디스플레이 분야 국가핵심기술(디스플레이, TFT-LCD, AMOLED)을 키워드로 4가지 문헌 종류에서 텍스트 데이터(뉴스, 논문, 보고서, 특허) 수집하여 텍스트 마이닝 분석기법인 TF-IDF와 LDA 토픽 모델링 분석기법을 적용하는 파일럿 테스트를 진행 하였다. 테스트 결과의 신뢰도 검증을 위하여 각 키워드에 대한 사전적 정의를 분석 후 개념 핵심어를 추출하여 공통성 분석을 진행하였다.

본 연구 결과로 특허 데이터에 LDA 토픽 모델링을 적용한 결과가 디스플레이 분야 국가핵심기술과 관련있는 중요기술 관계망 구축에 적합한 것으로 확인되었으며, 디스플레이 키워드에서는 “액정 표시”, “발광 소자”, “수지 조성물 및 헤드업”이, TFT-LCD 키워드에서는 “OLED”, “마이크로 LED”, “반도체 유리기관”이, AMOLED 키워드에서는 “OLED”, “곡면 LED”, “유기발광소자”가 각각 토픽으로 추출 되었다. 도출된 중요기술을 관계망으로 구축하여 기술간 연관성까지 구체적으로 확인 할 수 있었다. 관계망을 통해 시각화가 이루어졌기 때문에 단어만 나열한 것 보다 연관기술의 범위 설정에 더욱 효과적이다.

특히 TFT-LCD와 AMOLED 키워드의 연구 결과에 “OLED” 토픽이 추출된 것은 세 기술간의 상호 연결성을 확인해주는 결과이다. 이를 근거로 OLED 기술에 대해 경쟁국으로부터 심각한 유출위협이 발생한 이유를 유추할 수

있으며, OLED기술에 대한 보호로 국가핵심기술을 보호하는 효과를 볼 수 있다.

본 연구의 결과를 통하여 국가핵심기술과 연관된 중요기술을 조기에 선별하여 적재적소에 보호조치를 수행하는 등 국가핵심기술이 가지는 제한적인 정보 공개 특성과 기술 범위의 모호성을 극복할 수 있다. 더 나아가 신기술에 대한 전문적 정보를 가장 빨리 얻을 수 있는 특허를 활용한 분석방법을 제시함으로써 새롭게 특허에 출원되는 기술 중 국가핵심기술과 연관성 있는 것을 판별할 수 있다. 그 결과 연구 및 개발자들에게 새로운 학술적 영감을 줄 수 있고 제조 현장에는 효과적인 보안체계를 적용할 수 있다.

연구의 한계점은 다음과 같다. 첫째, 국내 문헌만을 분석 자료로 채택하여 국내외 사례에 적용하기 어렵다는 점이다. 향후연구에서는 해외문헌을 중점으로 분석하거나, 국내 문헌과 해외문헌을 모두 분석하여 연구의 범위를 넓힐 수 있다. 둘째, 네 종류의 문헌 수집 수가 일정하지 않고 문헌의 종류에 따라 추출 가능한 단어의 편차가 존재하였던 점이 한계이며 향후 연구에서는 자료 수집에 대한 기준을 마련하여 보완하거나 국가핵심기술 분석을 위한 자료수집 기준 자체에 대한 연구를 수행할 수 있다. 셋째, 특허에서는 핵심어 추출만 나올 수 있어 향후 연구에서는 이러한 결과를 실무에 직접적으로 활용할 수 있는 방안(용도)에 대해서 고민할 필요가 있다. 넷째, 기술의 전문적인 체계를 파악할 수 있는 단어가 추출되지 않았기 때문에 향후 연구에서는 국가핵심기술과 밀접한 기술들의 체계와 연관성을 파악할 수 있도록 연구 방향을 제시하여 보완할 수 있다. 다섯째, 본 연구는 국가핵심기술로 지정된 분야들 중 디스플레이 기술만을 대상으로 연구를 진행한

되었다는 것이 한계이다. 향후 연구에서는 다른 산업 분야를 대상으로 기술의 핵심어를 추출하고 해당 핵심어간 연관성 네트워크를 분석하여 학문적, 실무적 시사점을 제시할 수 있을 것이다.

References

- [1] Act on Prevention of Divulgence and Protection of Industrial Technology, Statutes of Republic Korea, Published Oct. 27, 2006.
- [2] Beattie, R. S., Kim, S., Hagen, M. S., Egan, T. M., Ellinger, A. D., and Hamlin, R. G., "Managerial coaching: A review of the empirical literature and development of a model to guide future practice," *Advances in Developing Human Resources*, Vol. 16, No. 2, pp. 184-201, 2014.
- [3] Blei, D. M., Ng, A. Y., and Jordan, M. I., "Latent dirichlet allocation," *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022, 2003.
- [4] Blei, D.M. and Lafferty, J.D., "Topic models," *Text Mining: Classification, Clustering, and Applications*, Vol. 10, No. 71, pp. 34, 2009.
- [5] Blei, D.M., "Probabilistic topic models," *Communications of the ACM*, Vol. 55, No. 4, pp. 77-84. 2012.
- [6] Braam, R. R., Moed, H. F., and Van Raan, A. F., "Mapping of science by combined co-citation and word analysis," *Journal of the American Society for Information science*, Vol. 42, No. 4, pp. 233-251, 1991.
- [7] Byeon, J.K., "[Exclusive] Only 36% of Companies Which Have National Core Technology Operating Their Own Security team," *Dong-A Ilbo*, 2020.
- [8] Chang, E. Y., "Foundations of Large-Scale Multimedia Information Management and retrieval: mathematics of perception," *Springer Science & Business Media*, 2011.
- [9] Chang, S. W., "Oracle Bigdata," *Oracle Korea*, 2013.
- [10] Cho, E. and Chang T. W., "Patent Analysis on 5G Technology Trends from the Perspective of Smart Factory," *The Journal of Society for e-Business Studies*, Vol. 25, No. 3, pp. 95-108, 2020.
- [11] Choi, E. R., Song, B. G., and Rhee, Y. I., "Analysis of Protection of National Core Technologies Using the Delphi Method," *Korean Journal of Public Safety and Criminal Justice(KJPC)*, Vol. 21, No. 2, pp. 305-340, 2012.
- [12] Choi, H. C., Kwahk, K. Y., and Kim, N. G., "Discovering Promising Convergence Technologies Using Network Analysis of Maturity and Dependency of Technology," *Journal of Intelligence and Information Systems(JIIS)*, Vol. 24, No. 1, pp. 101-124. 2018.
- [13] Choi, J. M., Kim, Y. N., and Park, D. C., "A Study on the Desirable Direction for Korean Community Gardens-Using Case Studies in Japan and Questionnaire Survey of Korean Experts-," *Journal of The*

- Residential Environment Institute of Korea, Vol. 16, No. 1, pp. 83-104, 2018.
- [14] Griffiths, T. L. and Steyvers, M., "Finding scientific topics," Proceedings of the National Academy of Sciences, Vol. 101, pp. 5228-5235, 2004.
- [15] Herh, M., "21 Key National Technologies Leaked over Past 6 Years," BusinessKorea, 2017.
- [16] Jeong, J. H., "Big data Policy Tasks in the Era of the 4th Industrial revolution," Isyu wa nonjeom No. 1482, National Assembly Research Service, 2018.
- [17] Kang, J. W., "Research Trends in Managerial Coaching Leadership in Korea Using Topic Modeling and Keyword Network Analysis," Andragogy Today: Interdisciplinary Journal of Adult and Continuing Education(IJACE), Vol. 23, No. 4, pp. 181-210, 2020.
- [18] Kim, D. H., Kim, H. H., Kim, D. G., and Jo, J. N., "Social network analysis of keyword community network in IoT patent data," The Korean Journal of Applied Statistics, Vol. 29, No. 4, pp. 719-728, 2016.
- [19] Kim, J. C., Lee, J. H., Park, S. S., and Jang, D.S., "Technology Strategy based on Patent analysis," Journal of Korean Institute of Intelligent Systems, Vol. 26, No. 2, pp. 141-146, 2016.
- [20] Kim, J.S., "Emotion Prediction of Paragraph using Big Data Analysis," Journal of Digital Convergence, Vol. 14, No. 11, pp. 267-273, 2016.
- [21] Kim, M. A. and Suh, C. K., "Keyword Network Analysis for Patent Information : Using SCM Patent Data," The Journal of Internet Electronic Commerce Research, Vol. 17, No. 3, pp. 159-175, 2017.
- [22] Kim, S. B., "An Exploratory Study of VR Technology using Patents and News Articles," Journal of Digital Convergence, Vol. 16, No. 11, pp. 185-199, 2018.
- [23] Kim, S. G., Cho, H. J., and Kang, J. Y., "The Status of Using Text Mining in Academic Research and Analysis Methods," The Journal of Information Technology and Architecture, Vol. 13, No. 2, pp. 317-329, 2016.
- [24] Kim, T. K., Choi, H. R., and Lee, H. C., "A Study on the Research Trends in Fin-tech using Topic Modeling," Journal of Korea Academia-Industrial Cooperation Society(JKAIS), Vol. 17, No. 11, pp. 670-681, 2016.
- [25] Korea Institute of Science and Technology Evaluation and Planning, "A Study on the Alignments Analysis Between Strategy and Budget of the Government R&D's Major Sectors," 2019.
- [26] Korea Trade Insurance Corporation, "Display Industry Technology · Market Trend -Form Factor Innovation and OLED Competition, Industry Trend Report," 2019.
- [27] Lee, D. H., Choi, H. Y., Jeong, B. K., and Yoon, J. H., "Monitoring Bio-fuel Technology Using Patent Text Mining," The Journal of Intellectual Property, Vol. 13, No.

- 1, pp. 285-312, 2018.
- [28] Lee, J. H., Lee, M. B., and Kim, J. W., "A Study on Korean language processing using TF-IDF," *The Journal of Information Systems*, Vol. 28, No. 3, pp. 105-121, 2019.
- [29] Lee, S. J. and Kim, H. J., "Keyword Extraction from News Corpus using Modified TF-IDF," *The Journal of Society for e-Business Studies*, Vol. 14, No. 4, pp. 59-73, 2009.
- [30] Lee, S. S., "A Study on the Application of Topic Modeling for the Book Report Text," *Journal of Korean Library and Information Science Society(JKLISS)*, Vol. 47, No. 4, pp. 1-18, 2016.
- [31] Lim, H. U., Cho, S. S., Quan, J., Kim, J. J., Lee, C. U., and Lee, G. M., "Prediction of Aviation Security Technologies through Cluster Analysis of Patent Big Data," *Journal of the Aviation Management Society of Korea*, Vol. 15, No. 5, pp. 43-63, 2017.
- [32] Lim, J. H., "Environmental change and development plan of the display industry," Bank of Korea, 2020.
- [33] Lim, S. Y., Lee, M. S., Jin, K. H., and Shin, D. B., "A Study on the Research Trends in the Area of Geospatial-Information Using Text-mining Technique Focused on National R&D Reports and Theses," *Spatial Information Research*, Vol. 22, No. 4, pp. 11-20, 2014.
- [34] Matsuo, Y. and Ishizuka, M., "Keyword extraction from a single document using word co-occurrence statistical information," *International Journal on Artificial Intelligence Tools*, Vol. 13, No. 1, pp. 157-169, 2003.
- [35] Ministry of Trade, Industry and Energy, "[Amendments] Public Notice for Designate of National Core Technology," 2020.
- [36] National Disaster Management Institute. "Development of Foresight Method for Future Disaster Through the Analysis of Complex Foresight Methodology," 2013.
- [37] Noh, B. J., Xu Z. S., Lee, J. U., Park, D. H., and Chung, Y. H., "Keyword Network Based Repercussion Effect Analysis of Foot-and-Mouth Disease Using Online News," *Journal of Korean Institute of Information Technology(JKIIT)*, Vol. 14, No. 9, pp. 143-152, 2016.
- [38] Panichella, A., Dit, B., Oliveto, R., Di Penta, M., Poshynanyk, D., and De Lucia, A., "How to effectively use topic models for software engineering tasks? an approach based on genetic algorithms," *International Conference on Software Engineering (ICSE)*, pp. 522-531, 2013.
- [39] Salton, G. and Buckley, C., "Term-weighting approaches in automatic text retrieval," *Information Processing & Management*, Vol. 24, No. 5, pp. 513-523, 1988.
- [40] Steyvers, M. and Griffiths, T., "Probabilistic topic models," *Handbook of Latent Semantic Analysis*, Vol. 427, No. 7, pp. 424-440, 2007.
- [41] Won, J. Y. and Kim, D. G., "Deduction

- of Social Risk Issues Using Text Mining,” *Crisisonomy*, Vol. 10, No. 7, pp. 33-52, 2014.
- [42] Yoo, L. N. and Hwang, S. C., “A Trend Analysis of Agricultural and Food Marketing Studies Using Text-mining Technique,” *Journal of Korea Academia-Industrial Cooperation Society(JKAIS)*, Vol. 18, No. 10, pp. 215-226, 2017.
- [43] Yun, N. S. and Ji, I. Y., “An Analysis of Patent Co-Classification Network for Exploring Core Technologies of Firms: An Application to the Foldable Display Sector,” *Journal of Korea Academia-Industrial cooperation Society(JKAIS)*, Vol. 20, No. 4, pp. 382-390, 2019.
- [44] Yun, S. Y. and Han, K. S., “A Study on Patent Data Analysis and Competitive Advantage Strategy using TF-IDF and Network Analysis,” *Journal of Digital Contents Society*, Vol. 19, No. 3, pp. 529-535, 2018.

저 자 소 개



박세희
2021년
2021년~현재
관심분야

(E-mail: saeh95@cau.ac.kr)
중앙대학교 융합보안학과 (석사 졸업)
중앙대학교 융합보안학과 (박사과정)
보안 데이터 분석, 산업융합보안



윤원석
2021년
2021년~현재
관심분야

(E-mail: i3629i@cau.ac.kr)
성결대학교 컴퓨터공학과 (졸업)
중앙대학교 융합보안학과 (석사과정)
블록체인서비스, 빅데이터 보안, 산업융합보안



장항배
2006년
2007년~2012년
2012년~2013년
2014년~현재
관심분야

(E-mail: hbchang@cau.ac.kr)
연세대학교 정보시스템관리 (박사)
대진대학교 경영학과 조교수
상명대학교 경영학과 조교수
중앙대학교 산업보안학과 교수
보안 데이터 분석, 블록체인서비스, 산업융합보안