

미래 유망기술의 Weak Signal 탐지 방안

Detecting Weak Signals of Emerging Technologies

최수길 (S.G. Choi) 미래기술전략연구 1실 선임연구원
김기영 (K.Y. Kim) 미래기술전략연구 1실 실장
오진태 (J.T. Oh) 미래기술연구부 부장

지능정보사회를 대비한
이머징 기술이슈 특집

- I. 서론
- II. 미래 유망기술 개념 및 발굴 방법론
- III. Weak Signal의 개념 및 연구동향
- IV. Weak Signal 탐지를 위한 정량적 데이터 분석방법
- V. 결론

국가 경제의 지속적인 발전이 가능하게 하려면 미래 유망기술의 발굴이 중요함은 널리 알려져 있다. 하지만, 과학기술의 융·복합화와 불확실성의 증가 그리고 기존 기술의 틀을 깨는 혁신적 기술의 등장 등으로 인해서 실효성 높은 미래 유망기술의 발굴이 어려워지고 있다. 미래 유망기술 선점을 위해서는 fast follower 전략에서 first mover 형으로 전환이 필요하며, 이를 위해서는 기술태동 초기의 weak signal을 파악해야 한다. 지금까지 weak signal 탐지는 주로 전문가 기반으로 이루어져 왔으나 분석 시간이 오래 걸릴 뿐만 아니라 대상 분야가 넓어지고 분석할 데이터의 양이 급격히 증가하면서 정량적 데이터 분석을 보완적으로 사용하는 방향으로 패러다임의 변화가 일어나고 있다. 하지만, 텍스트 분석을 통한 weak signal 탐지기술은 아직 기초적인 수준에 머물러 있어서 관련 연구에 대한 투자가 필요하다.

I. 서론

국가 차원에서 미래사회의 성장동력을 발굴하고 이를 실현하여 경제·사회의 지속적인 발전이 가능하도록 하는 것이 중요하다. 기업 차원에서도 미래 기술 경쟁력을 확보하기 위하여 유망기술 발굴을 통한 효율적인 전략 수립이 필요하다. 이러한 중요성은 전 세계적으로 인식되고 있으며 EU Horizon 2020, 미국의 Defense Advanced Research Projects Agency(DARPA), 일본 New Energy and Industrial Technology Development Organization(NEDO) 등에서 미래 유망기술을 발굴하고 이에 대한 연구를 지원하는 프로세스를 정립하고 수행하고 있다. 유망기술의 발굴, 연구자 선정 및 지원, 산업화에 이르기까지 전체적인 프로세스가 구축되어야 하지만, 본고에서는 새롭게 부상하는 유망기술의 발굴에 집중한다.

미래 유망기술 발굴 방법은 [1]에 따르면 미래사회 니즈 분석 기법, 미래예측 기법, 설문조사, 전문가 의견활용, 데이터 분석방법 등으로 분류할 수 있다. 이러한 방법들 중에서 설문조사와 전문가의견 등이 주로 사용됐으나, 과학기술의 융·복합화와 불확실성의 증가 등으로 인하여 이러한 방법을 통한 실효성 있는 유망기술의 발굴에 어려움이 증가하고 있다. [2]에 따르면 미래란 과거의 여러 과정과는 전혀 다른 예측할 수 없었던 새로운 행동양식들이 불연속적인 변환이나 단절을 통해 나타나는 경향이 있다고 하며, 이러한 불연속적인 변화는 탐지하기가 더욱 어렵다.

이러한 변화가 어느 정도 구체화되고 선진 연구자들에 의해 가시적인 결과가 나온 후 fast follower 전략을 통해 국가 경쟁력을 유지하는 것이 이전에는 가능했지만, 이제는 first mover만이 시장 주도권을 가질 수 있게 되었다. 따라서 변혁적 기술의 태동초기에 나타나는 징후나 조짐을 파악할 수 있어야 하며, 이렇게 단초가 되는 정보를 weak signal이라고 표현한다. 지금까지

weak signal의 파악은 주로 높은 수준의 전문성을 가진 전문가 기반으로 이루어져 왔으나, 기술의 발전속도가 빨라지고 기술의 범위가 넓어지면서 전문가 기반의 분석이 점점 어려워지고 있다. 따라서, weak signal 탐지 과정을 좀 더 객관적으로 지원할 수 있는 정량적 데이터 분석방법이 필요하다. Weak signal 탐지를 위한 기존의 정량적 분석방법을 조사해본 결과 전문가 검토를 위한 후보 정보의 일부로 사용되는 기초적인 수준이므로 weak signal 상시 모니터링 체계 구축을 위해서는 관련 연구에 집중 투자할 필요가 있다.

II. 미래 유망기술 개념 및 발굴 방법론

1. 유망기술의 개념

유망기술의 추상적인 개념에 대한 공감대는 어느 정도 형성되어 있지만, 구체적인 정의는 연구목적에 따라 다양하다. 이러한 개념을 정리하고자 하는 시도가 이루어지고 있으며, [3]에서는 유망기술의 개념을 다음과 같이 정의하고 있다. 유망기술의 개념에서 가장 먼저 언급되어야 할 것은 ‘최근 들어 급격한 성장을 보이는 영역’이라는 것이다. 이는 신규 연구를 가능하는 잣대로 사용될 수 있을 뿐만 아니라 최근 들어 부상하는 연구를 파악하는 데 유용한 개념이다. 유망기술은 기존과 비교하여 다른 분야로의 전이나 변화를 의미한다. 이러한 전이나 변화는 동태적인 것으로서 변화가 멈추었다면 더는 유망기술로 보기 어렵다. 따라서 특정 시점에서의 유망기술은 전이 또는 변화하는 과정에 있어야 한다. 이러한 전이나 변화는 점진적일 수도 있으며 급진적일 수도 있다.

[4]에 따르면 유망기술에 대한 명명은 어떤 관점에서 보느냐에 따라 다음과 같이 다양하게 정의될 수 있다고 한다. 미래기술(future technology), 잠재성이 있는 기술(promising technology), 신흥 기술(emerging technology), 신기술(new technology), 돌파 기술

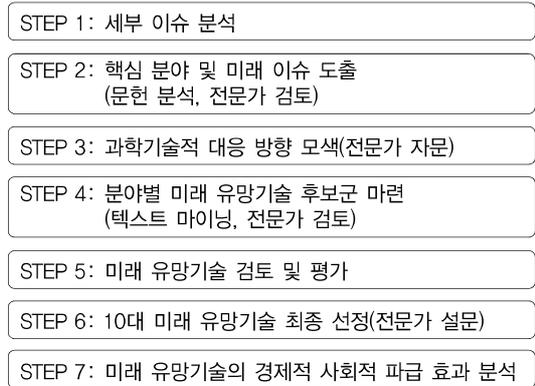
(breakthrough technology), 그리고 핵심 기술(key technology) 등으로 정의될 수 있다. 유망기술에 대한 정의는 해석하는 시각에 따라서 달라질 수 있는데, 기술 개발의 주체가 민간부문이면 사업화로 추진하여 수익을 낼 수 있는지가 기준이 되고 공공부문의 경우는 기술이 상대적으로 다른 국가보다 비교우위를 점할 수 있을 뿐만 아니라 현재 자국의 기술 경쟁력을 향상시켜줄 수 있는지가 기준이 된다.

세계 여러 기관에서 매년 미래유망기술을 선정하여 발표하고 있으며 대표적인 기관의 유망기술 선정 사례와 선정 기준이 [5]에 간략히 정리되어 있다.

2. 유망기술 발굴 방법

[1]에 따르면 미래 유망기술 발굴 방법은 <표 1>과 같이 분류할 수 있다. 이러한 기법들을 이용하여 유망기술 후보군이 도출되면, 평가 기준에 따라 후보기술의 특성을 분석하고 우선순위를 평가하여 최종 유망기술을 도출하게 된다. 이 과정에서 analytic hierarchy process와 quality function deployment 등의 의사결정 지원기법과 index analysis 등 계량적 데이터 분석기법, 그리고 패널 토론 등의 전문가 활용법 등의 기술평가 기법이 활용될 수 있다.

실제 유망기술 발굴 프로세스를 보면 목적이나 상황에 맞게 하나 혹은 그 이상의 기법을 조합하여 사용하는 것이 보통이다. 예를 들어, KISTEP의 미래 유망기술 선



(그림 1) KISTEP 미래 유망기술 선정 절차[5]

정 절차는 (그림 1)과 같이 미래 니즈 분석, 전문가 의견, 데이터 분석 등이 복합적으로 사용된다[5]. KISTEP 유망기술 선정방법의 매년 변화과정을 분석해보면 주목할 만한 점은 뉴스와 웹데이터 활용 등 데이터 분석에 기반을 둔 방법의 비중이 증가하고 있다는 것이다.

III. Weak Signal의 개념 및 연구동향

1. Weak Signal의 의미

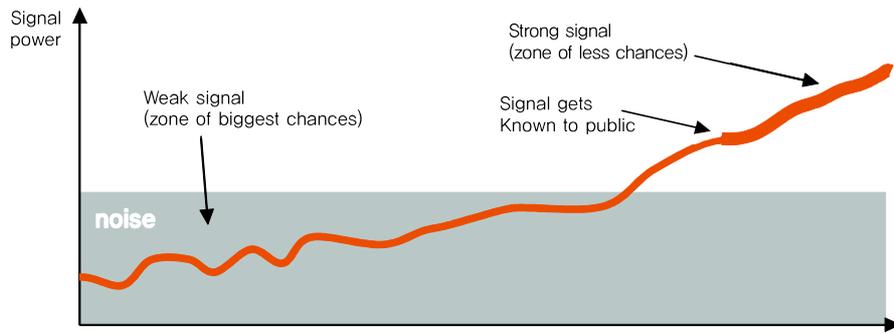
미래 유망기술을 발굴한다는 것은 과학기술에서 미래를 예측한다는 것이고, 결국 미래 기술 변화의 징후(signal)를 탐지하는 것이다. [6]에 따르면 미래 변화의 징후는 weak signal, strong signal 그리고 wild card의 관계에서 설명할 수 있다. Weak signal은 미약하고 중요하지 않은 것처럼 보이지만 미래에 일어날 일들에 대한 정보를 담고 있는 징후이다. Weak signal은 대개 noise와 섞여 있어서 구분하기 힘들지만, noise로 감춰지기 어려운 정도로 구체성을 띄게 되면 strong signal이 된다. Wild card는 예측하기 어려웠던 중대 사건의 발생을 의미하며 weak signal이 의미하던 현상이 현실로 나타나는 것이고, wild card 이후에는 기회가 거의 없거나 아주 적어진다고 볼 수 있다.

(그림 3)에 따르면 strong signal이 나타난 후 wild card까지의 시간이 짧으므로 대처할 시간이 부족하다.

<표 1> 유망기술 발굴 방법론[1]

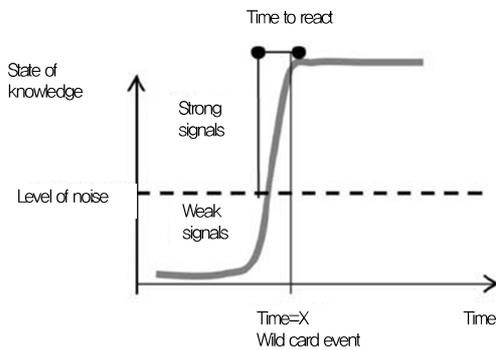
방법	세부 기술
Future needs	Needs analysis
Forecasting techniques	Scenario, Cross-impact analysis, Trend analysis
Experts	Expert panel, Delphi
General public	Survey
Technological data	Patent analysis, Keyword analysis, Citation analysis

<출처>: 한국통신학회논문지, “2020년 미래 무선통신 유망기술 발굴,” 2013. 01.

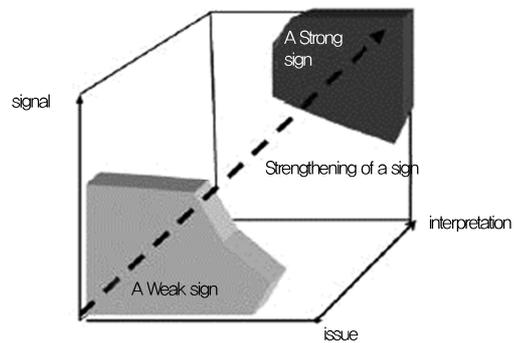


(그림 2) Weak Signal의 변이 과정

〈출처〉: R, Eckhoff et al., “Detecting Innovation Signals with Technology-Enhanced Social Media Analysis – Experiences with a hybrid approach in three branches,” *Inter. J. Innovation and Scientific Research*, vol. 17, no. 1, Aug. 2015, pp. 120-130.



(그림 3) Weak signal과 wild card의 관계[6]



(그림 4) Weak Signal과 Wild Card의 관계[7]

하지만, weak signal은 오랜 시간 유지되는 경향이 있기 때문에 초기에 징후를 포착한다면 미래의 변화를 준비할 충분한 시간을 가질 수 있는 이점이 있다.

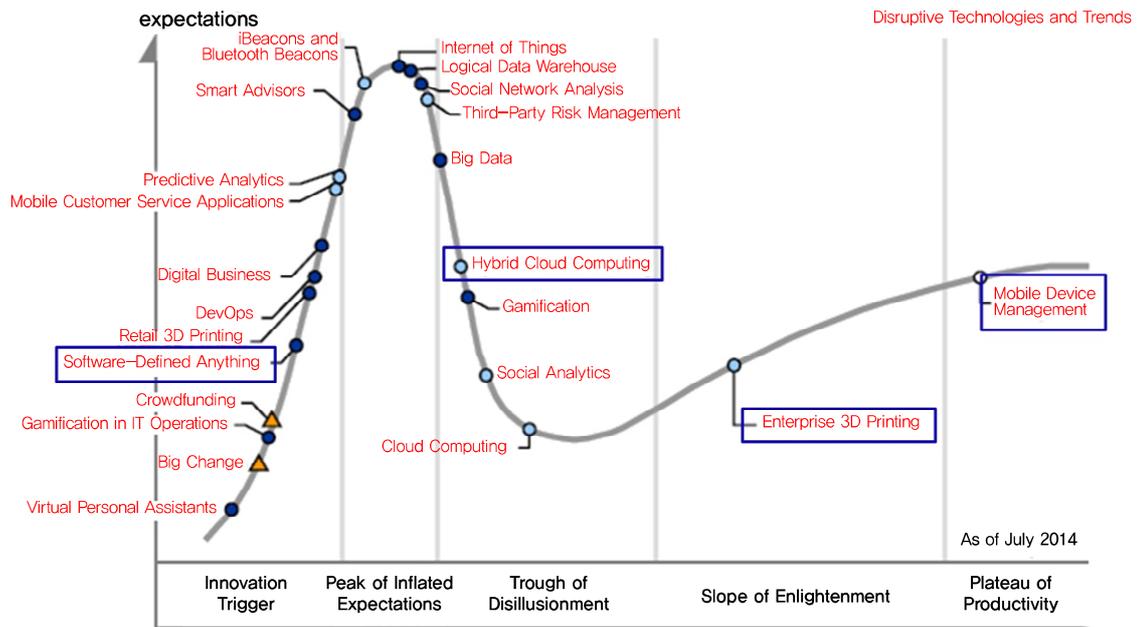
Weak signal의 정의를 좀 더 명확히 하기 위해서 많은 연구가 이루어져 왔고 (그림 4)와 같이 Hiltunen이 제시된 3차원 공간 모델이 주목받고있다[7].

이 모델에서 각 축은 신호(signal), 이슈(issue), 그리고 해석(interpretation)으로 구성되며, [2]에서는 각 요소의 의미를 다음과 같이 해석한다. 신호는 시그널의 수 또는 가시성을 나타내며, 이슈는 사건의 수(예를 들어 신문뉴스, TV 보도, 루머 등)를 나타내며, 해석은 정보 수용자들에게 미치는 미래 신호로써 인지하는 정도를 나타낸다. 신호와 이슈 축은 객관적 차원이라고 할 수 있고, 해석은 주관적 차원이라 할 수 있다. 신호의 수가 증가하고 사건의 수가 많아질수록 그리고 정보 수용자

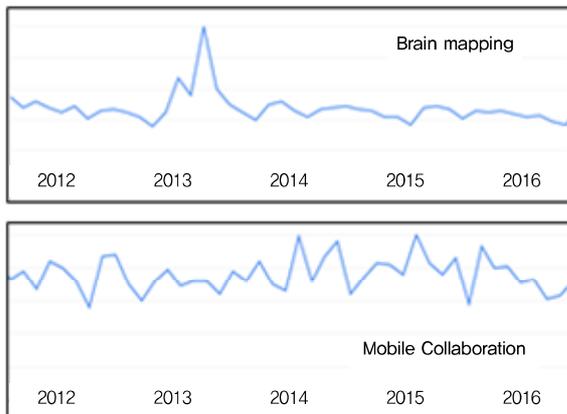
들에게 주어진 신호가 미래 신호로써 큰 영향력을 미칠 때 weak signal 에서 strong signal로 이동하게 된다.

2. Weak Signal 관점에서 본 유망기술 사례

유명기관에서 발굴한 유망기술 발굴 사례를 보면 weak signal 보다는 strong signal에 가까운 기술들이 포함되어 있다. [5]에 정리된 2014년 가트너 10대 유망 기술 중에서 2014년 가트너 hype cycle의 Major Disruptive Technologies and Trends 분야에 제시된 기술들이 (그림 5)에 표시되어 있다. Software Defined Anything 기술은 성숙기까지 5년에서 10년 정도 남은 비교적 초기단계 기술로 볼 수 있지만, mobile device management 기술은 성숙기까지 2년도 채 남지 않았고 Hybrid Cloud와 3D Printing은 5년 이내 성숙기 도달이 예상되므로 weak signal로 보기 어렵다. 2014년 MIT



(그림 5) 가트너 선정 2014 유망기술의 Hype Cycle(Major Disruptive Technologies and Trends) 성숙도



(그림 6) 2014년 MIT 선정 유망기술의 2012~2015년 기간 trends 분석(Google Trends 분석 결과)

선정 유망기술 중에서 Brain Mapping과 Mobile Collaboration 기술은 Google Trends로 분석해보면 (그림 6)에 나타난 것처럼 2012년부터 꾸준히 검색되었기 때문에 weak signal로 보기 어렵다.

매년 발표되는 유망기술들이 미래에 경제·사회적으로 큰 영향을 줄 기술들이라는 측면에서는 맞을지 모르지만 strong signal이거나 wild card를 지난 기술들도

포함되어 있다. 따라서, 이러한 기술들의 연구에 무조건 뛰어드는 것은 기술 경쟁력과 시장 주도권 확보 가능성 관점에서 옳다고 보기 어렵다. 변화의 조짐이 어느 정도 구체화되고 가시적인 결과가 나온 후 fast follower 전략을 통해 국가 경쟁력을 유지하는 것이 이전에는 가능했지만, 이제는 first mover가 기술개발 결과로 생기는 혜택의 대부분을 가져가기 때문이다. 따라서 변혁적 기술의 태동초기에 나타나는 징후나 조짐인 weak signal을 탐지하는 것이 중요하다.

3. Weak Signal 탐지 프로세스 사례

Weak signal을 조기에 발견하고 이를 기술개발로 연계하기 위한 프로세스를 정립하고 이를 실제로 운영하는 사례가 증가하고 있으며, 대표적인 사례는 다음과 같다.

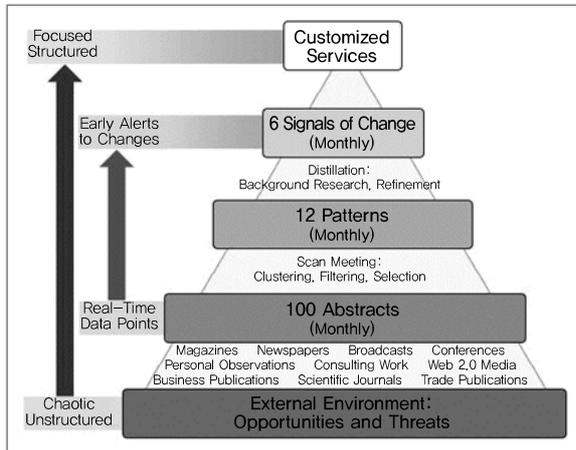
가. Scanning for Emerging Science and Technology Issues(SESTI)

SESTI[8] 프로세스는 (그림 7)과 같이 크게 3 단계로

	Phase 1: Identification of weak signals	Phase 2: Processing of Weak signals	Phase 3: Analysis and interpretation
Focused expert review	High	High	High
Wiki	Low	Low	Low
Twitter	High	Low	Low
Surveys	Low	High	High
Conferences	Low	Medium	High
Text-mining	Low	Medium	Medium

(그림 7) SESTI 단계별 세부방법 비교

〈출처〉: Science and Public Policy, "On concepts and methods in horizon scanning: Lessons from initiating policy dialogues on emerging issues" 2012.



(그림 8) SBI의 Weak Signal 탐색 과정

〈출처〉: <http://strategicbusinessinsights.com/scan>

구성된다. 1단계는 exploratory scan으로써 분석 분야와 분석 데이터에 제한을 두지 않고 weak signal을 찾는다. 2단계는 weak signal을 군집화하고 군집별로 중요도 분석 등을 거쳐서 중요 weak signal을 파악한다. 3단계는 정책적 관점에서 weak signal을 분석하고 해석하는 과정이다. (그림 7)에 정리된 단계별로 사용된 세부방법 비교를 보면 전문가 의견이 모든 단계에서 가장 효과적이며, text mining과 같은 자동화 방법은 상대적으로 효율성이 낮다.

나. Strategic Business Insights(SBI)

세계적인 컨설팅 기관인 SRI의 business intelligence

부문을 모태로 하는 SBI의 Scan 프로세스[9]는 아직 관심을 받지 못하고 있는 숨어있는 이슈들을 발굴하여 사업 기회로 활용할 수 있도록 한다. SBI의 Scan은 전문가들이 소비자 행동 패턴, 정치, 경제, 과학 기술 등 다양한 분야에서 signal을 발굴하여 여러 단계를 거치면서 주요 signal을 선정하게 된다.

IV. Weak Signal 탐지를 위한 정량적 데이터 분석방법

Weak signal 탐지 프로세스 사례들을 분석해본 결과 (그림 1)과 같이 weak signal 탐지를 명시하지 않은 유망기술 발굴 프로세스와 유사한 것을 알 수 있다. 후보 기술들을 여러 단계에 걸쳐서 좁혀 나가며, 그 과정에 전문가 자문이나 데이터 분석이 사용되는 것은 동일하다. 결국 weak signal을 탐지하는지 아니면 strong signal에 가까운 것을 발견하는지를 결정하는 것은 세부 실행방안과 참여 인력이라고 할 수 있다. 전문가 검토나 설문조사와 같이 사람에게 의존한 방법은 오랫동안 시행해온 사례가 있지만, 세부 발굴 방법에 있어서 정량적 데이터 분석은 최근 들어 관심을 받는 분야이다. 지금까지 weak signal의 탐색은 주로 높은 수준의 감각과 전문성을 갖는 전문가의 내용 분석에 의존해 왔다. 하지만, 그러한 분석이 어려울 뿐만 아니라 범위도 넓고 그 분석결과의 해석 또한 전문가 개인의 판단에 맡길 수밖에 없다는 한계가 있었다. 따라서 이러한 작업을 객관적으로 지원하기 위해서 정량적 데이터 분석방법이 필요하다. 그리고, weak signal뿐만 아니라 대부분의 유망기술 선정이 일부 전문가의 주관적·정성적 판단에 의존하는 경우가 많기 때문에 합리성, 객관성, 투명성 측면에서 부족한 점을 보완하기 위해서 정량적 분석방법이 사용되어야 한다.

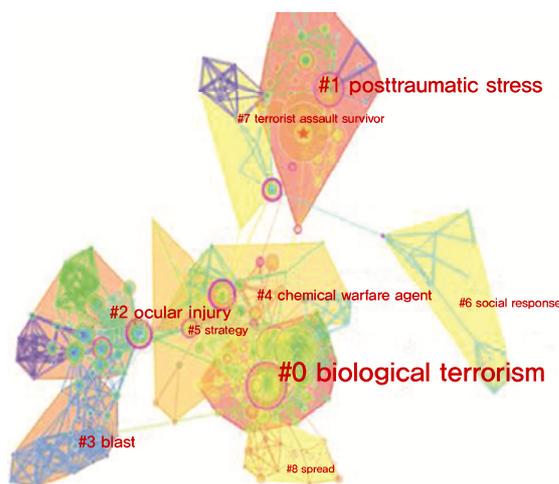
Weak signal 탐지를 위해서 분석할 데이터는 유망기술 발굴을 위해서 분석하는 데이터와 차이가 있을 수 있

고, 동일한 데이터를 분석하더라도 그 분석방법에서 차별화가 되어야 한다. 가령, 신문기사의 제목으로 언급될 정도이면 strong signal에 가깝기 때문에 신문기사의 헤드라인을 분석하는 것은 weak signal 탐지에 적합하지 않다. Weak signal을 탐지하기 위해서는 좀 더 폭넓은 정보인 기사 전체를 분석하는 것이 필요하다. 그리고, 논문, 특허, 웹 데이터 등을 분석하는 경우에는 가시적인 트렌드와 같은 거시적 동향을 파악하기 보다는 시계열 분석에서 1년 단위보다 더 세분화된 시간을 설정하여 분석을 해야 한다.

분석대상 문서의 특성에 따라서 인용정보를 분석하는 방법과 자연어 처리 기술을 기반으로 하여 숨겨진 의미를 추론하는 텍스트 분석으로 크게 구분할 수 있다.

1. 인용정보 분석

대표적인 정형 데이터로는 특허와 논문이 있다. 특허에서는 출원인 정보와 인용정보와 같은 명확한 정보를 활용할 수 있으며, 논문도 저자 정보와 인용정보를 포함하고 있다. 논문과 특허 분석에서 가장 많이 사용되는 정보가 인용정보며, 이를 이용하여 유사 문서끼리 군집화를 할 수 있고 인용이 많이 되는 핵심 문서를 발굴할 수도 있다. (그림 9)는 CiteSpace[10]을 이용하여 핵심



(그림 9) CiteSpace를 이용한 논문 인용정보 분석

연구분야를 파악하고 연구분야들 간의 관계를 도시한 것이다.

하지만, 이와 같이 인용정보를 이용한 분석은 충분한 시간이 지나야만 유의미한 분석을 할 수 있는 데이터가 축적되기 때문에 어떠한 기술들이 유망해졌는지를 파악하는 목적으로는 유용할 수 있지만, weak signal을 탐지하기에는 부적합한 점이 있다. 그리고 인용정보 분석에 많이 사용되는 특허는 출원 후 공개되기까지 18개월이 소요되고, 논문 중에서 저널 논문은 심사에 오랜 시간이 걸리는 경향이 있어서 conference 논문이 주로 분석대상이 된다. 이와 같은 문제점을 가능한 줄이면서 인용정보 분석을 통해 weak signal을 파악하고자 하는 연구들이 진행됐고, 특허 인용정보를 이용하는 대표적인 사례를 살펴보고자 한다.

[11]에 따르면 특허 인용정보 분석을 통해서 weak signal에 가까운 정보를 포함하는 특허를 탐색하는 과정은 (그림 10)과 같이 진행된다. 이렇게 발굴된 특허들이 weak signal이라고 보기는 어렵지만, 기존의 방법보다 유망기술을 내포하는 특허를 좀 더 조기에 발굴할 수 있게 된다. Hot patent는 근래에 citation이 빨리 증가하는 특허를 의미한다. Hot patent가 빠르게 성장하는 기술분야의 내용이라면 최근 특허일 가능성이 크지만, 오래 전에 발견되었지만 새로운 활용 분야를 찾은 경우일 수도 있어서 꼭 최근의 특허일 필요는 없다. 이 hot patent가 next generation patent가 되는 것이 아니고, hot patent를 인용하는 최신 특허가 next generation patent가 된다. Next generation patent를 유망기술을 내포하는 특허로 볼 수 있다. 이렇게 발굴된 특허의 개

STEP 1: hot patent 탐색

STEP 2: next-generation patent 결정

STEP 3: patent 군집화

STEP 4: next-generation patent 순위 결정

(그림 10) Emerging Patent Cluster 탐색 과정

수가 많을 수 있어서 유사 특허끼리 군집화하는 단계를 거치게 된다.

군집화를 거쳐서 선정된 특허들이 얼마나 의미 있는 혁신적인 기술을 다루는지에 대한 점수를 산출하고 이에 따라 우선순위를 부여하게 된다. 이 과정을 거치면 기존 특허의 내용을 점진적으로 개선한 특허들은 제외되게 된다. 유망 특허 점수 산출 기준은 public sector 비율, science index, reference index, originality index의 4가지이다. 이 중에서 public sector 비율과 science index가 주요 기준이다. Public sector 비율은 각 유망 특허 군집 별로 정부기관이나 비영리 기관 등에서 출원한 특허가 차지하는 비율을 의미한다. 정부기관이나 비영리 기관 등에서 출원하는 특허는 성숙기에 이르지 않은 비교적 초기단계의 기술들에 대한 내용들이 많다는 특성에 기반한 기준이다. Science index는 특허에서 인용하는 자료 중에서 특허가 아닌 자료의 수를 의미한다. 특허가 아닌 논문과 같은 자료를 인용하는 특허는 최신 기술동향을 반영할 확률이 높고, 이전 특허만을 인용하는 경우에는 기존 특허나 기술을 점진적으로 개선한 경우가 많다는 연구 결과로부터 만들어진 기준이다.

2. 텍스트 분석

텍스트 분석은 논문이나 특허의 인용정보나 저자 정보와 같은 정형화된 정보 없이 텍스트 자체를 분석하는 것이므로 기본적인 텍스트 마이닝 기법에 weak signal 탐지를 위한 방법이 추가적으로 사용된다. 분석대상은 논문이나 특허뿐만 아니라 웹 데이터와 SNS 정보까지 아주 다양하다. 웹 데이터는 과학기술 관련 뉴스 기사나 블로그 그리고 MIT Technology Review와 같은 온라인 상의 정보를 포함하며, 최신의 정보가 가장 빨리 반영되는 공간이기 때문에 weak signal 탐지를 위해서 가장 적합한 데이터이다.

텍스트 분석은 기본적인 텍스트 마이닝 과정을 거쳐

서 통계정보를 생성하는 것으로부터 시작된다. 텍스트 마이닝의 첫 단계는 입력 텍스트를 대상으로 형태소 분석을 하여 단어를 추출하는 것이다. 추출된 단어 중에서 불용어를 제거하고 유사한 의미의 단어는 통합하는 등의 처리를 거쳐서 의미 있는 단어를 선별하게 된다. 이렇게 추출된 단어들에서 단순히 단어별 빈도를 산출하기 보다는 Term Frequency-Inverse Document Frequency (TF-IDF)와 같은 의미 있는 통계정보를 생성한다. 현재 까지 사용되는 weak signal 탐색을 위한 텍스트 분석은 대부분 출현 단어의 빈도수에 기반을 두고 있다. 이전과 비교하여 새롭게 나타났거나 빈도가 급격히 증가하는 키워드들을 파악하여 weak signal로 보는 것이다. 키워드 기반의 빈도 변화를 주제 변화로 확장할 수 있지만 주제 분석 기술 자체의 기술적 어려움으로 인해서 현재는 키워드 빈도를 이용한 방법이 주로 연구되고 있다.

텍스트 분석을 이용하여 유망기술을 파악하는 사례로는 [2][12][13] 등이 있다. [2]는 연도별 키워드의 출현 빈도수를 이용하여 출현 빈도의 연도별 증가율이 큰 키워드들이 변화의 징후를 나타내는 weak signal로 판단한다. [12]는 인터넷 뉴스에서 키워드의 빈도 또는 키워드가 포함되는 기사의 빈도가 낮으면서 키워드 빈도의 증가율이 높은 키워드가 weak signal을 의미하는 것으로 본다. [2]에서는 빈도의 증가율만을 고려했으나, [12]에서는 빈도의 절대값까지 포함한 차이가 있다. [12]는 인터넷 뉴스 전체를 대상으로 하지 않고 solar energy와 같은 특정 키워드로 검색되는 분야로 제한하여 분석한 한계를 보이며, 키워드 추출을 자동화하지 않고 전문가에 의존함으로써 확장성에 문제가 있고 전문가 의견에 의존함으로써 weak signal 보다는 strong signal에 가까운 키워드가 선택될 가능성이 크다. [13]은 논문의 요약문과 conference session 제목을 동시에 분석하는 특징이 있다.

하지만, 이러한 방법들은 데이터가 수년 간 누적되어야 가능하므로 weak signal 탐지로는 적합하지 않거나,

분석대상 데이터가 특정 분야로 제한되어 해당 분야의 weak signal만 탐지할 수 있는 한계가 있다. [14]는 분석 기간을 몇 주 단위로 아주 짧게 지정할 수 있고, 뉴스 기사를 특정 분야로 제한하지 않고 전체 뉴스 기사를 대상으로 분석할 수 있는 효율적인 알고리즘을 제안한다. 기본적으로 키워드의 빈도에 기반을 두지만, 키워드들의 동시 출현 빈도도 고려함으로써 유의미한 signal의 탐지 확률을 높인다. [14]는 일반적인 뉴스 기사나 Twitter를 대상으로 개발된 것으로서 과학 기술 분야의 weak signal 탐지에 그대로 적용하여 효과를 보기는 어려울 수 있다. 사건 사고와 같은 뉴스 기사는 어떤 일이 생기면 순간적으로 많은 기사가 나오지만, 과학 기술 분야의 초기단계 연구가 이렇게 순간적으로 집중을 받지 않는 않기 때문이다.

위와 같이 텍스트 분석을 통해 과학 기술 분야의 weak signal을 찾으려는 연구는 키워드 빈도 기반의 기초적인 수준에서 머물러 있고, 과학 기술 분야의 발전 특성으로 인해서 다른 분야에서 연구된 분석방법론을 바로 적용하여 효과를 보기 어려울 것으로 예상된다. 현재는 텍스트 분석방법이 weak signal 탐색에서 의미 있는 역할을 하지 못하고 있으며, 탐색 과정에서 중요한 요소로 자리 잡기 위해서는 한 단계 도약할 수 있는 획기적인 기술개발이 필요하다.

V. 결론

과학기술의 발전속도가 점점 빨라지고 있으며 융·복합화와 불확실성의 증가 등으로 인하여 실효성 있는 유망기술의 발굴이 어려워지고 있다. 기술혁신과 글로벌화의 급속한 진전이라는 환경변화에 대처하고 경쟁에서 우위를 점하기 위해서는 변화를 더 빨리 감지하여 연구 및 개발을 위한 시간을 확보하는 것이 중요하다. 따라서 strong signal로 나타나기 전에 weak signal을 탐지하기 위한 프로세스의 정립이 필요하다.

기존 유망기술 발굴 과정에서는 전문가 의견 또는 설문조사와 같은 정성적 방법이 주를 이루었고, weak signal 탐지를 목표로 하는 프로세스도 여전히 정성적 방법에 크게 의존하고 있다. 그러나 대량의 정보로부터 결과를 도출하는데 많은 시간과 비용이 소요될 뿐만 아니라 전문가의 주관적인 가치가 반영될 가능성이 존재한다. 따라서 인용관계 분석이나 텍스트 마이닝과 같은 정량적 데이터 분석방법을 상호 보완적으로 활용하는 방식으로 패러다임의 변화가 시도되고 있다.

하지만, 정량적 데이터 분석에 의한 weak signal 탐색 결과는 전문가 검토를 위한 후보 정보 중의 일부로만 사용되고 있으며, 분석대상이 되는 데이터도 미리 정해진 특정 분야로 한정하는 경우가 많다. 따라서, 특정 분야로 한정하지 않고 대량의 데이터 속에서 weak signal을 탐색할 수 있도록 확장성이 있고 일반화 성능이 뛰어난 분석 기술을 연구해야 한다. 변화의 조짐을 조기에 발견하기 위해서는 논문이나 특허보다는 최신성이 보장되는 인터넷으로 공개되는 온라인 데이터가 유용할 것으로 예상되므로 인용정보와 같은 정형화된 정보의 분석이 아니라 비정형 정보인 텍스트의 분석에 집중해야 한다.

용어해설

Weak Signal 미약하고 중요하지 않은 것처럼 보이지만 미래에 일어날 일들에 대한 정보를 담고 있는 징후. 노이즈 또는 다른 시그널과 혼재되어 탐지가 어려우며, 성숙되기까지 상당한 시간이 걸리는 경우가 많음.

Wild Card 미리 예측하기 어려웠던 중대 사건의 발생을 의미하며 weak signal이 의미하던 현상이 현실로 나타나는 것이고, wild card 이후에는 기회가 거의 없거나 아주 적어진다고 볼 수 있음.

약어 정리

DARPA	Defense Advanced Research Projects Agency
NEDO	New Energy and Industrial Technology Development Organization
TF-IDF	Term Frequency-Inverse Document Frequency

참고문헌

- [1] 이경표 외, “2020년 미래 무선통신 유망기술 발굴,” 한국통신학회논문지, 제38A권 제1호, 2013. 1, pp. 108-126.
- [2] 유선희, 박현우, 김경호, “계량정보분석을 활용한 기술혁신의 Weak Signal 탐색에 관한 연구,” 기술혁신연구, 제17권 제2호, 2009, pp. 109-130.
- [3] 여운동 외, “국가적 차원의 유망연구영역 탐색: Scopus 데이터베이스를 이용한 과학계량학적 접근,” 정보관리연구, 제39권 제3호, 2008, pp.95-113.
- [4] 한국과학기술정보연구원, “유망기술 도출방법론 개발,” 2007. 7.
- [5] KISTEP, “2015 KISTEP 10대 미래 유망기술 선전에 관한 연구,” 2015. 5. 19.
- [6] E. Hiltunen, “Was It a Wild Card or Just Our Blindness to Gradual Change,” *Journal of Futures Studies*, vol. 11, no. 2, Nov. 2006, pp. 61-74.
- [7] E. Hiltunen, “The Future Sign and Its Three Dimensions,” *Futures*, vol. 40, no. 3, April 2008, pp. 247-260.
- [8] E. Amanatidou et al., “On Concepts and Methods in Horizon Scanning: Lessons from Initiating Policy Dialogues on Emerging Issues,” *Science and Public Policy*, vol. 39, no. 2, 2012, pp. 208-221.
- [9] <http://strategicbusinessinsights.com/scan/>
- [10] C. Chen, “CiteSpace II: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature,” *J. American Society for Information Science and Technology*, vol. 57, no. 3, Feb. 2006, pp. 359-377.
- [11] A. Breitzman and P. Thomas, “The Emerging Clusters Model: A Tool for Identifying Emerging Technologies Across Multiple Patent Systems,” *Research Policy*, vol. 44, no. 1, 2015, pp. 195-205.
- [12] J. Yoon, “Detecting Weak Signals for Long-Term Business Opportunities Using Text Mining of Web News,” *Expert Systems with Applications*, vol. 39, no. 16, Nov. 2012, pp. 12543-12550.
- [13] T. Furukawa et al., “Identifying the Evolutionary Process of Emerging Technologies: A Chronological Network Analysis of World Wide Web Conference Sessions,” *Technological Forecasting & Social Change*, vol. 91, Feb. 2015, pp. 280-294.
- [14] E. Schubert, M. Weiler, H.-P. Kriegel, “SigniTrend: Scalable Detection of Emerging Topics in Textual Streams by Hashed Significance Thresholds,” *20th ACM SIGKDD Conference on KDD*, Aug. 2014, pp. 871-880.