

2023년 미래 불확실성에 대응하는 환경변화 탐색 및 대응방안 연구

2023 study on the exploration of environmental changes and its responding measures in response to future uncertainties

임현 외

IV. 기술뉴스 DB 구축 및 신흥핵심기술 도출



IV. 기술뉴스 DB 구축 및 신흥핵심기술 도출

제1절 개요

1. 연구 목적 및 내용

1) 연구 배경

- 주요국들은 신흥·핵심기술을 핵심 국가안보 영역으로 인식하고 기술을 확보하고 보호하며 국제적 표준과 규범 마련을 선도하는 등 기술블록화가 심화
 - 국제적 기술 우위를 선점하기 위한 경쟁이 심화되는 가운데 신흥·핵심기술의 조기경보와 발굴이 어느 때보다 중요한 시점
- 혁신적인 신흥기술의 우위를 바탕으로 기술선도자의 위치를 확보하기 위해서는 기술의 초기단계 동향 파악이 중요
 - 신흥기술은 매우 심각한 안보 위협으로 작용할 수 있으며 사회경제적 파급효과 이상으로 파괴적인 영향을 미칠 가능성이 존재하여 신흥·핵심기술의 조기경보와 발굴이 중요한 시점
 - 기술·산업·정책 환경변화에 대응하여 선제적으로 NEXT 국가전략기술 분야 후보군 탐색 필요

2) 연구 목적

- 미래에 파괴적이고 혁신적인 신흥·핵심기술로 발전할 가능성이 있는 기술을 사전적으로 탐지
 - 글로벌 기술환경의 급속한 변화에 기민하게 대처하여 경쟁우위를 확보하기 위해 기술의 발전을 조기에 인지
 - 후발국에서 선도국으로의 도약을 위한 기술 우위 선점을 위한 상시 정보 융합·생산이 가능한 신흥기술 후보군 발굴 시스템 마련
- 토픽모델링(Topic Modeling) 등 AI 기반 분석 기법을 활용하여 신흥기술 관련 동향 모니터링 프로세스를 기존 수작업 중심에서 AI 기반 데이터 분석 프로세스로 전환

- 공급망·통상, 신산업, 외교·안보 등 기술패권 구도 속 대내·외 환경 변화를 신속히 파악할 수 있는 데이터 기반의 모니터링 체계 구축 필요
- 신기술 후보풀 마련을 위한 해외 주요 기관의 발간자료, 해외 뉴스 DB 구축
- 구축된 데이터의 심층 분석을 통한 국가전략기술 신기술 후보풀 마련

3) 연구 내용

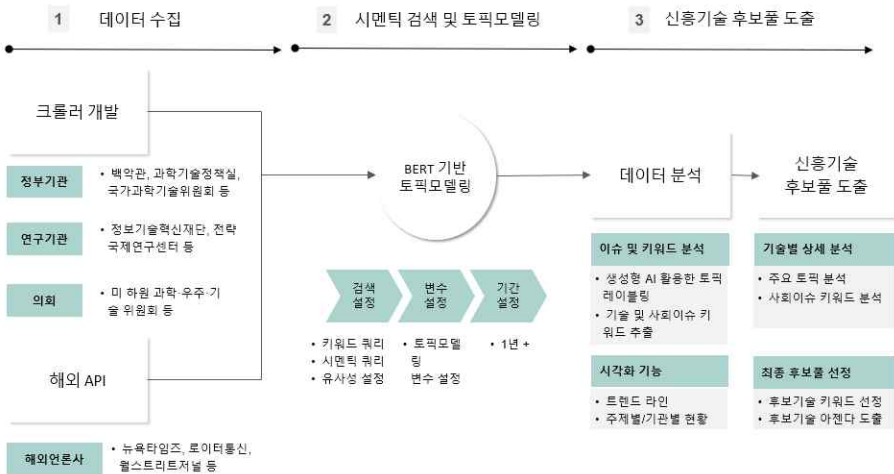
- 신기술·핵심기술 모니터링을 위한 DB 구축과 인공지능 기술을 통해 신규 전략기술 선정에 활용할 수 있는 정보를 추출하는 데이터 기반 분석시스템을 확립
 - 웹 크롤링을 기반으로 해외 주요 기관(주요국 정부기관, 연구기관, 의회 등)의 발간자료를 함께 수집하여 DB화
 - 생성형 AI 기반 빅데이터 분석을 통해 기술뉴스 DB를 활용하여 매체별 특정 카테고리(Technology 또는 Science)에 대한 대규모 뉴스 데이터를 수집하여 DB화
 - 구축한 DB 중심으로 인공지능 빅데이터 분석 방법론, 생성형 AI 기반의 주요토픽 추출 및 심층 분석을 통해 신기술 후보풀 도출
 - 과학기술관련 클러스터링 결과 및 트렌드 관련 시각화 데이터 도출

〈표 4-1〉 KISTEP ‘미래유망기술’과의 차별성

구분	(가칭) 신기술 센싱	KISTEP 미래유망기술
핵심가치	공급망·통상, 신산업 선점, 외교안보	미래사회의 핵심 이슈
전망 기간	5년 이내	5~10년
방법	인공지능(딥러닝) 활용, 전문가 검토	문헌조사, 설문조사
시기	연 2회(상, 하반기)	연 1 회
산출물	신기술·핵심기술 후보군	10개 내외 유망기술

2. 연구 절차

- (1단계) 주요 기술 정보 및 정책 데이터의 수집, (2단계) 시멘틱 검색 및 토픽모델링을 통한 사회 이슈 파악, (3단계) 신흥기술 후보폴 도출 등 신흥·핵심기술 모니터링 과정을 AI 기반의 정량적 데이터 기반으로 수행
- (1단계) 웹 크롤링, 뉴스 API를 활용하여 신흥기술동향 관련 해외 주요 기관 및 뉴스 데이터를 수집하고 AI 기반 시멘틱 검색을 활용하여 신흥기술 동향 관련 자료 필터링 수행 및 DB 구축
- (2단계) AI기반 토픽모델링 및 생성형 AI모델을 활용하여 클러스터별 공통 주제 요약 및 자동 토픽 레이블링
 - BERT 기반 토픽모델링을 통해 구축한 DB 내 신흥·핵심 기술 관련 자료를 추출
 - 추출된 자료를 중심으로 생성형 AI를 활용하여 토픽레이블링, 기술이슈 및 사회이슈 키워드 추출 실시
- (3단계) 생성형 AI 기반으로 기술 키워드와 연관된 세부 아젠다에 대한 국가전략기술 관점의 선정사유 요약



[그림 4-1] 연구 절차

제2절 연구 결과

1. 주요 기술 정보 및 정책 데이터의 수집 결과

- API 활용 및 크롤링을 통하여 '22.06.01.부터 '23.09.18.까지의 해외 주요 기관 및 뉴스 발간 자료 전수인 700,682건의 raw data를 수집
 - 전처리 전 raw data는 704,122건으로 언론사 자료 683,401건, 연구기관 자료 6,197건, 의회 자료 80건, 정부기관 자료 11,004건으로 구성
 - 수집한 raw data는 중복 문서, 본문이 지나치게 짧은 문서 등을 제거하는 전처리 작업을 거쳐 700,682건으로 정리
 - raw data는 신흥-핵심 기술 선정과 무관하더라도 과학기술정책 관련해 자료를 발간하는 기관 및 언론사들의 자료를 모은 것으로 향후 타 과제 수행을 위한 자료로도 활용 가능
- 텍스트분석 및 생성형 AI를 활용하여 raw data 내 유효문서를 추출하기 위한 77건의 시멘틱 쿼리 마련
 - 유효문서를 검증하기 위하여 문서가 신흥-핵심기술과 관련되어 있는지 검토가 필요하며 시멘틱 쿼리는 단순히 키워드의 존재 유무로만 관련성을 검증하지 않고 문맥 등을 함께 고려해 관련성을 검증하는데 활용
 - ※ 시멘틱 검색은 사용자 검색 쿼리(문장)의 의미를 이해하고 이와 유사한 의미를 가진 정보를 검색하는 방식
 - 백악관, 청와대, 과학기술정보통신부 등에서 공개한 국가전략기술 관련 문건에 대한 텍스트분석을 통해 주요 키워드 및 문구를 추출
 - 텍스트분석은 참고한 문건 내 키워드 및 문구의 출현 빈도수를 확인하기 위해 시행하였으며 출현 빈도수가 높은 키워드 및 문구 선정
 - 생성형 AI를 활용하여 추출한 주요 키워드 및 문구와 유사한 키워드 및 문구를 생성하고, 생성된 키워드 및 문구를 전문가 검토를 통해 검증
 - 77건의 시멘틱 쿼리는 공급망-통상 관련 쿼리 21개, 외교-안보 관련 쿼리 27개, 신산업 관련 쿼리 18개, 기술패권 관련 쿼리 11개로 구성

2. 시맨틱 검색 및 토픽모델링 기반 이슈 분석 결과

- 1단계에서 마련한 시맨틱 쿼리를 활용하여 700,682건의 문서 중 신흥핵심기술 관련 문서 추출을 위한 시맨틱 검색을 실시*한 결과, 6,098건의 유효문서를 확보

* 검색 쿼리 문장과 검색 대상 데이터간의 시맨틱 공간의 유사성(cosine similarity 0.4 이상)을 기준으로 유사성 낮은 노이즈 데이터 제거

- 유효문서를 대상으로 BERT 기반 토픽모델링*을 수행하여 클러스터링을 위한 최소 데이터 수 5건을 만족하는 227개의 토픽을 포착

* BERT 기반 토픽모델링은 사전학습된 BERT 언어모델을 활용한 임베딩과 클래스 기반(class-based) TF-IDF 기술을 활용하여 주제 설명에 필요한 중요한 단어를 통해 주제를 추출하는 분석 방법

- 토픽모델링 분석 결과 6,098건의 유효문서 중 227개의 토픽과 연관된 문서 4,116건이 정리되었으며, 주요 토픽과 관계없는 1,982건의 문서는 상세 분석에 미활용

- 227개 토픽에 대하여 생성형 AI 기반으로 토픽 해석을 진행하고, 관련된 기술과 사회이슈를 추출

- 토픽모델링 분석 결과는 인간이 이해하기 어려운 형태로 결과가 추출되어, 기존 연구들은 별도 전문가 검토작업을 거쳐 토픽을 인간의 언어로 해석하는 작업을 거쳤으나, 금번 연구에서는 생성형 AI를 활용하여 작업 수행

- 생성형 AI를 활용하여 최종 유효문서 4,116건을 검토한 결과, 최종 유효문서에서 언급된 기술은 총 1,734개이며, 언급된 사회 이슈는 총 2,033개

- 기술 1,734개, 사회 이슈 2,033개는 키워드 표준화를 거친 결과로, 임베딩 벡터 클러스터링을 활용하여 키워드들을 대표할 수 있는 “표준 키워드”를 생성

- “표준 키워드”는 공급망-통상, 신산업, 외교-안보, 기술패권 관점으로 구분하고, 신흥핵심기술 후보(안)과 매칭하여 관련 이슈를 파악하는데 활용

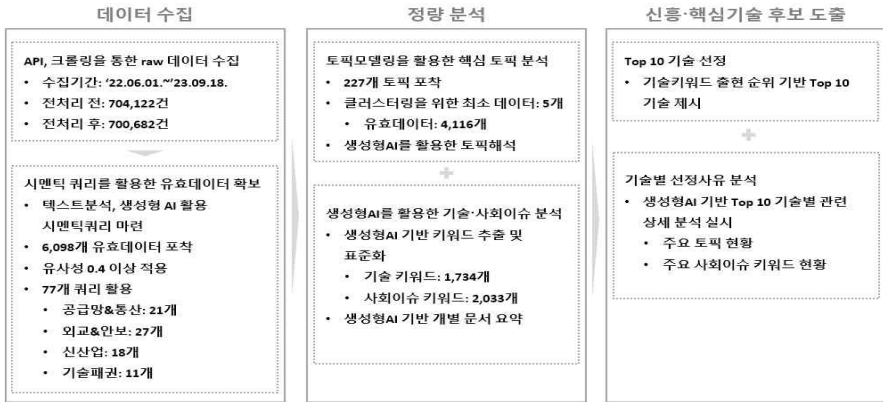
3. 신흥·핵심기술 도출 결과

- 생성형 AI를 활용하여 파악한 기술 1,734개를 유효문서에서 언급된 빈도수 기반으로 나열하여 상위 10개 기술을 신흥·핵심기술 후보(안)으로 추출
- 추출된 기술은 순서대로 △반도체, △인공지능, △사이버보안, △양자컴퓨팅, △칩 제조, △악성소프트웨어, △전기자동차, △바이오, △클라우드 컴퓨팅, △친환경에너지
- ※ 클라우드 컴퓨팅과 친환경 에너지 기술은 각각 54회 언급되어 공동 9위로 신흥핵심기술 후보로 선정

〈표 4-2〉 신흥·핵심기술 후보(안)

기술명	언급 빈도	언급 빈도 백분율
반도체	536	9.03
인공지능	405	6.83
사이버보안	242	4.08
양자컴퓨팅	129	2.17
칩 제조	97	1.63
악성 소프트웨어	83	1.40
전기자동차	74	1.25
바이오	55	0.93
클라우드 컴퓨팅	54	0.91
친환경 에너지	54	0.91
총합	1,729	29.14%

- 신흥·핵심기술 후보(안)에 대하여 생성형 AI를 활용한 최종 유효문서 검토를 실시하여 공급망통상, 신산업, 외교·안보, 기술패권 관점의 주요 토픽을 정리
- 주요 토픽은 10개 신흥·핵심기술 후보(안)과 연관된 유효문서에서 언급된 토픽의 언급 빈도를 카운트하여 선정
- ※ 신흥·핵심기술 후보별로 공급망통상, 신산업, 외교·안보, 기술패권 관점의 유효문서가 없는 경우도 있어, 모든 후보가 공급망통상, 신산업, 외교·안보, 기술패권 관점의 토픽 및 사회이슈가 존재하는 것은 아님



[그림 4-2] 연구 결과 요약

- 10개 신흥·핵심기술 후보(안)과 주요 토픽 매칭 결과 후보 기술 선정과 연관된 토픽이 잘 추출되어 향후 전문가 평가에 활용 가능할 것으로 판단

〈표 4-3〉 반도체 분야 주요 토픽

<ul style="list-style-type: none"> Dutch restrictions on chip tech TSMC's semiconductor expansion plans CHIPS for America Program U.S. restrictions on China chips Taiwan's Semiconductor Trade 	<ul style="list-style-type: none"> CHIPS for America Program U.S. Semiconductor Funding Crisis Advancing American Semiconductor Research Investment in Clean Industrial Economy CHIPS Act and semiconductor manufacturing
공급망·통상	신산업
<ul style="list-style-type: none"> U.S. restrictions on China chips China export restrictions US-China Chip Restrictions Chip equipment export controls Microchip Supply Chain Fix 	<ul style="list-style-type: none"> CHIPS for America Program CHIPS and Science Act Advancing American Semiconductor Research CHIPS Act and semiconductor manufacturing Britain's semiconductor investment strategy
외교·안보	기술패권

〈표 4-4〉 인공지능 분야 주요 토픽

<ul style="list-style-type: none"> • Infrastructure Cybersecurity • Outbound Investment Screening • EU Economic Security and China • US-China Export Controls • Japan-South Korea cooperation 	<ul style="list-style-type: none"> • AI Governance Configurations • Regional Tech Hubs Program • Net-zero future technologies • Data Localization Impact Analysis • Tech Layoffs Increase Hiring
공급망통상	신산업
<ul style="list-style-type: none"> • Defense Acquisition Technologies • Infrastructure Cybersecurity • AI Governance Configurations • Federal Cybersecurity and AI • Federal AI Research and Development 	<ul style="list-style-type: none"> • AI Governance Configurations • Self-Driving Science Labs • Regulation of AI • USPTO and Intellectual Property • Responsible AI Commitments
외교·안보	기술패권

〈표 4-5〉 사이버보안 분야 주요 토픽

<ul style="list-style-type: none"> • Infrastructure Cybersecurity • Foreign Port Security Assessment • Securing 5G Wireless Networks • Regulation of Crypto Exchanges • Partnership for Global Infrastructure 	<ul style="list-style-type: none"> • Infrastructure Cybersecurity • Manufacturing USA and Smart Manufacturing • USPTO and Intellectual Property • China's Tech Talent and Industrial Policies • Regulation of Crypto Exchanges
공급망통상	신산업
<ul style="list-style-type: none"> • Infrastructure Cybersecurity • NIST Cybersecurity Framework Updates • Russian cyberattacks in Ukraine • Chinese hacking activity • Nuclear Weapons Cybersecurity 	<ul style="list-style-type: none"> • Infrastructure Cybersecurity • Tech Layoffs Increase Hiring • Securing 5G Wireless Networks • Regional Tech Hubs Program • CIO 2023 Priorities
외교·안보	기술패권

〈표 4-6〉 양자컴퓨팅 분야 주요 토픽

<ul style="list-style-type: none"> • Japan-South Korea cooperation • US-China Export Controls • EU Economic Security and China • U.S.-EU Trade Cooperation • Biden's Executive Order on Chinese Technology 	<ul style="list-style-type: none"> • Data Localization Impact Analysis • Advancing American Semiconductor Research • Regional Tech Hubs Program • CHIPS and Science Act • Infrastructure Cybersecurity
공급망·통상	신산업
<ul style="list-style-type: none"> • Post-Quantum Encryption Standardization • Quantum Technology Research and Applications • Defense Acquisition Technologies • Outbound Investment Screening • India-US Relations 	<ul style="list-style-type: none"> • Quantum Technology Research and Applications • Tech Layoffs Increase Hiring • Achieving America's STEMM Equity • CHIPS and Science Act • CHIPS for America Program
외교·안보	기술패권

〈표 4-7〉 칩 제조 분야 주요 토픽

<ul style="list-style-type: none"> • Dutch restrictions on chip tech • U.S. restrictions on China chips • TSMC's semiconductor expansion plans • ASML chip demand recovery • Chinese chip industry curbs 	<ul style="list-style-type: none"> • ASML chip demand recovery • U.S. Semiconductor Funding Crisis • Dutch restrictions on chip tech • Biden's Executive Order on Chinese Technology • TSMC's Expansion in Arizona
공급망·통상	신산업
<ul style="list-style-type: none"> • U.S. restrictions on China chips • Japan restricts chip equipment • US-China Chip Restrictions • US-China chip restrictions • Chinese chip industry curbs 	<ul style="list-style-type: none"> • CHIPS for America Program • Taiwan's Semiconductor Trade • U.S. Semiconductor Funding Crisis • TSMC's Expansion in Arizona • Britain's semiconductor investment strategy
외교·안보	기술패권

〈표 4-8〉 악성 소프트웨어 분야 주요 토픽

<ul style="list-style-type: none"> • Infrastructure Cybersecurity 	-
공급망·통상	신산업
<ul style="list-style-type: none"> • Biden's Prohibition on Government Spyware • Infrastructure Cybersecurity • Russian cyberattacks in Ukraine • Cybersecurity Collaboration in Abraham Accords • NIST Cybersecurity Framework Updates 	<ul style="list-style-type: none"> • Infrastructure Cybersecurity
외교·안보	기술패권

〈표 4-9〉 전기자동차 분야 주요 토픽

<ul style="list-style-type: none"> • Taiwan's Semiconductor Trade • China's Industrial Heft • Critical Mineral Supply Chain • EU Green Deal Industrial Plan • Auto sales supply chain woes 	<ul style="list-style-type: none"> • Critical Mineral Supply Chain • Net-zero future technologies • Investment in Clean Industrial Economy • India's Clean Energy Transition • China's Industrial Heft
공급망·통상	신산업
<ul style="list-style-type: none"> • China export restrictions • Biden's visit to South Korea • Japan-South Korea cooperation • Russian cyberattacks in Ukraine 	<ul style="list-style-type: none"> • Net-zero future technologies • China's Industrial Heft • Malaysia's Economic Growth Outlook • Digital Habits and Work Efficiency • Investment in Clean Industrial Economy
외교·안보	기술패권

〈표 4-10〉 바이오 분야 주요 토픽

<ul style="list-style-type: none"> • Outbound Investment Screening • Japan-South Korea cooperation • US-China Trade Relations • Biomanufacturing Action Plan 	<ul style="list-style-type: none"> • Biomanufacturing Action Plan • Data Localization Impact Analysis • Regional Tech Hubs Program • Defense Acquisition Technologies • Advancing American Semiconductor Research
공급망·통상	신산업
<ul style="list-style-type: none"> • Defense Acquisition Technologies • EU Economic Security and China • U.S. restrictions on China chips • Japan-South Korea cooperation • Infrastructure Cybersecurity 	<ul style="list-style-type: none"> • Space exploration and satellite market • Digital Habits and Work Efficiency • Regional Tech Hubs Program • Net-zero future technologies • Advancing American Semiconductor Research
외교·안보	기술패권

〈표 4-11〉 클라우드 컴퓨팅 분야 주요 토픽

<ul style="list-style-type: none"> • Infrastructure Cybersecurity • Tech's Terrible Week, Stocks • EU Microsoft antitrust probe • U.S. restrictions on China chips • Taiwan's Chip Supply Chains 	<ul style="list-style-type: none"> • Digital Platform Competition and Antitrust • Net-zero future technologies • Defense Acquisition Technologies • Product Innovation by Owners • Data Localization Impact Analysis
공급망·통상	신산업
<ul style="list-style-type: none"> • Infrastructure Cybersecurity • U.S. restrictions on China chips • Defense Acquisition Technologies • Securing 5G Wireless Networks • Tech's Terrible Week, Stocks 	<ul style="list-style-type: none"> • Tech Layoffs Increase Hiring • Digital Platform Competition and Antitrust • China's Tech Talent and Industrial Policies • Data Localization Impact Analysis • EU Microsoft antitrust probe
외교·안보	기술패권

〈표 4-12〉 친환경 에너지 분야 주요 토픽

<ul style="list-style-type: none"> • Outbound Investment Screening • Critical Mineral Supply Chain • USMCA Transparency • US-China Trade Relations • North American Semiconductor Conference 	<ul style="list-style-type: none"> • Net-zero future technologies • Critical Mineral Supply Chain • Investment in Clean Industrial Economy • EU Green Deal Industrial Plan • Infrastructure Cybersecurity
공급망·통상	신산업
<ul style="list-style-type: none"> • NSF Science Partnerships Collaborations • Infrastructure Cybersecurity • USPTO and Intellectual Property • Securing 5G Wireless Networks • Biden's visit to South Korea 	<ul style="list-style-type: none"> • Investment in Clean Industrial Economy • Net-zero future technologies • CHIPS for America Program • Digital Habits and Work Efficiency • AGOA trade-US Africa
외교·안보	기술패권

제3절 결론 및 개선방향

1. 결론

- 생성형 AI를 활용하여 최종 DB에 가장 많이 언급된 기술 상위 10개를 추출
 - 반도체, 인공지능, 사이버보안, 양자컴퓨팅 기술 등이 상위 10개 기술로 선정되었으며, 정량 분석에 따라 도출한 Top 10 기술 중 9개 기술이 12대 국가전략기술*과 매칭되며 정량 분석을 통한 신흥핵심기술 후보군 발굴의 유효성 확인
 - * △사이버보안, △인공지능, △양자, △반도체·디스플레이, △이차전지, △첨단 모빌리티, △차세대 원자력, △첨단 바이오, △우주항공·해양, △수소, △차세대 통신, △첨단로봇·제조
 - 최종 DB에 기술 관련 멘션은 5,933건이 있었으며, 이 중 상위 10개 기술과 관련된 멘션은 1,729건으로 약 29.14%의 비중을 차지
- AI 기반 분석 기법을 활용할 경우, 신흥기술 관련 동향 모니터링 프로세스를 기존 전문가 및 문헌분석에 기반한 수작업 방식에서 AI 기반 데이터 분석 프로세스로 전환 가능
 - 주요 기관 및 언론사에 대한 전체 자료를 상시 수집하고 DB화하여, 모니터링을 위한 자료 수집 및 검토하는 과정을 최소화하여 빠른 환경에 보다 신속하게 대응 가능
 - 향후 수집데이터의 범위를 논문 및 특허로 확장하여 12대 국가전략기술 단위가 아닌 50대 핵심기술 단위로 신규 후보(안) 도출하는 작업도 가능할 것으로 판단
 - 수집한 자료를 기반으로 AI 분석 방식을 활용해 인력 투입을 최소화한 초안 마련이 가능해져 모니터링 결과 고도화에 많은 인력 및 시간 투입 가능
 - AI 기반의 정량 분석 방식으로 100% 대체는 어렵더라도 AI 분석을 통해 단기간에 신흥핵심기술 후보풀을 확보하고 전문가는 후보풀을 검토만 하는 방식으로 작업 방식 변경 가능

2. 개선방향

1) 한계점

- 데이터 소스, 연구기관 성격에 따른 기술 편향 가능성 존재
 - 정부기관, 연구기관, 의회, 주요 언론사 등 대표성 높은 기관의 자료임에도 불구하고 데이터 출처의 선택에 따른 기술 전망에 대한 편향성 존재 가능
- Strong signal 중심 분석에 따른 메가트렌드에 포커스하여 선제적 기술 포착의 어려움
 - 신흥기술 포착을 위해서는 데이터 속에 “잠재된” 기술 요소를 선제적으로 파악하는 것이 중요

2) 개선방향

- 모니터링 대상기관을 넓혀 데이터 소스의 다양성을 확보하고 기관의 수를 확장하여 다양한 기술과 이해관계자의 관점 반영 필요
 - 연구기관 성격에 따른 데이터 출처 균형 필요
 - 연구 성격에 맞게 데이터 소스를 선택할 수 있는 기능 필요
- 잠재적 유망기술 포착을 위한 weak signal 지표 개발 및 분석 필요
 - 분석 주기를 짧게 설정 후ダイナミック한 시계열적 변화를 모니터링
 - 세부적인 이슈 파악을 위해 토픽 군집 기준치 축소 필요
- 이종 데이터 결합을 통한 기술 혁신성 파악
 - 특허 데이터 분석을 통한 신흥기술 키워드 및 니즈 파악뿐만 아니라 기술 혁신 역량 등을 파악하여 기술의 전략적 중요도 파악