

정책연구 25-29-04

ICT 주요이슈 분석체계 구축

Establishment of Analysis System for ICT Issue

2025. 12

연구기관 : 정보통신정책연구원

정책연구 25-29-04

ICT 주요이슈 분석체계 구축

(Establishment of Analysis System for ICT Issue)

손녕선/박지원/진정민/신현호/박재홍

2025. 12

연구기관 : 정보통신정책연구원



과학기술정보통신부
Ministry of Science and ICT

이 보고서는 2025년도 과학기술정보통신부 정보통신진흥기금 정보통신방송연구개발사업의 연구결과로서 보고서 내용은 연구자의 견해이며, 과학기술정보통신부의 공식입장과 다를 수 있습니다.

제 출 문

과학기술정보통신부 장관 귀하

본 보고서를 『ICT 주요이슈 분석체계 구축』의 연구결과보고서로 제출합니다.

2025년 12월

연구기관 : 정보통신정책연구원

총괄책임자 : 손녕선 연구위원

참여연구원 : 박지원 연구위원

진정민 전문연구원

신현호 위촉연구원

박재홍 교수

목 차

요약문	xi
제1장 서론	1
제1절 연구의 필요성 및 목적	1
제2장 ICT 주요 이슈 분석	3
제1절 개요	3
제2절 디지털 헬스케어에 대한 온라인 인식 변화	7
1. 디지털 헬스케어의 정의	7
2. 디지털 헬스케어의 기술 동향	9
3. 디지털 헬스케어의 국내외 정책 동향	11
4. 텍스트 분석 개요	14
5. 디지털 헬스케어 언급 추이	14
6. 연관어 분석	17
7. 감성분석	18
8. 결론	19
제3절 멀티모달 인공지능 기술 트렌드 분석	21
1. 멀티모달 인공지능의 정의 및 기술 트렌드	21
2. 분석 개요	25
3. 멀티모달의 언급 추이	25
4. 멀티모달 연관어 분석 결과	28
5. 결론	29
제4절 양자컴퓨팅 키워드의 언론 인식 변화	32
1. 양자컴퓨팅의 정의	32

2. 분석 개요	33
3. 양자컴퓨팅의 언급 추이	34
4. 양자컴퓨팅 연관어 분석 결과	38
5. 결론	41
제5절 개방형 AI 모델 동향 분석	43
1. 개방형 AI 모델의 정의	43
2. 개방형 AI 모델의 발전 과정	43
3. 개방형 AI 모델과 온프레미스 추세	49
4. 개방형 AI 키워드 텍스트 분석	51
5. 결론	57
제3장 ICT 뉴스경기지수 구축 및 고도화	59
제1절 개요	59
제2절 ICT 뉴스경기지수 작성 방법론	60
1. 키워드 선정 및 자료수집	60
2. ICT 뉴스경기지수 작성 방법	62
3. 문장 분류: KoBERT	63
제3절 ICT 뉴스경기지수 구축 결과	65
제4절 국내 최신 트렌드를 반영한 ICT 키워드 발굴	69
1. 개요	69
2. 연구범위 및 방법론	69
3. 기존 키워드의 변화 양상 분석	72
4. 최신 키워드 발굴	87
5. 최종 선정 키워드	96
6. 신규 키워드 기반 ICT 뉴스경기지수와 BSI 간 비교	98
제4장 ICT 키워드 도출 자동화 체계 구축	102
제1절 개요	102
제2절 국내 ICT 관련 뉴스 기반 키워드 자동화 체계 구축 방법론	105

1. 크롤링 자동화	107
2. 뉴스기반 키워드 추출	107
3. ICT 키워드 선정	110
제 3 절 결과 및 평가	114
제 5 장 온라인 노동지수 구축	120
제 1 절 개 요	120
제 2 절 온라인 정보를 통한 노동지수 구축: 해외사례	120
제 3 절 온라인 노동지수 구축 방법	125
제 4 절 온라인 노동지수 분석	126
제 5 절 IT 직종 공고 분석	132
참고문헌	141

표 목 차

〈표 2-1〉 국내외 디지털 헬스케어 정의	8
〈표 2-2〉 국내외 디지털 헬스케어 기술 동향	9
〈표 2-3〉 CES 2025 디지털 헬스케어 부문 혁신상 수상 목록(일부)	10
〈표 2-4〉 국내 디지털 헬스케어 정책 동향	11
〈표 2-5〉 해외 디지털 헬스케어 정책 동향	13
〈표 2-6〉 연도별 디지털 헬스케어 관련 텍스트 분석 결과	17
〈표 2-7〉 디지털 헬스케어 키워드 감성분석 결과	18
〈표 2-8〉 멀티모달 인공지능의 핵심과제 5가지 분류	22
〈표 2-9〉 멀티모달 인공지능의 세대별 기술 트렌드	24
〈표 2-10〉 2021년~2025년 5월 멀티모달 관련 온라인 언급량 추이	25
〈표 2-11〉 연도별 멀티모달 관련 주요 키워드	28
〈표 2-12〉 2018년~2025년 8월 양자컴퓨팅 관련 뉴스 언급량 추이	34
〈표 2-13〉 연도별 양자컴퓨팅 관련 주요·사업자 관련 키워드	38
〈표 2-14〉 개방형 AI 모델과 폐쇄형 AI 모델 비교	44
〈표 2-15〉 LLaMA 1의 벤치마크 결과	46
〈표 2-16〉 연도별 등장 빈도 상위 키워드	54
〈표 2-17〉 토픽별 주요 키워드	56
〈표 3-1〉 키워드 리스트	61
〈표 3-2〉 키워드 리스트	73
〈표 3-3〉 트렌드가 지난 키워드 변화 양상	73
〈표 3-4〉 꾸준히 언급되는 키워드 변화 양상	75
〈표 3-5〉 최근에 나타난 키워드 변화 양상	81
〈표 3-6〉 신규 키워드 후보 리스트: 출처	88

<표 3-7> 신규 키워드 후보 리스트: 등장횟수 89

<표 3-8> 최종 선정 키워드 분류 결과 89

<표 3-9> 선정된 키워드 목록 96

<표 3-10> 최종 키워드 목록 97

<표 3-11> 2024년 시차별 ICT와 BSI간 상관계수 101

<표 4-1> 프롬프트 예시(일부 발췌) 111

<표 4-2> 최종 출력 예시 - IT 일반 카테고리 ICT 키워드 선정 결과 112

<표 4-3> IT 일반 카테고리 키워드 추출 결과 114

<표 4-4> 타 카테고리 키워드 추출 결과 115

<표 4-5> 보고서 기반 키워드와 뉴스 기반 ICT 키워드 비교 117

<표 5-1> 직무 분류 및 키워드 목록 132

그림 목 차

[그림 2-1] 감성분석 예시	4
[그림 2-2] 연관어 분석 예시	4
[그림 2-3] KOSIS open API 제공 서비스 예시	5
[그림 2-4] 한국은행 open API 제공 서비스 예시	6
[그림 2-5] 글로벌 디지털 헬스케어 산업 규모 전망	7
[그림 2-6] 디지털 헬스케어 뉴스 언급량 추이	15
[그림 2-7] 디지털 헬스케어 매체별 언급 비중 추이	16
[그림 2-8] 멀티모달 인공지능의 핵심과제	23
[그림 2-9] 멀티모달 온라인 언급량 추이	27
[그림 2-10] 멀티모달 매체별 언급량 및 비중 추이	27
[그림 2-11] 위성통신 언론 언급량 추이	33
[그림 2-12] 양자컴퓨팅 뉴스 언급량 추이	36
[그림 2-13] 양자컴퓨팅 매체별 언급량 및 비중 추이	37
[그림 2-14] 개방형 AI 모델과 폐쇄형 AI 모델간 MMLU 벤치마크 비교	45
[그림 2-15] LLaMA 3 70B와 GPT-4o 벤치마크 비교	48
[그림 2-16] DeepSeek와 OpenAI-o1 벤치마크 비교	48
[그림 2-17] 기업이 오픈소스 모델을 사용하는 이유	50
[그림 2-18] 월별 기사수 추이	52
[그림 2-19] BERTopic 알고리즘	55
[그림 2-20] 토픽별 유사도 행렬	56
[그림 3-1] BERT와 타 자연어 처리 모델 비교	63
[그림 3-2] ICT 뉴스경기지수 결과	66
[그림 3-3] 네트워크 분석 결과: 2025년 10월	67

[그림 3-4] 네트워크 분석 결과: 2025년 11월	68
[그림 3-5] '자동화' 키워드 시계열 분석 결과	74
[그림 3-6] '네트워크' 키워드 시계열 분석 결과	76
[그림 3-7] '보안' 키워드 시계열 분석 결과	77
[그림 3-8] '로봇' 키워드 시계열 분석 결과	78
[그림 3-9] '우주' 키워드 시계열 분석 결과	79
[그림 3-10] '플랫폼' 키워드 시계열 분석 결과	80
[그림 3-11] '6G' 키워드 시계열 분석 결과	82
[그림 3-12] '메타버스' 키워드 시계열 분석 결과	83
[그림 3-13] '모빌리티' 키워드 시계열 분석 결과	84
[그림 3-14] '플랫폼 엔지니어링' 키워드 시계열 분석 결과	85
[그림 3-15] '디지털 안전' 키워드 시계열 분석 결과	86
[그림 3-16] '디지털 트윈' 키워드 시계열 분석 결과	90
[그림 3-17] '딥페이크' 키워드 시계열 분석 결과	91
[그림 3-18] '퀀텀 컴퓨팅' 키워드 시계열 분석 결과	92
[그림 3-19] '생성형 AI' 키워드 시계열 분석 결과	94
[그림 3-20] '드론' 키워드 시계열 분석 결과	95
[그림 3-21] ICT 뉴스경기지수 (기존 키워드, 최신 키워드, 종합)	99
[그림 3-22] ICT뉴스경기지수와 BSI 시각화	100
[그림 4-1] BERTopic 구조	104
[그림 4-2] 네이버 뉴스 IT/과학 섹션 내 서브 카테고리(모바일, 인터넷/SNS, 통신/뉴미디어, IT일반, 보안/해킹, 컴퓨터, 게임/리뷰, 과학 일반)	105
[그림 4-3] 뉴스 기반 키워드 도출 자동화 프레임워크	106
[그림 4-4] 시계열 추세 결과물 예시 - 생성형 AI 키워드	113
[그림 5-1] 국가/산업별 온라인 구인구직 광고 비중: Cedefop	121
[그림 5-2] 국가/직종별 온라인 구인구직 광고 비중: Cedefop	122
[그림 5-3] 온라인 노동 지수(수요): OLI	123
[그림 5-4] 온라인 노동지수(수요)-국가/직종별: OLI	123

[그림 5-5] 온라인 노동지수(공급)-국가별: Oxford Internet Institute	124
[그림 5-6] 온라인 노동지수 플랫폼 시스템 구성도	125
[그림 5-7] 온라인 노동지수 웹페이지	126
[그림 5-8] 전산업 온라인 노동지수 현황	128
[그림 5-9] 전산업 온라인 노동지수 현황: 2025년	128
[그림 5-10] 산업별 온라인 노동지수 현황	129
[그림 5-11] 산업별 온라인 노동지수 현황: 2025년	130
[그림 5-12] 직종별 온라인 노동지수 현황	131
[그림 5-13] 직종별 온라인 노동지수 현황: 2025년	131
[그림 5-14] 직무별 공고수 합계	134
[그림 5-15] 직무별/월별 공고 변화	134
[그림 5-16] 기술별 공고수 합계(상위 20개)	135
[그림 5-17] 직무별/기술별 공고수 합계(상위 10개)	136

요 약 문

1. 제 목

ICT 주요이슈 분석체계 구축

2. 연구 목적 및 필요성

디지털 전환이 경쟁력의 핵심 원천이 되고 ICT 기술에 기반한 융합으로 다양한 신산업이 등장하면서, 기반이 되는 ICT 산업의 중요성은 점차 더 커지고 있다. 디지털 플랫폼, 인공지능 등 ICT와 관련된 주제들에 대한 관심도와 영향력을 살펴보면 이러한 흐름을 분명하게 느낄 수 있다. 더불어 ICT 기반의 융합과 신산업이 활발해짐에 따라 ICT와 관련된 이슈들도 빠른 속도로 증가하고 있고, 그 종류 또한 다양해지고 있다. 이처럼 다양한 ICT 관련 이슈를 신속하고 효과적으로 파악하기 위해서는 이를 가능케 하는 분석체계를 구축하는 것이 중요하다. 이러한 배경 하에서, 본 연구는 온라인 기반 빅데이터 자료와 통계 자료를 함께 활용하여 ICT 관련 이슈를 시의성 있게 다룰 수 있는 분석 방법론을 개발하고 분석 결과를 제공하는 것을 목표로 한다.

3. 연구의 구성 및 범위

2장에서는 온라인 플랫폼 기반으로 수집한 뉴스기사와 통계자료를 함께 활용하여 ICT와 관련된, 관심도가 높은 이슈들을 선정하고 분석을 수행하였다. 올해 보고서에서는 디지털 헬스케어, 멀티모달 인공지능, 양자컴퓨팅, 개방형 AI 모델 현황을 주제로 선정하였다. 각 주제별로 관련 통계를 활용한 현황 및 전망 분석을 수행하였고, 온라인 자료를 통해 다양한 텍스트 분석을 수행하여 함의점을 제공하였다.

3장에서는 ICT 산업의 경기를 신속하게 파악할 수 있는 대안으로 만든 ICT 뉴스경기 지수 구축 방법론과 주요 결과를 제시하였다. ICT 뉴스경기지수를 통해 확인한 ICT 산업 경기를 요약하여 현황을 제시하였으며, 뉴스기사 제목에 기반한 네트워크 분석을 통해 월별로 주요 이슈를 제시하였다. 또한, ICT 뉴스경기지수 산출에 기반이 되는 ICT 키워드를 재점검하고 최신 기술 발전과 산업 트렌드를 반영하여 최신화 작업을 수행하였다.

4장에서는 ICT 키워드 도출 자동화 체계 구축 방안을 설명하였다. 민간의 관심 흐름을 시의성 있게 파악하기 위해 뉴스 데이터와 딥러닝 방법론을 활용하여 ICT 트렌드를 반영한 키워드를 안정적으로 자동 산출 및 관리할 수 있는 시스템 체계를 마련하였다.

5장에서는 온라인의 구인구직 정보를 활용하여 작성한 온라인 노동지수에 대해 설명하였다. 온라인에서 제공되는 구인구직 정보를 활용한 다양한 선행연구들을 정리하고, 한국형 온라인 노동지수를 구축한 방법 및 산업/직종별 결과를 제시하였다. 특히, IT 개발자 관련 채용공고 변화 분석을 통해 디지털 전환과 생성형 AI 활성화 등에 따른 IT 관련 직무의 수요 재편을 확인하였다.

4. 연구 내용 및 결과

가. ICT 주요 이슈 분석

디지털 헬스케어에 대한 온라인 인식을 분석한 결과, 코로나 19 팬데믹 이후 대기업의 시장 진입이나 정부의 정책 발표 등에 기인하여 언급량이 증가하였고, 기술 중심의 키워드가 점차 빈번하게 등장하며 의료서비스 영역에서 디지털 기술 활용이 가속화되고 있는 것으로 나타났다. 또한 고령화 사회에 건강관리가 핵심 요소로 자리 잡으며 보건의료와 디지털 기술의 접목에 대한 기대감이 높아지며 디지털 헬스케어에 대한 긍정적 인식이 점차 강화되고 있음을 확인하였다.

멀티모달 인공지능의 경우, 생성형 AI의 고도화와 함께 주목을 받으며 1세대(모달리티 연결형), 2세대(LLM 기반 확장형), 3세대(범용 통합형)의 기술 트렌드 변화를 겪는 것으로 보인다. 연구단계에서 실용화 단계의 서비스 형태로 발전하였으며, 인공지능 모달리티 영역의 확장이 이루어지고 있는 것으로 나타났다. 멀티모달 인공지능 기술 발전에 대한 관심이 증가하면서 향후 멀티모달 데이터 구축 및 응용 서비스 고도화를 위한 노력이 필요

할 것으로 보인다.

양자컴퓨팅은 글로벌 기업의 기술발표, 정부의 R&D 투자, 국제 협력 등의 이슈에 기인하여 뉴스 언급량이 급격하게 증가하였으며, 연구단계에서 산업·정책, 투자시장 중심 단계로 담론이 변화하는 양상을 보였다. 향후 산업과 사회 전반에 새로운 가치를 창출할 핵심 기술임에 따라, 지속적인 정책 지원과 산업 생태계 조성을 병행하여 글로벌 경쟁에서 선도적 위치를 확보할 필요가 있는 것으로 보인다.

오픈소스 또는 오픈웨이트 형태로 공개된 인공지능 모델을 의미하는 개방형 AI 모델은 폐쇄형 모델과 달리 구조·가중치·데이터 등을 활용해 자유롭게 커스터마이징과 파인튜닝이 가능하며, 2023년 메타의 LLaMA 공개를 계기로 고성능 생태계가 급속히 확산되었다. 2025년 2월 DeepSeek 모델 발표와 함께 관련 기사량이 급증하였으며, 이후 기술 중심 논의에서 정책·산업 전략 중심의 담론으로 전환되며 국가차원의 산업정책 및 기술패권 논의의 핵심의제로 자리잡음을 보여주었다. 향후 개방형 AI 모델은 AI주권의 실현 수단으로도 주목받는 등 다양한 측면에서 중요도가 더욱 높아질 것으로 전망된다.

나. ICT 뉴스경기지수

온라인 플랫폼을 통해 확보한 뉴스기사에 기반하여 ICT 산업의 경기를 시의성 있게 측정할 수 있는 ICT 뉴스경기지수를 구축하였다. 먼저 ICT 뉴스를 대표할 수 있는 키워드를 선정하여 ICT 관련 뉴스를 정의하고 기사를 수집하였다. 각 기술 및 산업 분야를 대표할 수 있는 키워드 선정을 위해 공신력이 있는 기관과 기업의 자료, 네이버 Datalab, Google Trends 등을 활용하였고, 재점검 및 최신화 작업을 거쳐 최종적으로 20개의 키워드를 선정하였다. 해당 키워드가 포함된 뉴스 기사를 수집 후 KoBERT 모델을 활용하여 공부정 분석을 실시, 월별 ICT 뉴스경기지수를 작성하였다. 이후 월별로 ICT 뉴스의 핵심 이슈들을 네트워크 분석 결과를 도출하여 시각화하였다.

ICT 뉴스경기지수 검토 결과, ICT 산업의 생산과 수출의 움직임과 뉴스경기지수가 유사한 흐름을 보이는 것을 확인하였다. 2025년 3월 이후 ICT 산업은 전반적으로 긍정적인 신호를 보였다. 네트워크 분석 결과 ICT 뉴스 기사에서 AI가 차지하는 비중이 가장 높았으며, 미국 정부의 관세 등이 ICT 관련 중요한 이슈로 등장하였다는 점도 확인할 수 있었다.

또한, ICT 관련 주요 키워드를 시의성 있게 도출하고 상시적으로 점검할 수 있는 자동화 체계 구축 방안을 마련하였다. 뉴스 데이터를 활용한 딥러닝·LLM 기반 체계 구축을 통해 기사 수집부터 임베딩, 군집화, 최종 키워드 추출까지의 전 과정을 자동화하고 분석 기간의 핵심 트렌드를 반영할 수 있을 것으로 기대된다.

다. 온라인 노동지수

국내 주요 구인구직 정보를 바탕으로 온라인 노동지수를 구축하고 분석을 진행한 결과는 다음과 같다.

국내 온라인 플랫폼을 통해 확보한 구인구직 정보를 사용하여 온라인 노동지수를 구축하였고, 산업/직종별로 정보를 세분화하였다. 대표 노동지수 결과를 살펴보면, 2025년 전 산업 온라인 노동지수는 2024년 대비 소폭 감소하는 추세를 보이며, 전체 기간으로도 하향세를 보이고 있다. 산업별 결과에서는 비 ICT서비스업의 경우 2023년부터 타 산업 대비 채용공고 수 증가가 두드러지며, 2025년 기준 최고치를 기록하였다. 반면, ICT 서비스업은 2022년까지 가장 높은 값을 기록하였으나, 2025년에 상대적으로 낮은 수준에 머물렀다. 직종의 경우 2023년 들어 하락 추세를 보이다 2025년 중반 이후 전 직종에서 상승세를 보이고 있으며, IT 직종은 2020년 대비 공고 수가 크게 증가하지 않은 것을 확인할 수 있었다.

더불어, IT 개발자 관련 공고를 분석한 결과 서버/백엔드 개발자 관련 공고 다음으로 AI/ML 엔지니어 공고 수가 많았으며, 이를 통해 AI 관련 직종에 대한 수요가 꾸준히 상승하고 있음을 확인할 수 있었다. 또한 대부분의 직종에서 Python에 대한 기술 수요가 높다는 점을 알 수 있었다.

5. 정책적 활용 내용

본 연구는 온라인에서 수집 가능한 빅데이터와 통계를 활용, 다양한 ICT 관련 주제들을 탐색하고 분석할 수 있는 체계를 구축하였다. 이러한 분석 체계는 빠르게 등장하고 변화하는 ICT 관련 이슈들을 신속하게 파악하고 확인할 수 있는 방법을 제공한다는 점에서 중요한 의미를 가진다. 연구 결과는 ICT 산업의 최신 흐름과 중요 주제들을 파악하고, 이에 대한 정책 대응을 선제적으로 마련하는 데 기여할 것으로 예상된다.

6. 기대효과

본 연구에서 구축한 분석체계는 신속하게 변화하는 온라인 기반의 정보와 기존 통계 정보를 함께 활용하여, 빠르게 변화하는 ICT 이슈들을 파악하고 분석할 수 있도록 한다. 연구의 결과는 디지털 전환의 가속화로 더욱 중요해지고 있는 ICT 관련 주제들을 시의성 있게 확인하고 파악할 수 있는 기반을 제공하는 데 기여할 것으로 기대된다.

SUMMARY

1. Title

Establishment of Analysis System for ICT Issue

2. Purpose and Necessity of Research

The importance of the ICT industry is increasing as digital transformation becomes a core source of competitiveness and diverse new industries emerge through convergence based on ICT technologies. The growing interest in and influence of ICT-related topics, such as digital platforms and artificial intelligence, clearly reflect this trend. Furthermore, as ICT-based convergence and new industries expand, ICT-related issues are increasing rapidly in both volume and variety. To promptly and effectively identify and analyze these diverse ICT-related issues, it is essential to establish an analytical framework that enables such examination. Against this backdrop, this study aims to develop an analytical methodology that integrates online big data and statistical data to examine ICT-related issues in a timely manner and to provide analytical results.

3. Contents and Scope of Research

Chapter 2 presents an analysis of high-interest ICT-related issues by jointly utilizing news articles collected from online platforms and relevant statistical data. In this year's report, digital healthcare, multimodal artificial intelligence, quantum computing, and open AI model developments were selected as the core topics. For each topic, analyses of

current status and future prospects were conducted using relevant statistical indicators, while various text analyses based on online data were performed to derive policy and industry implications.

Chapter 3 outlines the methodology for constructing the ICT News Economic Index as an alternative tool for the timely assessment of business conditions in the ICT industry and presents its key findings. The current state of the ICT industry, as identified through the index, is summarized, and major monthly issues are identified through network analysis based on news article headlines. In addition, the ICT keywords underlying the index were re-examined and updated to reflect recent technological advancements and evolving industry trends.

Chapter 4 describes the construction of an automated system for ICT keyword extraction. To timely capture shifts in public interest, a systematic framework was developed that utilizes news data and deep learning methodologies to stably automate the generation and management of ICT trend-reflective keywords.

Chapter 5 explains the Online Labor Index constructed using online job vacancy and recruitment data. The chapter reviews prior studies that utilized online job market information and presents the methodology for developing a Korea-specific Online Labor Index, along with industry- and occupation-level findings. In particular, an analysis of changes in job postings for IT developers confirms a restructuring of demand for IT-related occupations driven by digital transformation and the diffusion of generative AI.

4. Research Results

4.1. Analysis of Major ICT Issues

The analysis of online perceptions of digital healthcare indicates that, following the COVID-19 pandemic, mentions have increased due to the market entry of large corporations and government policy announcements. Technology-driven keywords have appeared with growing frequency, suggesting that the utilization of digital technologies in

the healthcare service sector is accelerating. In addition, as health management has become a core element in an aging society, expectations regarding the integration of healthcare and digital technologies have risen, confirming a gradual strengthening of positive perceptions toward digital healthcare.

Multimodal artificial intelligence has gained attention alongside the advancement of generative AI and appears to be undergoing technological evolution across three stages: first-generation modality-bridging systems, second-generation large language model (LLM)-based expansion systems, and third-generation general-purpose integrated systems. The technology has progressed from the research phase to practical service-oriented forms, accompanied by the expansion of AI modality domains. As interest in multimodal AI continues to increase, further efforts will be required to develop multimodal datasets and enhance applied services.

Quantum computing has experienced a sharp increase in news coverage, driven by technology announcements from global corporations, government R&D investments, and international cooperation. The discourse has shifted from a research-focused stage to one centered on industry, policy, and investment markets. As it is regarded as a key technology capable of creating new value across industries and society, sustained policy support alongside the development of an industrial ecosystem will be necessary to secure a leading position in global competition.

Open AI models, defined as artificial intelligence models released in open-source or open-weight formats, differ from closed models in that they allow flexible customization and fine-tuning through access to model architectures and weights. Following Meta's release of LLaMA in 2023, a high-performance ecosystem expanded rapidly. In February 2025, the announcement of the DeepSeek model led to a sharp increase in related media coverage, and the discourse subsequently shifted from technology-centered discussions to policy- and industry strategy-oriented debates. This shift suggests that open AI models have emerged as a key agenda in national-level industrial policy and technological sovereignty discussions. Looking ahead, the importance of open AI models is expected to

further increase, as they are also being recognized as a means of realizing AI sovereignty.

㉑. ICT News Economic Index

Based on news articles collected through online platforms, an ICT News Economic Index was developed to measure economic conditions in the ICT industry in a timely manner. ICT-related news was first defined by selecting representative keywords, and relevant articles were subsequently collected. To identify keywords representing major technologies and industrial sectors, data from authoritative institutions and corporations, as well as Naver DataLab and Google Trends, were utilized. Following review and updating processes, a final set of 20 keywords was selected. News articles containing these keywords were then analyzed using the KoBERT model to conduct sentiment analysis, and a monthly ICT News Economic Index was constructed. In addition, key monthly ICT news issues were identified through network analysis, and the results were visualized.

The results show that the ICT News Economic Index exhibits trends similar to movements in ICT industrial production and exports. Since March 2025, the ICT industry has shown overall positive signals. Network analysis results indicate that AI accounted for the largest share of ICT-related news coverage, and that U.S. government tariffs emerged as a significant ICT-related issue.

Furthermore, a framework for establishing an automated system was developed to enable the timely extraction and continuous review of key ICT-related keywords. By building a deep learning- and LLM-based system utilizing news data, it is expected that the entire process, from article collection and embedding to clustering and final keyword extraction, can be automated, thereby reflecting key trends during the analysis period.

㉒. Online Labor Index

Based on major domestic job posting data, an Online Labor Index was constructed and analyzed, and the results are as follows.

Job posting information collected from domestic online platforms was used to develop the Online Labor Index, with further segmentation by industry and occupation. The overall Online Labor Index across all industries in 2025 shows a slight decrease compared to 2024 and also demonstrates a downward trend over the entire period. At the industry level, non-ICT service sectors have exhibited a notable increase in the number of job postings since 2023 compared to other industries, reaching their highest level in 2025. In contrast, ICT service industries recorded the highest values through 2022 but remained at relatively lower levels in 2025. By occupation, job postings showed a declining trend beginning in 2023 but shifted to an upward trend across all occupations from mid-2025 onward. Notably, the number of IT-related job postings did not increase substantially compared to 2020.

Further analysis of IT developer-related postings indicates that, after server/backend developer positions, AI/ML engineer roles accounted for the next largest share of postings. This confirms that demand for AI-related occupations has been steadily increasing. In addition, high demand for Python-related skills was observed across most occupational categories.

5. Policy Suggestions for Practical Use

This study developed a system for exploring and analyzing various ICT-related topics using big data and statistics collected from online sources. This analytical framework is significant as it provides a method for quickly identifying and understanding the rapidly emerging and changing ICT-related issues. The findings are expected to contribute to identifying the latest trends and key topics in the ICT industry and proactively formulating policy responses.

6. Expectations

The analytical framework developed in this study combines rapidly changing online-based information with existing statistical data, enabling the identification and analysis of fast-evolving ICT issues. The results of this research are expected to contribute to providing a foundation for timely identification and understanding of ICT-related topics, which are becoming increasingly important due to the acceleration of digital transformation.

CONTENTS

Chapter 1. Introduction

Section 1. Necessity and Purpose of the Study

Chapter 2. Analysis of Major ICT Issues

Section 1. Overview

Section 2. Changes in Online Perceptions of Digital Healthcare

Section 3. Analysis of Technological Trends in Multimodal AI

Section 4. Changes in Media Perceptions of Quantum Computing Keywords

Section 5. Analysis of Trends in Open AI Models

Chapter 3. Development and Advancement of the ICT News Economic Index

Section 1. Overview

Section 2. Methodology for ICT News Economic Index

Section 3. Results of the ICT News Economic Index Development

Section 4. Identification of ICT Keywords Reflecting Latest Domestic Trends

Chapter 4. Establishment of an Automated Framework for ICT Keyword Extraction

Section 1. Overview

Section 2. Methodology for Establishing an Automated Keyword Extraction
Framework Based on Domestic ICT-Related News

Section 3. Results and Evaluation

Chapter 5 Development of the Online Labor Index

Section 1. Overview

Section 2. Development of a Labor Index Using Online Information: International
Cases

Section 3. Methodology for Developing the Online Labor Index

Section 4. Analysis of the Online Labor Index

Section 5. Analysis of IT Occupation Job Postings

References

제1장 서론

제1절 연구의 필요성 및 목적

- 디지털 전환의 확대에 따라 핵심 기반으로서 ICT 산업과 기술의 역할이 중요해지고 있으며, ICT 관련 주제들의 영향력이 점차 커지고 있음
 - 전자상거래와 온라인 교육 등 디지털 플랫폼에 대한 의존도 증대, 인공지능의 범용화, IoT 등에 기반한 스마트제조 등 주요 혁신은 ICT와 연결되어 있음
 - ICT와 관련된 이슈들이 시대의 발전을 상징하는 지표로 자리매김하고 있음
- ICT 기술의 활용과 이에 따른 신산업 창출이 가속화되면서 경제사회적으로 중요한 주제들이 빠르게 생성되고 있기에, 이를 신속하게 다루고 분석할 수 있는 체계가 필요
 - 전통적인 조사방식의 경우 신뢰성 있는 결과를 제공하지만, 시의성에서는 약점을 보이기에 최신 이슈를 발빠르게 대응하기에는 어려움이 있음
 - 신속하게 변화하는 ICT 관련 이슈들을 다루기 위해서는 빠른 대응을 통해 현황을 파악할 수 있는 분석 체계가 필요
- 이러한 필요성에 근거하여 본 연구에서는 온라인 기반의 빅데이터 자료를 활용하여 ICT 관련 주제들을 시의성 있게 다룰 수 있는 방법론을 구축하고, 분석 결과를 제공하는 것을 목표로 함
 - 온라인상의 뉴스 기사, 소셜 데이터 등의 실시간 자료를 사용하여 ICT와 관련된 이슈들을 빠르게 파악하고 관련된 동향을 분석
 - 기존 통계와 텍스트 기반의 분석 방법을 연결하여 이슈와 관련된 세부 현황들을 도출하고, 주요 함의를 파악하여 제공
 - 국가데이터처, 한국은행 등 주요 통계 제공 기관의 open API를 활용, 중요도가 높은 ICT 산업 통계들을 정기적으로 업데이트하고 최신 정보를 빠르게 확인 가능한 시스템을 구축

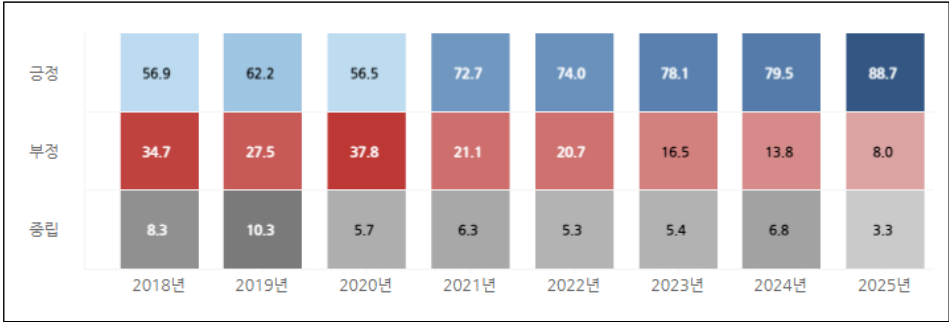
- 온라인 구인구직 정보를 사용해서 온라인 노동지수를 구축하고 주요 산업/직종 별로 고용 현황을 빠르게 확인할 수 있는 보조 지표 제공
- ICT 관련 뉴스를 기반으로 ICT 뉴스경기지수를 도출하고, 네트워크 분석을 통해 최신 흐름과 주요 키워드를 분석
- o 본 연구를 통해 도출한 결과들은 온라인 기반의 빅데이터를 활용, ICT의 중요 이슈들을 다양한 방법으로 살펴볼 수 있는 분석체계 구축에 기여
- o 연구의 성과를 통해 향후 더 고도화된 온라인 빅데이터 분석 방법을 만들어 갈 수 있을 것으로 기대됨

제 2 장 ICT 주요 이슈 분석

제1절 개 요

- 본 장에서는 ICT와 관련된 중요 이슈들을 발굴 및 분석하기 위해 마련한 체계를 설명하고, 이를 통해 분석한 결과들을 제시하였음
- ICT와 관련되어 중요도가 높은 최신 이슈를 확인하고 현황을 파악하기 위해 온라인 플랫폼에서 제공되는 뉴스 기사를 중심으로 분석을 수행하고 관련 통계들을 수집
- 분석 주제와 관련된 최신 뉴스 기사들을 수집한 후, 뉴스 기사의 본문을 활용하여 텍스트 분석을 수행하였음
 - 텍스트 분석에 사용된 방법에는 빈도 분석, 감성분석, 연관어 분석 등이 포함
 - 빈도 분석은 문서에서 가장 자주 등장하는 단어들을 식별하는 기본적인 방법으로, 분석 주제의 핵심 키워드를 파악하고 시간에 따른 키워드의 흐름 변화를 추적할 수 있음
 - 감성분석은 문서에 등장하는 단어들의 긍정/부정 비율을 계산하여, 해당 주제에 대한 대중이나 언론의 인식을 확인할 수 있게 해줌
 - 연관어 분석은 특정 단어와 함께 자주 등장하는 단어들을 찾는 방법으로, 단어간 연관성을 확인하여 주제의 흐름 및 이슈의 변화를 파악할 수 있음
 - 분석을 위해 파이썬에 기반한 자체 프로그램을 개발하였으며, 일부 분석에서는 썬트렌드에서 제공하는 기능을 활용함(<https://some.co.kr/>)

[그림 2-1] 감성분석 예시



출처: ITSTAT 이슈포커스, “디지털 헬스케어에 대한 온라인 인식 변화”

(https://www.itstat.go.kr/itstat/kor/board/BoardDetail.html?board_seq=12097&rootId=2010003&searchType=searchAll&pageIndex=1&board_class=BOARD27&menuId=2010131)

[그림 2-2] 연관어 분석 예시



출처: ITSTAT 이슈포커스, “양자컴퓨팅 키워드의 언론 인식 변화”

(https://www.itstat.go.kr/itstat/kor/board/BoardDetail.html?board_seq=12242&rootId=2010003&searchType=searchAll&pageIndex=1&board_class=BOARD27&menuId=2010131)

- ICT 산업의 중요 통계들을 신속하게 분석하기 위한 방안으로, 국가데이터처와 한국은행의 open API를 통한 업데이트 시스템 구축
 - 국가데이터처와 한국은행은 각 기관에서 제공하는 주요 통계들의 활용 편의성을 높이기 위해 open API를 제공
 - open API는 통계들의 업데이트, 확장, 분석을 자동화할 수 있는 기반이 되기에 분석의 시의성을 향상시키는데 기여할 수 있음
 - ICT의 산업별 매출액, 고용 등 중요한 지표들을 open API로 저장할 수 있는 시스템을 구축하여 활용

[그림 2-3] KOSIS open API 제공 서비스 예시

통계목록

통계목록 API를 이용하기 위한 요청변수, 출력정보, 샘플데이터를 제공합니다. 출력포맷으로 JSON과 SDMX를 제공합니다.

상세기능 - URL 생성 + 개발가이드 다운로드

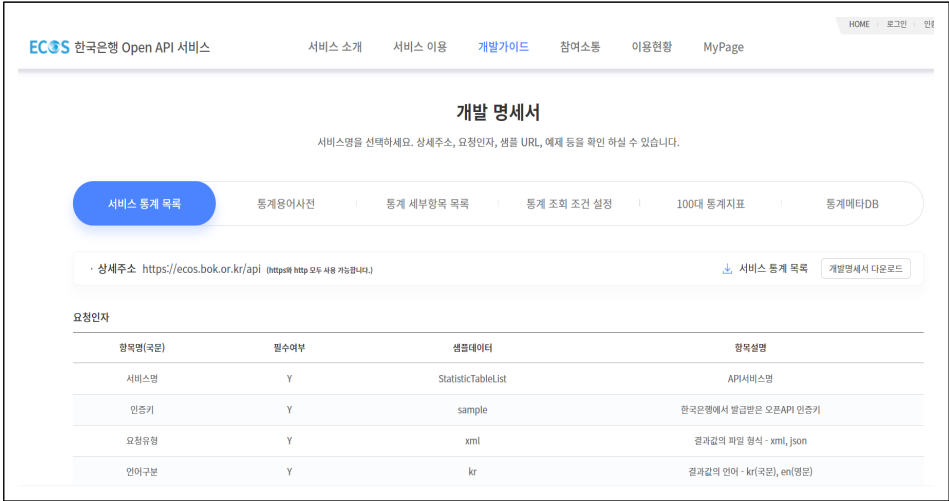
JSON SDMX

요청URL https://kosis.kr/openapi/statisticsList.do?method=getList

요청변수	요청변수	변수타입	설명	비고
apiKey		String	발급된 인증 key	필수
vwCd		String	서비스뷰 코드 - MT_ZTITLE : 국내통계 주제별 - MT_OTITLE : 국내통계 기관별 - MT_GTITLE01 : e-차량차표(주제별) - MT_GTITLE02 : e-차량차표(지역별) - MT_CHOSUN_TITLE : 광복이전통계(1908-1943) - MT_HANKUK_TITLE : 대한민국통계연감 - MT_STOP_TITLE : 작성중지통계 - MT_RTITLE : 국제통계 - MT_BUNHAN : 국외통계 - MT_TMI_TITLE : 대상별통계 - MT_TM2_TITLE : 이슈별통계 - MT_ETITLE : 영문 KOSIS	필수

출처: KOSIS 공유서비스, https://kosis.kr/openapi/devGuide/devGuide_0101List.do

[그림 2-4] 한국은행 open API 제공 서비스 예시



출처: 한국은행 Open API 서비스, <https://ecos.bok.or.kr/api/#/DevGuide/DevSpecification>

- 올해 분석을 위해 ICT와 관련하여 사회적인 관심도가 높은 주제를 중심으로 디지털 헬스케어, 멀티모달 인공지능, 양자컴퓨팅, 개방형 AI 모델 현황을 주제로 이슈 분석을 진행하였음
- 분석을 통해 작성한 원고는 『ICT 통계분석』을 통해 발표하였으며 주요 결과는 ITSTAT의 이슈 포커스에서 시각화하여 제공
 - 본 보고서에서는 원고의 전문을 수록하였음

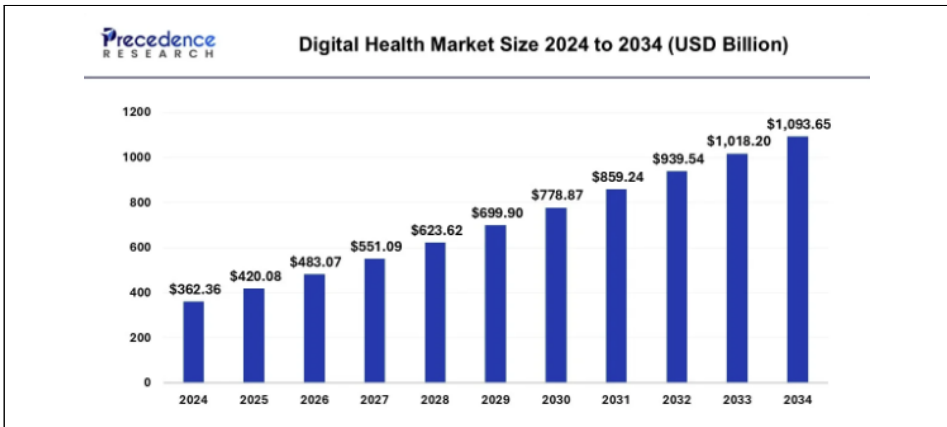
제 2 절 디지털 헬스케어에 대한 온라인 인식 변화¹⁾

1. 디지털 헬스케어의 정의

- (개요) 코로나19가 확산되면서 전 세계적 원격의료에 대한 수요가 증가하였고, 이에 따라 의료분야 ICT 기술이 비약적으로 발전하며 디지털 헬스케어가 미래 신성장 산업으로 부상
 - 시장 조사 기관 Precedence Research에 따르면, 글로벌 디지털 헬스케어 시장 규모는 '24년 3,624억 달러에서 '34년 1조 937억 달러로 연평균 11.7% 성장할 것으로 전망

[그림 2-5] 글로벌 디지털 헬스케어 산업 규모 전망

(단위: 10억 달러)



출처: Precedence Research(2025), "Digital Health Market Size and Forecast 2025 to 2034".

- 이에 본고에서는 디지털 헬스케어 산업의 급성장에 따른 디지털 헬스케어의 주요 동향 및 온라인 인식 변화를 분석하고자 함

1) 분석 내용은 신현호, 손녕선(2025), "디지털 헬스케어에 대한 온라인 인식 변화"로 ICT 통계분석 25-01호에 발표되었음

- (정의) 디지털 헬스케어에 대한 정의는 다양하나 공통적으로 건강 관련 서비스와 의료 IT기술이 융합된 종합 의료서비스 개념을 포함하며, 환자의 치료부터 질병 예방까지 환자 개개인의 고유한 특성에 적합한 맞춤형학(Personalized)을 제공하는 것을 목표로 함

〈표 2-1〉 국내외 디지털 헬스케어 정의

출처	정의
FDA (2020) ²⁾	모바일 헬스케어, 보건정보기술, 웨어러블 기기, 원격의료, 개인 맞춤형 의료 등 다양한 범주를 포괄하는 광범위한 개념으로 정의
WHO (2021) ³⁾	건강 증진을 위해 디지털 기술을 개발하고 활용하는 것과 관련된 지식과 실천의 분야로, 스마트 기기, IoT, 빅데이터, 로봇공학 등 다양한 디지털 기술의 활용 포함
국회보건복지위	지능정보기술과 보건의료 데이터를 활용해 질병 예방, 진단, 치료, 건강 관리, 연구개발, 사후관리 등 국민 건강 증진에 기여하는 일련의 활동과 수단으로 정의
KDI (2021) ⁴⁾	모바일 헬스(웨어러블, 건강 앱), 디지털 보건의료 시스템(건강 정보 저장·교환), 보건의료 분석(정밀 의료, 유전자 검사), 원격의료(ICT 기반) 등 4개 분야로 구분
정보통신산업진흥원(2019) ⁵⁾	헬스케어와 ICT가 융합하여 개인의 건강과 질병을 관리하는 산업 영역으로 정의하며, 모바일 헬스, 원격의료, 전자의료기록(EMR/EHR 시스템) 등으로 구분
한국보건산업진흥원(2020) ⁶⁾	(광의) ICT 기술이 적용된 모든 헬스케어 분야, (협의) 모바일 헬스케어, 원격의료, 인공지능 등이 포함된 헬스케어 분야

2) FDA(2020), "What is Digital Health?," <https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health> (검색일: 2025.04.01).

3) WHO(2021), "Global strategy on digital health 2020-2025", p.39

4) KDI경제정보센터(2021), 「디지털 헬스케어 산업 활성화 정책…발자취와 새로운 도약」, 《해외동향-디지털헬스케어편》, 2021-02호, pp.3~4

5) 문장원·윤형진·선미란(2019.12.11.), 「해외 디지털 헬스케어 규제개선 동향」, 《이슈리포트》, 2019-37호, 정보통신산업진흥원, p.2

6) 김지은 외(2020.12.), 「디지털 헬스 산업 분석 및 전망 연구」, 한국보건산업진흥원, p.6

2. 디지털 헬스케어의 기술 동향

- (기술 동향) 디지털 헬스케어는 AI, 클라우드, 디지털 트윈 등 ICT 기술을 기반으로, 정밀 진단과 개인 맞춤형 치료, 의료 접근성과 서비스 효율성을 향상시키는 방향으로 발전하고 있음
- 기존 모바일 앱과 웨어러블 기기를 활용한 운동량, 혈압, 수면 상태 등의 건강 관리가 일상화되며, 개인의 건강 데이터를 자가 모니터링하는 방식이 보편화 되었음
- 최근 데이터의 수집, 저장, 분석 기술 고도화로 예방 및 정밀 의료 분야에도 점차 활용⁷⁾되고 있으며, 이는 환자 중심의 진료 환경 조성⁸⁾과 보건의료 서비스의 질적 향상에 기여할 것으로 기대

〈표 2-2〉 국내외 디지털 헬스케어 기술 동향

구분	내용
AI 기술 적용 ⁸⁾	혈액 검사, 이미지 등의 데이터 분석을 통한 진단 분석 고도화에 활용되며 그 외에도 신약 후보 물질 제시를 통한 임상시험 과정 효율화
AR/VR 활용 ⁹⁾	외과 수술 훈련, PTSD·공포증 등 정신질환 치료, 만성통증 관리 등에 활용되어 시뮬레이션 기반 몰입형 치료를 통해 심리적 장애 극복과 임상적 효과 향상
클라우드 활용 데이터 통합 ¹⁰⁾	데이터를 안전하게 저장·분석하고, 진단 정보를 통합하여 진단의 정밀성과 진료 효율성을 제고함. 다만 개인정보 보호 및 보안 등의 위험 요소 존재
디지털 트윈 기반 정밀 진단 ¹¹⁾	환자의 장기나 신체를 디지털로 복제해 예방, 진단, 치료 등 전 주기적 시뮬레이션 진단. 현재는 초기 수준의 기술이 활용되고 있으며 향후 모델링 정확도 향상, 윤리적 문제 등이 해결된다면 정밀한 맞춤형 의료 실현 기대 ※ 디지털 트윈은 구현 수준에 따라 모사(복제), 관제, 모의(최적화)의 3단계로 구분 * 1단계 모사: 현실 객체의 외형과 속성을 반영한 정적인 3D 가상 모델을 구현한 수준 * 2단계 관제: 현실 세계와 실시간으로 연동되어 모니터링 및 제어 가능한 수준 * 3단계 모의: 현실과 가상 모델 간 상호작용이 가능하며, 시뮬레이션을 통해 예측 및 의사결정 지원이 가능한 수준

7) 디지털비즈온(2024.07.19.), ““디지털 헬스케어”… 이미 시작된 미래”,
<https://www.digitalbizon.com/news/articleView.html?idxno=2336495>

구분	내용
글로벌 트렌드	<p>세계 최대 IT전시회 CES 2025에서 디지털 헬스 분야가 핵심 주제로 부상, 국내에서 23개 제품이 혁신상 수상. 그 중 최고 혁신상을 수상한 한양대 'TD 스캐어' 이외에도 생성형 AI기반 내시경, 보행지원 웨어러블 로봇, 비침습 약물 주입기 등 다양한 제품 주목 받음</p> <p>※ TD 스캐어: 이명 환자에게 이명 발생 위치를 귀 바깥으로 끄집어내어 인식할 수 있도록 다중 감각을 자극하는 가상환경 제공하여 이명 완화</p>

〈표 2-3〉 CES 2025 디지털 헬스케어 부문 혁신상 수상 목록(일부)

회사명	제품명	주요기능
뉴다이브	버디인 (buddy-in)	자폐, 의사소통 장애아를 위한 디지털 사회 기술 훈련 게임
더웨이브톡	Real-Time Bacteria Sensor for Water	물속 박테리아 오염 감지 자동장치
데카사이트	ARTIS (AR Training for Intervention & Surgery)	침습적 치료 및 수술을 위한 AR 트레이닝 시뮬레이터
메디코스바이오텍	큐어실크 (CureSilk)	상처 사진을 AI앱이 분석 후 거미줄 단백질 봉대 사용
바디프렌즈	Standing Row (SR) 733	독립적 팔, 다리제어, AI헬스케어 서비스 마사지 의자
세라젠크	Home Medicare Bed 2.0 with AI Concierge	건강 모니터링, 측정, AI 비서 연동 스마트 헬스 침대
에이슬립	SleepBoard Powered by AsleepTrack	AI 기술 자동수면 추적, 수면 환경최적화 태블릿

- 8) Martyn Eeles(2023.01.01.), "The Role of AI in Digital Health: How to Turn This Technology Into A Competitive Advantage", <https://www.linkedin.com/pulse/role-ai-digital-health-how-turn-technology-advantage-martyn-eeles>(검색일: 2025.04.01.)
- 9) MassChallenge(2022.05.25.), "Rising Trends in Digital Health: 5 Technologies That Will Define the Future of Healthcare", <https://masschallenge.org/articles/digital-health-trends/>(검색일: 2025.04.01.)
- 10) digitalhealth(2023.06.09.), "Cloud technology in healthcare benefits highlighted in new report", <https://www.digitalhealth.net/2023/06/cloud-technology-in-healthcare-benefits-highlighted-in-new-report>
- 11) 강유진·김주원(2022.11.30.), 「헬스케어 디지털 트윈」, 《KISTEP브리프》, 46호, 한국과학기술기획평가원, p.6

회사명	제품명	주요기능
지브레인	핀스팀(Phin Stim)	무선 뇌임플란트
카멜로테크	카멜레온(Cameleon)	AI 기반 한방 제약 자동화 시스템
한양대학교	TD 스퀘어 (TD Square)	디지털 이명 치료기기

출처: 메디게이트뉴스(2025.01.07.), “국내 디지털 헬스 CES혁신상 수상 제품·서비스 23개”

3. 디지털 헬스케어의 국내외 정책 동향

- (국내 정책 동향) 국내 디지털 헬스케어 정책 동향을 살펴보면, 코로나19 팬데믹을 계기로 디지털 기반의 보건의료 체계 구축이 필수적인 화두로 떠오르면서 다양한 정책 추진을 통한 단계적인 산업 육성 기조가 강화되고 있음
- 특히, '20년 한국판 뉴딜 프로젝트를 기점으로 디지털 전환이 본격화되고 R&D 사업, 바이오 빅데이터 활용 방안 등 새로운 산업 육성에 대한 계획이 주를 이루었으며, '22년 새정부 국정과제에 디지털 헬스케어가 포함되어 신시장 창출을 목표로 공공 중심의 정책과 법·제도 고도화가 추진됨

〈표 2-4〉 국내 디지털 헬스케어 정책 동향

구분	내용
한국판 뉴딜 종합계획 (관계부처 합동, 2020.7.)	<ul style="list-style-type: none"> • 한국판 뉴딜 10대 대표 과제 중 '디지털 뉴딜' 분야에 '스마트의료 인프라'가 포함 • 디지털 기반 스마트병원 구축, 12개 질환별 AI 정밀진단이 가능한 S/W 개발 실증 등
혁신성장 BIG3 추진회의(18차) (기획재정부, 2022.1.)	<ul style="list-style-type: none"> • 바이오헬스 산업 집중육성을 위한 22년 중점 추진 과제에 디지털 헬스케어 포함 • 마이헬스케어 실증 추진, 의료기관 진료 정보 디지털 전환 확대, 100만 명 규모 국가 통합 바이오 빅데이터 구축 등
디지털 헬스케어 산업의 육성 및 지원에 관한 법률안 발의 (산업통상중소벤처 위원회, 2022.2.)	<ul style="list-style-type: none"> • 우수기업 인증, 국가연구개발사업 우대, 조세특례, 기술·관리기준 표준화, 해외진출 지원 내용 포함
디지털 헬스케어 서비스산업 육성 전략 (다부처, 2022.2.)	<ul style="list-style-type: none"> • 디지털 헬스케어 시장 진출 지원 강화, 데이터 기반 융복합 헬스케어 기기 개발, 활성화 기반 조성 등 10대 중점 추진 과제 발표

구분	내용
윤석열 정부 110대 국정과제(2022.4.)	<ul style="list-style-type: none"> 서비스 혁신 및 민간사업 활성화 제도 도입과 지원, 디지털병원, 디지털 의료 전문인력 양성을 위한 의료서비스 혁신 병행 추진 등 산업 육성 관련 내용 포함
디지털 헬스케어 정책방향 발표 ¹²⁾ (보건복지부, 2022.8.)	<ul style="list-style-type: none"> 디지털 기술을 활용한 국민 건강 증진을 목표로 '디지털기반 미래의료 실현', '디지털 헬스케어 신시장 창출', '빅데이터 기반 바이오헬스 산업 혁신' 3대 정책 방향 추진
디지털 헬스케어 진흥 및 보건의료데이터 활용 촉진에 관한 법률안 발의 ¹³⁾ (보건복지부위원회, 2022.10.)	<ul style="list-style-type: none"> 의료데이터의 가명 처리 기준을 명확히 하고 본인 및 제3자에 대한 전송 요구권 도입 데이터 전송지원 시스템 구축, 시범 사업 및 규제샌드박스 운영, 연구개발·인력 양성 포함
데이터 기반 디지털 헬스케어 산업 고도화 추진 ¹⁴⁾ (보건복지부, 2023.6.)	<ul style="list-style-type: none"> 공공 R&D 및 건강보험 데이터 개방·공유 및 EMR 표준화와 유전체 정보 활용 확대를 위한 제도 개선 추진 가명 정보 연구 활용 촉진을 위한 IRB 심의 간소화 및 중개 플랫폼 구축 등
디지털헬스 산업 생태계 ¹⁵⁾ 조성을 위한 4대정책방향 (산업통상자원부, 2024.1.)	<ul style="list-style-type: none"> 디지털 헬스 산업 생태계 조성을 위해 혁신 서비스 상용화, 빅데이터 활용, 융합인재 양성, 민간 투자·수출 지원 등

출처: (한국판 뉴딜 종합계획 ~ 윤석열 정부 110대 국정과제) 연미영(2023.06.30.), 「디지털 헬스케어 정책 현황: 소비자 참여 기반의 디지털 헬스케어 활성화를 위한 검토」, p.5 (기타 각주 참고)

- (해외 정책 동향) 최근 해외 주요국의 정책 동향 역시 고령화, 만성질환 증가, 기술 진보에 따른 보건의료 수요 변화에 대응하기 위해 디지털 기반의 의료 혁신을 추진하고 있음

12) 보건복지부 보도자료(2022.06.27.), “데이터 기반으로 바이오·디지털헬스 글로벌 도약 추진”.

13) 강기운 외 2022, “디지털 헬스케어 진흥 및 보건의료데이터 활용 촉진에 관한 법률안 (의안번호 17775)”.

14) 보건복지부 보도자료(2023.06.01.), “국민건강 증진을 위해, 디지털 헬스케어 산업을 육성 하겠습니다”.

15) 산업통상자원부 보도자료(2023.01.30.), “디지털헬스 산업 생태계 조성을 위한 4대 정책 방향 제시”.

- 특히, 신규 의료기기 및 의약품에 대한 인허가 제도 정비, 건강보험 연계, 의료 데이터 활용 및 디지털 치료제 도입 등 제도 기반 강화에 집중
- 국내의 경우 코로나19 팬데믹을 계기로 디지털 헬스케어 정책이 본격화된 반면, 미국과 유럽에서는 그 이전부터 관련 제도와 전략을 선제적으로 마련해 기술 상용화 및 시장 조성을 적극 추진해 온 점에서 차이를 보임

〈표 2-5〉 해외 디지털 헬스케어 정책 동향

국가	구분	내용
미국	21세기 치료법 제도 등 발표 (2016)	신의료기술의 개발 및 인허가 가속화를 위한 ① 21세기 치료법(더 나은 치료 선택과 접근성 제공을 위한 규제 완화), ② Pre-Cert 프로그램(사전인증 기업에 대한 인허가 제출 의무 면제), ③ 획기적 의료기기 프로그램(혁신의료기기 개발 기간 단축) 도입
	FDA 디지털헬스 가이드라인 발표 (2017)	AI 의료기기의 안전성 개선을 위한 사전 변경관리계획, 코로나 시기 한시적 규제 전환 계획, 사이버 보안 기준, 약물 임상시험의 환자 데이터 활용 기준 포함
독일	디지털 헬스케어법 제정 (2019)	디지털 건강 앱(DiGA)의 시장 진입 촉진을 위한 법안 제정 및 패스트트랙 제도로 디지털 치료제 시장 활성화, 데이터 보안 기준 강화 등 ※ DiGA: 의사의 처방과 건강보험 적용이 가능한 인증 디지털 치료 앱
프랑스	개인건강 계획 발표 (2018)	환자 중심·품질 기반 의료체계 구축, 보건의료 디지털 전환, 지역 기반 의료 협업 체계 강화, 의료 인프라 재구축 내용 포함
영국	생명과학 비전 발표 (2021)	암, 치매 등 주요 질환의 예방·진단·치료를 위한 디지털 기술 및 임상 시험 가속화 7대 임무 설정, 산업계와 정부의 정밀 의료, 백신, 면역 치료 분야 공동 지원 내용 포함
	전장 유전체 분석 정책 추진	2012년 '10만 게놈 프로젝트'를 통한 유전체 해독 완료 후 특정 질환 중심 전국민 전장 유전체 분석 확대, 관련 인프라와 개인정보 보호, 생명과학에 집중 투자 내용 포함
일본	건강사회 실현 계획 발표 (2023)	헬스케어 DX(디지털 전환) 가속화 AI·IOT를 활용한 예방 중심 사회 구축, 고령자 및 만성 질환자 대응 강화 등

자료: 한국산업기술진흥원(2024.04.12.), "2024년 산업기술환경예측(디지털헬스케어)", pp.50~59.

4. 텍스트 분석 개요

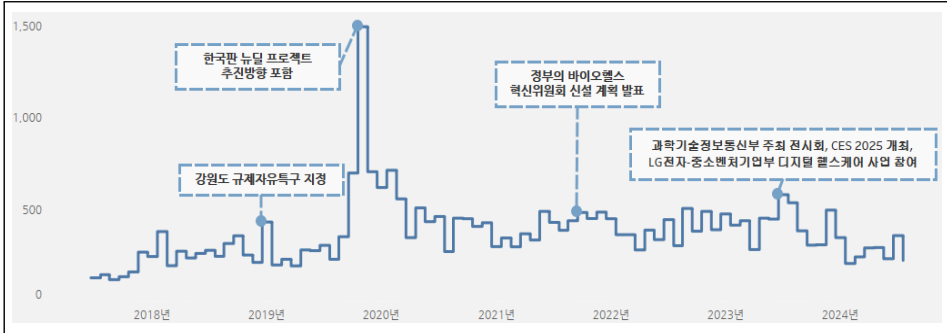
- (분석 목적) 코로나19 팬데믹 이후 신성장 산업으로 부상하고 있는 디지털 헬스케어에 대해 SNS 빅데이터 분석 서비스를 제공하는 썬트렌드를 활용하여 온라인에서 어떻게 언급되고 있는지를 분석하고자 함
- (분석 대상) '18년 1월부터 '25년 2월까지 썬트렌드가 제공하는 매체(뉴스, 블로그, 트위터, 인스타그램, 커뮤니티)의 디지털 헬스케어 관련 키워드의 언급 내용을 분석 대상으로 함
 - 분석을 위해 '원격의료', '모바일헬스', '보건의료분석학', '디지털보건의료시스템', '전자의료기록'을 '디지털 헬스케어' 키워드로 선정하였음
 - ※ 디지털 헬스케어 정의의 다양성을 고려하여, 구분된 범주 가운데 핵심 키워드 도출
 - 언급량 분석 결과는 뉴스 매체를 대상으로 하며, 매체별 언급량 및 텍스트 분석 결과(연관어, 감성분석)는 전체 매체를 대상으로 함
- (분석 내용) 디지털 헬스케어 관련 키워드가 온라인에서 얼마나 언급되고 있는지 그 추이를 살펴본 후, 해당 키워드에 대한 연관어 및 감성분석 결과를 통해 주요 이슈를 분석하고자 함

5. 디지털 헬스케어 언급 추이

- (뉴스 언급량) 디지털 헬스케어 뉴스 언급량은 코로나19 팬데믹 시기인 '20년 최대 건수(6,957건)를 기록하였으며, '18년 2,529건 수준에서 '24년 4,212건 언급되며 지속적으로 상승하는 추세를 보이고 있음

(그림 2-6) 디지털 헬스케어 뉴스 언급량 추이

(단위: 건)



※주1: 2018년 1월부터 2025년 2월까지 자료를 활용하여 분석

※주2: '원격의료', '모바일헬스', '보건의료분석학', '디지털보건의료시스템', '전자의료기록'을 '디지털 헬스케어' 키워드로 선정

※자료: 썬트렌드

- (뉴스 언급량 추이 변화 원인) 디지털 헬스케어 관련 뉴스 언급량이 증가한 시점들은 대기업의 시장 진입 및 정부의 정책 발표 등에 기인한 것으로 보임
 - '20년 2월 국내 코로나 첫 확진자 발생 이후 디지털 헬스케어 관심도가 증가함에 따라 언급량이 대폭 상승하기 시작했으며, '20년 5월 한국판 뉴딜 프로젝트 추진 방향에 디지털 헬스케어 내용이 포함¹⁶⁾되며 가장 많은 언급량을 기록
 - '22년 4월 정부의 바이오헬스 혁신위원회 신설 발표¹⁷⁾ 이후 디지털 헬스케어 정책 관련 논의가 활발해짐에 따라 뉴스 기사 언급량 증가
 - '24년 1월 LG전자가 중소기업벤처기업부의 디지털 헬스케어 사업에 참여¹⁸⁾함에 따라 뉴스 기사 언급량이 소폭 증가

16) 뉴스1(2020.05.05.), "이번주 '한국판 뉴딜' 밑그림 공개...디지털·의료·일자리 키워드", <https://www.news1.kr/economy/policy/3925148>

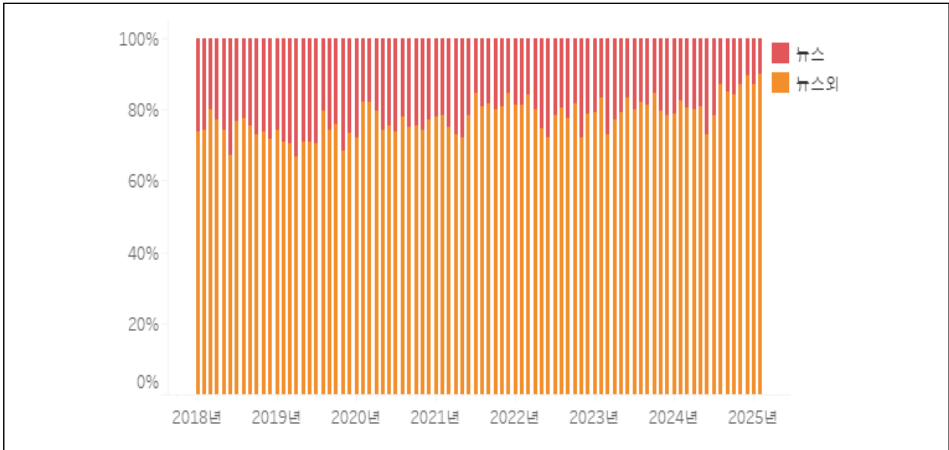
17) 이데일리(2022.04.25.), "바이오헬스 한류시대' 연다...제약바이오혁신위 신설", <https://n.news.naver.com/mnews/article/018/0005199828?rc=N&ntype=RANKING>

18) 조선비즈(2024.01.30.), "LG전자, 중기부와 '디지털 헬스케어' 스타트업 발굴 나선다", <https://n.news.naver.com/mnews/article/366/0000966526>

- (매체별 언급량) 디지털 헬스케어 대한 매체별 언급량 비중 추이를 살펴보면, 뉴스 보다는 뉴스 이외의 매체(블로그, 트위터, 인스타그램, 커뮤니티)의 언급 비중이 상대적으로 높은 것으로 나타남

[그림 2-7] 디지털 헬스케어 매체별 언급 비중 추이

(단위: %)



※주1: 2018년 1월부터 2025년 2월까지 자료를 활용하여 분석

※주2: 뉴스 외 매체는 뉴스를 제외한 블로그, 트위터, 인스타그램, 커뮤니티를 대상으로 분석

※자료 : 썬트렌드

- 특히 '21년 이후 뉴스 외 매체에서의 언급 비중이 증가하는 추세로 나타나는데, 이는 코로나19 팬데믹 이후 디지털 헬스케어에 대한 정책과 기업들의 투자가 늘어남에 따라 일반인들의 관심도가 높아진 것으로 보임

6. 연관어 분석

- (분석 결과) 디지털 헬스케어와 관련된 연관어를 구체적으로 살펴보기 위해 썬트렌드 분석 결과(뉴스, 뉴스 외 매체 포함)를 활용하여 연관어 분석을 수행

〈표 2-6〉 연도별 디지털 헬스케어 관련 텍스트 분석 결과

(단위: 건)

2018년	2019년	2020년	2021년	2022년	2023년	2024년	2025년
의료 (2,023)	의료 (1,841)	코로나 (15,525)	코로나 (5,008)	코로나 (6,398)	인공지능 (7,230)	인공지능 (6,743)	인공지능 (2,319)
성장 (1,738)	계획 (1,679)	의료 (6,850)	투자 (4,569)	건강 (4,187)	의료 (5,495)	건강 (4,731)	건강 (1,315)
정책 (1,452)	인공지능 (1,558)	경제 (4,893)	의료 (3,995)	의료 (4,139)	건강 (4,372)	의료 (4,494)	시스템 (1,049)
산업혁명 (1,340)	성장 (1,536)	투자 (4,499)	인공지능 (3,405)	투자 (4,072)	시스템 (4,054)	시스템 (3,787)	의료 (979)
건강 (1,285)	지원 (1,479)	시스템 (4,326)	건강 (3,367)	글로벌 (3,728)	글로벌 (4,048)	글로벌 (3,222)	로봇 (956)
연구 (1,266)	연구 (1,457)	정책 (4,074)	시스템 (3,238)	연구 (3,562)	연구 (4,018)	연구 (2,925)	ces (928)
인공지능 (1,189)	산업혁명 (1,403)	인공지능 (3,897)	성장 (3,210)	인공지능 (3,526)	투자 (3,674)	투자 (2,877)	글로벌 (796)
투자 (1,172)	투자 (1,360)	건강 (3,452)	글로벌 (2,968)	원격진료 (3,228)	성장 (3,312)	성장 (2,680)	성장 (716)
일자리 (1,013)	건강 (1,280)	글로벌 (3,274)	연구 (2,666)	성장 (3,029)	계획 (3,112)	건강 (2,084)	건강관리 (656)
규제완화 (871)	글로벌 (1,248)	뉴딜 (2,875)	로봇 (2,259)	메타버스 (1,828)	로봇 (2,794)	계획 (2,022)	혁신 (653)
개혁 (841)	빅데이터 (803)	원격진료 (2,873)	원격진료 (1,931)	로봇 (1,793)	코로나 (1,980)	발전 (1,891)	트렌드 (589)
원격진료 (717)	스마트시 터 (784)	비대면 (2,756)	클라우드 (1,819)	클라우드 (1,427)	임상 (1,837)	로봇 (1,630)	발전 (539)

※주1: 2018년 1월부터 2025년 2월까지 자료를 활용하여 분석

※주2: 2025년은 1-2월간 언급된 기사로 분석한 자료로, 이전 자료와 건수를 직접 비교 시 유의

※주3: 분석 매체는 뉴스, 블로그, 트위터, 인스타그램, 커뮤니티를 포함

※자료: 썬트렌드

- (키워드 변화) 분석 기간 초기에는 ‘성장’, ‘산업혁명’, ‘규제완화’ 등 산업 발전에 대한 기대감을 반영한 키워드가 포함되었으며, 후기에는 ‘인공지능’, ‘로봇’ 등 기술 중심의 키워드가 포함되며 의료서비스 영역에서 디지털 기술 활용이 가속화되고 있는 것으로 보임
- ‘18년 8월 정부의 원격의료 규제 개선 언급 및 ‘19년 7월 강원도가 디지털 헬스케어 규제 자유 특구로 선정됨에 따라 ‘계획’, ‘규제완화’와 같은 키워드가 상위에 언급
- ‘21년 12월 대선 후보 공약에 디지털 헬스케어 분야가 포함¹⁹⁾되며 대선 기간 전후로 관련 언급이 증가하며 ‘투자’, ‘성장’과 같은 키워드 관련도가 높은 것으로 분석
- 이외에도 ‘인공지능’ 키워드는 전 기간에 걸쳐 상위에 언급되고 있어 의료와 ICT 기술의 접목에 대한 관심도가 증가
- ‘25년에는 ‘CES’와 ‘트렌드’ 키워드가 상위에 언급되었는데, 이는 미국소비자기술회협(CTA)에서 개최한 세계가전전시회(CES)에서 디지털 헬스케어가 핵심 기술 트렌드 중 하나로 선정²⁰⁾되었기 때문으로 보임

7. 감성분석

- (분석 결과) 썬트렌드 자료를 활용하여 디지털 헬스케어에 대해 일반인들이 느끼는 감성(긍정, 부정, 중립) 결과를 연도별로 분석하였음

〈표 2-7〉 디지털 헬스케어 키워드 감성분석 결과

(단위: %)

	2018년	2019년	2020년	2021년	2022년	2023년	2024년	2025년
긍정	56.9	62.2	56.5	72.7	74.0	78.1	79.5	88.7
부정	34.7	27.5	37.8	21.1	20.7	16.5	13.8	8.0
중립	8.3	10.3	5.7	6.3	5.3	5.4	6.8	3.3

※주1: 2018년 1월부터 2025년 2월까지 자료를 활용하여 분석

※주2: 진한 색에 가까울수록 비중이 더 높다는 것을 의미

※자료: 썬트렌드

19) 프레시안,(2022.01.24.), “이재명·윤석열·안철수·심상정, 포스트 코로나 체제의 비전은?”, <https://n.news.naver.com/mnews/article/002/0002228489>

20) 머니투데이(2025.01.07.), “‘모빌리티·디지털헬스, 그리고 “...CES 2025 휩쓴 ‘이 기술’”, <https://n.news.naver.com/mnews/article/008/0005137693>

- 디지털 헬스케어에 대한 감성분석 결과, '18년 긍정적 인식 비율이 57%를 기록하였으나, 이후 긍정적 인식 비중이 꾸준히 증가하며 '25년 89% 최고치를 기록
- '18년 정부의 원격의료 제도 도입 논의²¹⁾에 대해 의료 영리화 및 진료 안전성 등의 우려가 지속되며 긍정적 인식 비율이 57%에 그친 것으로 나타남
- '20년~'22년 코로나19 팬데믹 상황 속 비대면 서비스 확대와 ICT 기술 발전으로 디지털 헬스케어 산업이 성장²²⁾함에 따라 긍정적 비율이 높아진 것으로 분석
- '23년 이후 디지털 헬스케어에 대한 긍정적 인식이 꾸준히 상승하여 '25년 89%로 나타났는데, 이는 고령화 사회에 건강관리가 핵심 요소로 자리 잡으며 보건의료와 디지털 기술이 접목된 다양한 서비스²³⁾에 대한 관심도와 기대감이 높아졌기 때문으로 보임

8. 결론

- 본고에서는 코로나19 팬데믹 이후 신성장 산업으로 부상하고 있는 디지털 헬스케어에 대한 온라인 인식 변화에 대해 분석하였음
- 언급량 분석 결과, '18년 2,529건 수준에서 '24년 4,212건이 언급되며 지속적으로 상승하는 추세를 보이고 있음
 - (뉴스 언급량) 디지털 헬스케어 키워드의 뉴스 언급량이 증가한 시점들은 대기업의 시장 진입 및 정부의 정책 발표 등에 기인
 - (매체별 언급량) '디지털 헬스케어' 키워드는 뉴스보다 뉴스 이외의 매체에서의 언급이 상대적으로 높은 비중을 차지하며, 특히 '21년 이후 뉴스 외 매체의 언급 비중이 더욱 증가하는 것으로 나타남

21) 경향신문(2018.11.12.), "홍영표 원내대표 "원격의료 도입, 정기국회 내 처리",
<https://www.khan.co.kr/article/201811122218005>

22) 국민일보,(2022.02.06.), "코로나, 디지털헬스케어에 폭발력 더하다... IT기업·통신사 '눈독'",
<https://n.news.naver.com/mnews/article/005/0001503247>

23) 한겨레(2024.11.28.), "스마트 헬스케어, 건강관리 해결책 '주목'",
<https://n.news.naver.com/mnews/article/028/0002718409>

- 연관어 분석 결과, 분석 기간 초기에는 ‘성장’, ‘산업혁명’, ‘규제완화’ 등 산업 발전에 대한 기대감을 반영한 키워드가 포함되었으며, 후기에는 ‘인공지능’, ‘로봇’ 등 기술 중심의 키워드가 포함되며 의료서비스 영역에서 디지털 기술 활용이 가속화되고 있는 것으로 보임
- 감성분석 결과, '18년 원격의료 제도 도입 논의에 대해 의료 영리화 및 진료 안전성 등 우려로 긍정적 인식 비율이 57%로 나타났으나, 이후 지속적으로 상승하며 '25년 89%로 최고치를 경신
 - － 고령화 사회에 건강관리가 핵심 요소로 자리 잡으며 보건의료와 디지털 기술의 접목에 대한 기대감이 높아졌기 때문으로 보임
- 전 세계적으로 디지털 헬스케어의 수요가 급증함에 따라 다양한 정책이 추진되고, ICT 기술과 의료 분야의 융합에 대한 기대감으로 긍정적인 인식도 점차 확대되는 등 미래의 신성장 산업으로 더욱 주목받을 것으로 예상됨
 - － 이를 위해, 정부의 산학연 중심 전문 인력 양성과 산업 활성화를 위한 중장기적 정책이 논의될 필요가 있으며, 인공지능 기술 발전에 따라 동반되는 데이터 보안 및 윤리적 문제에 대한 다양한 해결 방안도 마련할 필요가 있음

제 3 절 멀티모달 인공지능 기술 트렌드 분석²⁴⁾

1. 멀티모달 인공지능의 정의 및 기술 트렌드

- (개요) 본 절에서는 생성형 인공지능의 고도화와 함께 주목받고 있는 멀티모달 인공지능(Multimodal AI)의 기술 트렌드를 분석하고 온라인에서 언급되는 현황을 살펴보았음
- (정의) 멀티모달 인공지능은 텍스트, 이미지, 오디오, 비디오 등 여러 모달리티(데이터 유형)의 정보를 처리하고 통합할 수 있는 머신러닝 모델을 의미²⁵⁾. 다양한 형태의 데이터를 동시에 처리하고 통합하여 인간과 유사한 인지 능력을 갖춘 인공지능 시스템을 구현하는 기술임
 - 멀티모달 인공지능의 목표는 다양한 데이터 유형의 정보를 효과적으로 통합·처리함으로써, 각 모달리티의 강점을 살리고 한계를 보완할 수 있는 모델을 구축하는 것임
- (핵심과제) Baltrušaitis, T. et al.(2019)은 멀티모달 인공지능의 핵심과제(Core-challenges)를 아래와 같이 5가지 유형으로 분류하고 각 핵심과제의 도전요소들을 도출
 - 멀티모달 인공지능의 5가지 핵심과제는 기능 구현의 목표일 뿐만 아니라, 각 과제를 해결하는 접근 방식에 따라 모델의 구조와 학습 전략이 결정되므로 모델 설계의 기준으로 작용
 - 따라서, 핵심과제에 대한 이해는 전체적인 기술 구조와 발전 방향을 파악하는 데에 도움이 됨

24) 분석 내용은 박지원, 손영선(2025), “멀티모달 인공지능 기술 트렌드 분석”으로 ICT통계 분석 25-02호에 발표되었음

25) IBM(2024.7.15.), “What is multimodal AI?”,
<https://www.ibm.com/think/topics/multimodal-ai>. (접속일: 2025.5.19.)

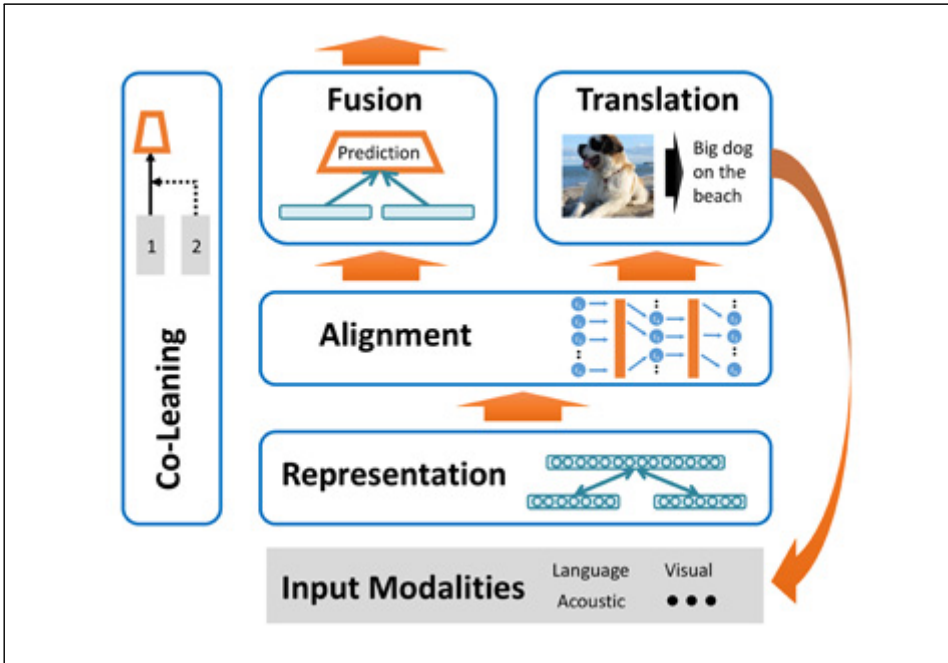
〈표 2-8〉 멀티모달 인공지능의 핵심과제 5가지 분류

유형	내용	도전 요소
표현 (Representation)	멀티모달 데이터의 상호보완성과 중복성을 활용할 수 있는 방식으로 데이터로 표현하고 요약하는 것	데이터 표현 방식의 이질성
정렬 (Alignment)	서로 다른 모달리티의 구성 요소 간 직접적인 관계를 파악하고 매칭하는 것	장거리 의존성 (long-range dependency), 정보 간 모호성 처리 필요
변환 (Translation)	한 모달리티의 데이터를 다른 모달리티로 변환하는 것	모달리티 간 관계의 주관성 및 비결정성
융합 (Fusion)	두 개 이상의 모달리티 정보를 결합하여 예측을 수행하는 것	모달리티 간 정보 품질 및 노이즈 특성 차이 존재
공동학습 (Co-learning)	한 모달리티에서 학습한 지식을 다른 모달리티에 전이하는 것	특정 모달리티의 정보가 부족한 경우 성능 보원에 효과적

자료: Baltrušaitis, T. et al.(2019)을 활용하여 저자 번역 및 재구성

- 핵심과제들은 서로 밀접하게 연관되어 있으며, 먼저 서로 다른 모달리티의 정보를 공통된 공간에 매핑하는 ‘표현(Representation)’ 단계가 이루어짐
- 표현된 정보를 모달리티의 구성 요소 간 ‘정렬(Alignment)’한 다음 다른 모달리티로 ‘변환(Translation)’하거나 정렬된 정보를 바탕으로 ‘융합(Fusion)’이 이루어짐
- 특정 모달리티에서 학습된 지식을 다른 모달리티로 전이하거나 상호보완하는 ‘공동학습(Co-learning)’ 과정을 통해 모델의 적응력과 학습 효율성을 높임

[그림 2-8] 멀티모달 인공지능의 핵심과제



자료: Baltrušaitis, T. et al.(2019), 재인용

- (기술트렌드) 멀티모달 인공지능 기술 트렌드를 모달리티 통합* 방식을 기준으로 1세대(모달리티 연결형), 2세대(LLM 기반 확장형), 3세대(범용 통합형)으로 구분하고, 각 세대에서 보이는 특징을 분석함
- * 멀티모달 인공지능 핵심과제 중 하나인 '융합(fusion)'의 구체적 실현 방식과도 연결
- 세대별 모델의 기술 구조와 출력 생성물의 형태를 비교함으로써 최근 멀티모달 인공지능 기술의 발전 흐름을 살펴보고자 함

〈표 2-9〉 멀티모달 인공지능의 세대별 기술 트렌드

구분	주요 내용
1세대 (모달리티 연결형)	<ul style="list-style-type: none"> - 모달리티별 별도의 인코더로 처리한 후, 공통 임베딩 공간에서 표현을 정렬(Alignment)하는 방식으로 학습되며 멀티모달 이해 및 유사도 기반 매칭 작업에 중점을 둔 구조 - 각 모달리티 전용 모델들을 구성하고 출력 단계에서 연결 - 이해 기반 분류·매칭 작업 수행을 목적으로 생성물 없음 - 대표모델: 오픈AI CLIP, 구글 ALIGN
2세대 (LLM 기반 확장형)	<ul style="list-style-type: none"> - 트랜스포머 기반 대규모 언어모델(LLM)의 등장 이후 언어모델 중심으로 확장 모듈을 연결하는 방식으로 진화 - 이미지 등의 모달리티 입력이 임베딩 형태로 변환된 후, 언어모델에 입력되어 응답 생성에 활용 - 주로 텍스트 형태로 출력 생성 - 대표모델: 오픈AI GPT-4V, Anthropic Claude 3
3세대 (범용 통합형)	<ul style="list-style-type: none"> - 다양한 모달리티의 입력을 하나의 모델에서 통합하여 동등한 우선 순위로 처리하는 구조로 발전 - 여러 모달리티의 입력을 하나의 통합 시퀀스로 변환한 후, 트랜스포머 내부에서 모달 간 정보가 상호작용하며 결합 - 단순히 멀티모달 입력을 처리하는 수준을 넘어 외부 생성 모델과의 연계를 통해 이미지 생성 등 다양한 모달리티 출력 생성 지원 - 대표모델: 오픈AI GPT-4o, 구글 Gemini 2.5 Pro

자료: 저자 작성

- 1세대(모달리티 연결형)는 모달리티별 전용 인코더를 구성하여 정보를 처리한 후, 최종 임베딩 결과를 비교적 단순한 방식으로 결합(late-fusion)을 진행
- 2세대(LLM 기반 확장형)은 2017년 트랜스포머 기반 대규모 언어모델(LLM)의 등장 이후, 사전 학습된 언어모델을 기반으로 이미지, 음성 등의 입력을 추가로 연결하여 확장하는 형태로 발전
- 최근에 대두되는 3세대(범용 통합형)는 다양한 모달리티 데이터를 공동으로 학습하여 네이티브 멀티모달 능력을 갖추며, 여러 모달리티의 정보를 하나의 모델 내부에서 통합적으로 처리하는 구조로 발전하였다는 점에서 이전 세대 모델들과 뚜렷한 차이를 보임

2. 분석 개요

- (분석 목적) 생성형 인공지능의 고도화와 함께 주목받고 있는 멀티모달 인공지능에 대해 온라인에서 언급되는 현황을 분석하고자 함
- (분석 대상) 2021년 1월부터 2025년 5월까지 썬트렌드가 제공하는 온라인 매체(뉴스, X(트위터), 인스타그램, 블로그, 커뮤니티)에서 멀티모달 인공지능이 언급된 내용을 분석 대상으로 함
 - 언급량 및 연관어 분석 시 핵심 용어인 ‘멀티모달’을 키워드로 선정하였음
- (분석 내용) 멀티모달이 온라인 매체에서 얼마나 언급되고 있는지 그 추이를 살펴본 후, 연도별 관련도가 높은 키워드를 추출하여 관련 주요 이슈를 분석하고자 함

3. 멀티모달의 언급 추이

- (언급량) 멀티모달의 온라인 언급량은 2021년 1,037건 수준에서 2025년 1~5월 14,334건이 언급되는 등 꾸준히 언급량이 증가하는 추세를 보이고 있음
 - ※ 2025년 1월~5월 기사 수는 14,334건으로 전년 동월(2024년 1월~5월) 대비 언급량(6,005건)이 대폭 증가(138.7%)하였으며, 멀티모달에 대한 관심은 2025년 하반기에도 이어져 전년 대비 언급량이 크게 증가할 것으로 예상

〈표 2-10〉 2021년~2025년 5월 멀티모달 관련 온라인 언급량 추이

(단위: 건)

구분	2021년	2022년	2023년	2024년	2025년 1월~5월
합계	1,037 (-)	2,193 (111.5%)	8,614 (292.8%)	15,351 (78.2%)	14,334 (138.7%)

주: 1) 2021년 이전에는 연간 ‘멀티모달’ 언급량 합계가 1,000건 미만으로 2021년부터 분석 시작
 2) 괄호 안은 전년대비 증감률을 의미. 단, 2025년 5월까지의 자료는 전년 동월(2024년 1월~5월 합계) 대비 증감률을 의미

자료: 썬트렌드 분석자료

- (추이 변화 원인) 멀티모달 관련 언급량이 증가한 시점들은 주요 기업들의 멀티모달 인공지능 모델 출시 발표에 기인한 것으로 보임
 - 2021년 12월, 카카오 브레인은 텍스트를 입력하면 이미지를 생성하는 'minDALL-E'을 오픈소스로 공개²⁶⁾
 - 2023년 3월, 오픈AI는 텍스트와 이미지를 입력받아 텍스트를 생성할 수 있는 'GPT-4' 공개²⁷⁾²⁸⁾
 - 2023년 12월, 구글에서 텍스트, 이미지, 음성 등 다양한 형태의 정보를 동시에 인식하고 이해할 수 있는 멀티모달 인공지능 모델 'Gemini' 공개²⁹⁾³⁰⁾
 - 2024년 5월, 오픈AI에서 실시간 대화형 멀티모달 인공지능 모델 'GPT-4o' 공개³¹⁾, 메타는 텍스트와 이미지를 단일 모델에서 동시에 처리하는 'Chameleon' 연구논문 공개
 - 2025년 4월, 메타는 네이티브 멀티모달 인공지능 모델 'Llama 4' 공개³²⁾, 네이버는 자사 모델의 경량모델 시리즈 '하이퍼클로바X SEED'을 오픈소스로 공개³³⁾

26) 카카오(2021.12.15.), "카카오브레인, 초거대 AI 두 번째 모델 'minDALL-E' 공개", <https://www.kakaocorp.com/page/detail/9638> (접속일: 2025.7.1.)

27) OpenAI(2023.3.14.), <https://openai.com/index/gpt-4-research/>. (접속일: 2025.7.1.)

28) GPT-4 출시 당시 텍스트와 이미지 입력이 가능한 멀티모달 모델로 공개하였으나, 이미지 입력 기능은 일반 사용자에게 제공하지 않고 기술 데모로만 활용하였으며 이후 2023년 하반기부터 멀티모달 기능 사용 가능

29) Google(2023.12.6.), "Introducing Gemini: our largest and most capable AI model" <https://blog.google/technology/ai/google-gemini-ai/> (접속일: 2025.7.1.)

30) Gemini 1.0 출시 당시 음성 데이터 입력 처리 기능은 서비스상에서 비활성화되어 있었으며, 2024년에 출시된 Gemini 1.5 Pro부터 오디오 입력을 포함한 멀티모달 처리 본격 지원

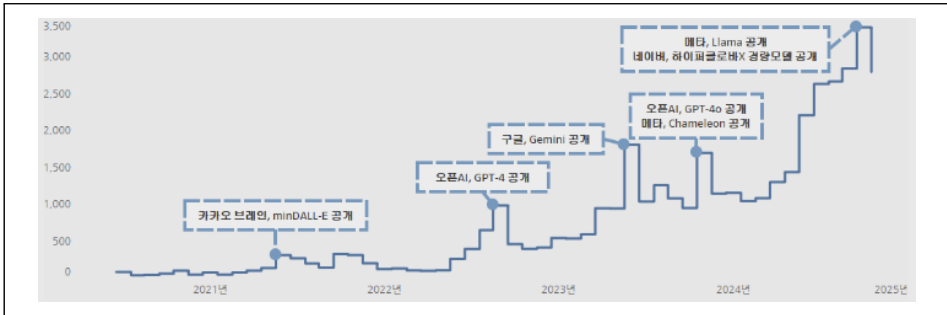
31) OpenAI(2024.5.13.), "Hello GPT-4o", <https://openai.com/index/hello-gpt-4o/> (접속일: 2025.7.1.)

32) Meta(2025.4.5.), "The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation", <https://ai.meta.com/blog/llama-4-multimodal-intelligence/> (접속일: 2025.7.1.)

33) 네이버클라우드(2025.4.23.), "AI 생태계에 씨앗을 뿌리다: 상업용 오픈소스 AI, HyperCLOVA X SEED", <https://tinyurl.com/4zbfscdp5> (접속일: 2025.7.1.)

[그림 2-9] 멀티모달 온라인 언급량 추이

(단위: 건)

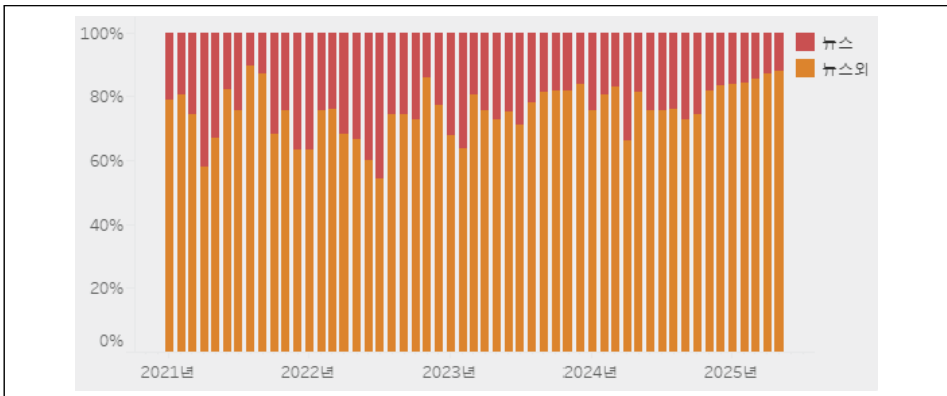


주: 1) 2021년 1월부터 2025년 5월까지의 자료를 활용하여 분석
 2) 분석 매체는 뉴스, X(트위터), 인스타그램, 블로그, 커뮤니티를 포함
 자료: 썬트렌드 분석자료

- (매체별 언급량) 썬트렌드 분석 결과, '멀티모달' 키워드는 뉴스 이외의 매체(X, 인스타그램, 블로그, 커뮤니티)에서의 언급이 상대적으로 높은 비중을 차지함
 - 특히, 오픈AI의 GPT-4o 모델 출시 시점인 2024년 5월 이후 뉴스 외 매체에서 언급 비중이 증가하는 추세로 나타나, 멀티모달 인공지능의 실용화 이후 일반 대중의 관심이 높아짐

[그림 2-10] 멀티모달 매체별 언급량 및 비중 추이

(단위: %)



주: 1) 2021년 1월부터 2025년 5월까지의 자료를 활용하여 분석
 2) 뉴스 외 매체는 X(트위터), 인스타그램, 블로그, 커뮤니티를 대상으로 분석
 자료: 썬트렌드 분석자료

4. 멀티모달 연관어 분석 결과

〈표 2-11〉 연도별 멀티모달 관련 주요 키워드

구분	2021년	2022년	2023년	2024년	2025년 1월~5월
주요 키워드	인공지능(AI) (440건)	인공지능(AI) (847건)	인공지능(AI) (5,822건)	인공지능(AI) (11,829건)	인공지능(AI) (10,692건)
	연구 (257건)	이미지 (479건)	이미지 (3,603건)	모델 (7,166건)	모델 (6,397건)
	지능 (252건)	모델 (466건)	모델 (3,330건)	이미지 (5,588건)	이미지 (5,817건)
	이미지 (226건)	연구 (443건)	텍스트 (2,470건)	지능 (4,541건)	기능 (4,563건)
	모델 (211건)	지능 (440건)	지능 (2,443건)	텍스트 (3,978건)	텍스트 (4,320건)
	학습 (204건)	LG (360건)	언어 (2,162건)	기능 (3,521건)	음성 (3,450건)
	언어 (179건)	텍스트 (353건)	서비스 (2,099건)	음성 (3,319건)	지능 (2,949건)
	영상 (175건)	초거대 (350건)	음성 (2,043건)	언어 (3,062건)	사용자 (2,640건)
	계획 (161건)	언어 (325건)	구글 (1,666건)	서비스 (2,695건)	영상 (2,304건)
	초거대 (152건)	서비스 (319건)	기능 (1,635건)	기업 (2,568건)	성장 (2,178건)

주: 1) 2021년 1월부터 2025년 5월까지 자료를 활용하여 분석

2) 2025년은 1~5월간 언급된 자료를 분석한 결과로, 이전 자료의 건수와 직접 비교 시 유의

3) 분석매체는 뉴스, X(트위터), 인스타그램, 블로그, 커뮤니티를 포함

자료: 썬트렌드 분석자료

- (발전 형태 변화) 분석 기간 초기에는 ‘연구’, ‘학습’, ‘계획’이 상위 키워드로 나타났으나, 이후에는 ‘서비스’, ‘기능’, ‘사용자’ 등이 상위 키워드로 언급되며 멀티모달 인공지능이 연구 중심 단계에서 사용자와 밀접한 서비스 형태로 발전하고 있음
- 주요 멀티모달 인공지능 모델이 출시되었던 2023년부터 ‘서비스’, ‘기능’ 키워드의 언급량이 크게 증가한 것으로 보아 실사용 중심의 기술로 전환된 모델들의 기능에 대한 관심이 늘어난 것을 알 수 있음

- (모달리티 영역 확장) 분석 기간 초기부터 언급되었던 ‘이미지’, ‘텍스트’는 꾸준히 상위 키워드에 포함되는 한편, 분석 기간 중·후기에 ‘음성’, ‘영상’ 등의 키워드가 언급되고 있어 멀티모달 인공지능의 발전에 따라 모달리티 영역이 점차 확장되고 있음을 보여줌
 - 분석 기간 전반에 걸쳐 ‘이미지’, ‘텍스트’ 키워드가 지속적으로 언급되고 있는 것으로 보아, 두 모달리티가 멀티모달 인공지능에서 가장 기본적이고 핵심적인 영역임을 알 수 있음
 - ※ (이미지) 2021년 226건 → 2022년 479건 → 2023년 3,603건 → 2024년 5,588건
→ 2025년 1월~5월 5,817건
 - (텍스트) 2021년 (-)건 → 2022년 353건 → 2023년 2,470건 → 2024년 3,978건
→ 2025년 1월~5월 4,320건
 - 또한, 분석 기간 후기로 갈수록 ‘음성’ 키워드 언급량이 증가하는 것은 멀티모달 인공지능에서의 활용도가 확대되고 있으며, 대화형 인터페이스 및 실시간 상호작용이 강화되고 있음을 보여줌
 - ※ (음성) 2021년 (-)건 → 2022년 (-)건 → 2023년 2,043건 → 2024년 3,319건 →
2025년 1월~5월 3,450건
 - 분석 기간 후기 ‘영상’ 키워드의 등장은 최근 멀티모달 인공지능 모델의 처리 능력이 고도화와 함께 영상 등 고차원 모달리티에 대한 주목이 확대되고 있음을 시사

5. 결론

- 본고에서는 멀티모달 인공지능의 기술 트렌드를 분석하고, 썬트렌드 자료를 활용하여 온라인에서 언급되는 관련 주요 이슈를 살펴보았음
- 멀티모달 인공지능은 텍스트, 이미지, 오디오, 비디오 등 여러 모달리티의 정보를 처리하고 통합할 수 있는 머신러닝 모델을 의미하며, 인간과 유사한 인지 능력을 갖춘 인공지능 시스템을 구현하여 자연스러운 상호작용과 고차원적인 추론 및 판단을 수행하는 것을 목표로 함

- 멀티모달 인공지능 기술 트렌드를 모달리티 통합 방식을 기준으로 1세대(모달리티 연결형), 2세대(LLM 기반 확장형), 3세대(범용 통합형)으로 구분하고, 각 세대에서 보이는 특징을 분석함
 - 1세대(모달리티 연결형)는 모달리티별 전용 인코더를 구성하여 정보를 처리한 후, 최종 임베딩 결과를 비교적 단순한 방식으로 결합(Late-fusion)을 진행
 - 2세대(LLM 기반 확장형)는 대규모 언어모델을 기반으로 이미지, 음성 등의 입력을 추가로 연결하여 확장하는 형태로 발전하였으며 주로 텍스트 형태의 출력을 생성
 - 3세대(범용 통합형)는 여러 모달리티의 정보를 하나의 모델 내부에서 통합적으로 처리하는 구조로 발전하였으며 외부 생성 모델과의 연계를 통해 다양한 모달리티의 출력 생성을 지원
- 언급량 분석 결과, 2021년 1,037건 수준에서 2024년 15,351건이 언급되며 지속적으로 상승하는 추세를 보임
 - (온라인 언급량) 온라인 매체에서 멀티모달 키워드의 언급량이 증가한 시점들은 주요 기업들의 멀티모달 인공지능 모델 출시 발표에 기인한 것으로 보임
 - (매체별 언급량) '멀티모달' 키워드는 뉴스보다 뉴스 이외의 매체에서의 언급이 상대적으로 높은 비중을 차지하며, 특히 2024년 5월 이후 뉴스 외 매체의 언급 비중이 더욱 증가하는 것으로 나타남
- 연관어 분석 결과, 연도별 멀티모달 관련 주요 키워드를 바탕으로 멀티모달 인공지능의 발전 형태 변화와 모달리티 영역 확장에 대해 파악이 가능함
 - 분석 기간 초기에는 '연구', '학습', '계획'이 상위 키워드로 나타났으나, 이후에는 '서비스', '기능', '사용자' 등이 상위 키워드로 언급되며 멀티모달 인공지능이 연구 중심 단계에서 사용자와 밀접한 서비스 형태로 발전하고 있음
 - '이미지', '텍스트'는 분석 기간 전반에 걸쳐 꾸준히 상위 키워드에 포함되는 한편, '음성', '영상' 등의 키워드는 분석기간 중·후기에 언급되고 있어 멀티모달 인공지능의 발전에 따라 모달리티 영역이 확장되고 있음을 보여줌
- 향후 멀티모달 인공지능이 추가적인 모달리티를 포함하여 처리할 수 있도록 확장될 뿐만 아니라, 영상 데이터 학습 및 분석을 위한 발전된 기술도 포함할 것으로 예상³⁴⁾

- 또한, 특정 산업(예: 법률, 금융, 의료 등)에 대한 데이터를 기반으로 학습된 특수화된 멀티모달 모델들이 이러한 분야에서 생산성을 높이고 새로운 가치를 추가할 것으로 전망
- 멀티모달 인공지능의 급속한 발전을 이루면서 온라인상에서의 언급량과 사회적 관심도 함께 증가하고 있음. 이에 따라, 향후 멀티모달 데이터 구축 및 응용 서비스 고도화를 위한 노력이 필요할 것으로 보임
- 2024년 2월 과학기술정보통신부가 개최한 ‘제2차 인공지능 데이터 융합 네트워크’에서 생성형 인공지능 트렌드가 언어 중심에서 멀티모달로 발전할 것으로 전망함에 따라, 다양한 유형의 데이터가 결합된 멀티모달 데이터를 본격적으로 구축하여 국내 생성형 인공지능 모델 및 서비스 고도화 지원 계획 발표³⁵⁾

34) Gartner(2025.4.15.), “Emerging Tech Horizon for Generative AI”

35) 과학기술정보통신부 보도자료(2024.2.16.), “제2차 인공지능 데이터 융합 네트워크 개략”

제 4 절 양자컴퓨팅 키워드의 언론 인식 변화³⁶⁾

1. 양자컴퓨팅의 정의

- (개요) 본 절에서는 차세대 핵심 기술로 주목받고 있는 양자컴퓨팅(Quantum Computing)이 언론에서 언급되고 있는 담론의 변화에 대해 살펴보고자 함
- (정의) 양자컴퓨팅은 양자역학의 고유한 현상인 중첩(Superposition)과 얽힘(Entanglement)을 계산에 활용하여, 기존의 고전 컴퓨터로는 해결하기 어려운 복잡한 문제들을 처리하도록 설계된 컴퓨팅 기술 및 연구 분야를 의미³⁷⁾
 - 고전 컴퓨터는 정보가 0과 1로 표현되는 문자열인 비트(bit)로 계산되는 반면, 양자 컴퓨터는 중첩의 원리를 활용해 0과 1이 동시에 존재할 수 있는 큐비트(qbit)로 정보를 처리하여 다양한 상태의 조합을 동시에 표현³⁸⁾
- (상용화 전망) Gartner(2025.07.)³⁹⁾는 양자컴퓨팅을 향후 산업·사회 전반에 변혁적(Transformational) 영향을 미칠 기술로 평가하며, 상용화까지는 10년 이상이 소요될 것으로 전망
 - 양자 시뮬레이션, 보안, AI, 최적화 문제 해결 등에서 고전 컴퓨터로는 불가능했던 가치를 창출할 것으로 예상
 - 글로벌 기업 및 각국 정부의 대규모 투자, QCaaS(Quantum Computing as a Service)의 확산, 오류정정 기술 고도화 등을 통해 생태계가 빠르게 성장하고 있으며, 신약 개발·신소재·암호해독·AI 최적화 등 일부 분야에서 가시적 발전을 전망

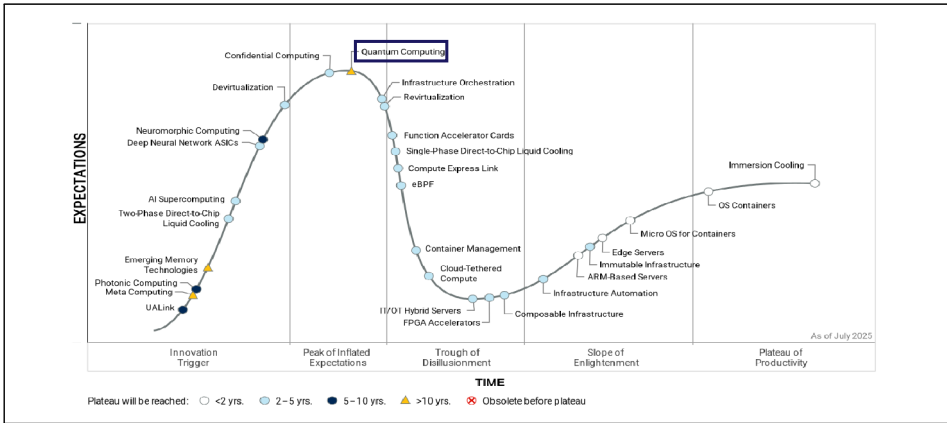
36) 분석 내용은 진정민, 손영선(2025), “양자컴퓨팅 키워드의 언론 인식 변화”로 ICT통계 분석 25-03호에 발표되었음

37) IBM 홈페이지, “What is quantum computing?”,
https://www.ibm.com/think/topics/quantum-computing?utm_source=chatgpt.com
(접속일자: 2025.09.10.)

38) Emanuel Knill(2010). “Quantum Computing”, Nature, Vol.463, pp.441-443,
<https://www.nature.com/articles/463441a>

39) Gartner(2025.07.), “Hype Cycle for Compute, 2025”,
<https://www.gartner.com/interactive/hc/6671434?ref=solrAll&refval=493038320>
(접속일자: 2025.09.16.)

[그림 2-11] 위성통신 언론 언급량 추이



주: Gartner의 기술 성숙도 모델인 Hype Cycle 상 양자컴퓨팅은 기술 촉발(Innovation Trigger)을 넘어서 기대의 정점(Peak of Inflated Expectations)에 해당하는 것으로 평가되었으며, 상용화까지는 10년 이상이 소요될 것으로 평가

출처: Gartner(2025.07.), "Hype Cycle for Compute, 2025"

2. 분석 개요

- (분석 목적) 양자컴퓨팅 관련 이슈가 언론에서 어떻게 언급되고 있는지를 분석하고, 관련 담론이 어떻게 변화했는지 분석하여 주요 이슈와 시사점을 도출하고자 함
- (2018년 1월*부터 2025년 8월까지 뉴스 및 SNS 빅데이터 분석 서비스를 제공하는 썬트렌드에서 양자컴퓨팅이 언급된 내용을 분석 대상으로 함

* 2018년은 미국이 양자정보과학(QIS) 연구개발 및 국가전략 수립을 위해 연간 약 10억 달러 규모로 5년간 투자를 진행하는 National Quantum Initiative Act⁴⁰⁾를 제정하고, EU가 양자컴퓨팅·통신·센서·시뮬레이션 등 4대 분야 경쟁력 확보를 위해 10년간 10억 유로 규모의 Quantum Technologies Flagship⁴¹⁾을 출범시킨 시기로 글로벌 양자기술 경쟁이 본격화된 시점임에 따라 해당 시점을 분석 시기로 설정

40) U.S. Congress, National Quantum Initiative Act, Public Law No: 115-368 (2018).

41) European Commission, Quantum Technologies Flagship (2018-2028), https://digital-strategy.ec.europa.eu/en/policies/quantum-technologies-flagship?utm_source=chatgpt.com(접속일자 : 2025.09.10.)

- 양자컴퓨팅의 영어식 표현인 퀀텀컴퓨팅 또한 혼용됨에 따라 동의어로 통합하여 분석을 진행
- (분석 내용) 양자컴퓨팅이 썬트렌드의 뉴스 매체에서 얼마나 언급되고 있는지 그 추이를 살펴본 후, 연도별 관련도가 높은 텍스트를 추출·분석하여 관련 주요 이슈를 분석하고자 함

3. 양자컴퓨팅의 언급 추이

- (언급량) 양자컴퓨팅의 뉴스 언급량은 2018년 346건 수준에서 2024년 2,536건 수준으로 언급되는 등 뉴스 언급량이 급격히 증가하는 추세를 보이고 있음
- ※ 2025년 1월~8월 뉴스 언급량은 3,608건으로 전년 동월(2024년 1월~8월) 대비 언급량(1,299건)이 177.8% 증가하는 등 2025년에도 양자컴퓨팅의 뉴스 언급량이 증가하는 추세를 보이고 있음

〈표 2-12〉 2018년~2025년 8월 양자컴퓨팅 관련 뉴스 언급량 추이

(단위: 건)

구분	2018년	2019년	2020년	2021년	2022년	2023년	2024년	2025년 1월~8월
합계	346	467 (35.0%)	528 (13.1%)	737 (39.6%)	1,074 (45.7%)	2,094 (95.0%)	2,536 (21.1%)	3,608 (177.8%)

주: 괄호 안은 전년대비 증감률을 의미. 단, 2024년 7월까지의 자료는 전년 동월(2023년 1월~7월 합계) 대비 증감률을 의미

자료: KISDI 분석자료

- (추이 변화 원인) 양자컴퓨팅 관련 뉴스 언급량이 증가한 시점들은 글로벌 기업의 기술발표, 정부의 R&D 투자, 국제 협력 등의 이슈에 기인한 것으로 보임
- 2019년 10월, 구글은 54qbit 시카모어(Sycamore) 프로세서를 이용해 슈퍼컴퓨터로 1만 년이 걸리는 연산을 200초 만에 해결했다고 발표하며, 양자 우위(Quantum Supremacy)* 달성에 성공⁴²⁾했다는 언론 보도가 집중되며 언급량이 증가

42) 조선비즈(2019.10.31), “구글 “10년 뒤엔 양자컴퓨터…슈퍼컴퓨터로 1만년 할 일을 200초만에”, <https://n.news.naver.com/mnews/article/366/0000446509>

- * Preskill, J. (2012)이 제안한 용어로, 양자컴퓨터가 고전 컴퓨터로는 실용적으로 풀 수 없는 문제를 해결하는 순간을 의미⁴³⁾
- 2022년 1월에는 과학기술정보통신부의 50qbit 양자 컴퓨터 개발·착수⁴⁴⁾와 더불어 LG전자⁴⁵⁾·현대차⁴⁶⁾ 등의 양자컴퓨팅 협력 소식이 잇따라 나오며 정책·산업 이슈가 집중적으로 보도되어 언급량이 증가
 - ※ (LG전자) IBM이 기업과 연구소 등 170여 회원사와 결성한 양자컴퓨팅 발전 협력체인 'IBM 퀀텀 네트워크'에 LG전자가 합류하여, IBM이 제공하는 양자 컴퓨팅 시스템을 포함한 양자 소프트웨어 개발도구를 이용 가능, (현대차) 미국의 양자컴퓨터 업체인 아이온큐와 협력하여 효율성과 안정성이 높아진 차세대 배터리 개발
 - 2023년 8월, 미국의 첨단 반도체, 양자컴퓨팅, 인공지능 등 3개 분야에 대한 중국 투자 금지 행정명령 발표⁴⁷⁾와 더불어, 한·미·일 정상외 캠프 데이비드 회담에서 공급망·기술안보·AI·양자컴퓨팅 등 신기술 분야 협력을 공고화하기로 합의⁴⁸⁾한 기사로 인해 언론 관심이 급증
 - 2025년 1월, 마이크로소프트가 2025년을 '양자컴퓨팅 준비의 해'라 선언⁴⁹⁾하고, LGU+는 양자컴퓨팅 위협에 대응할 수 있는 양자내성암호(Post-Quantum

43) Preskill, J. (2012). Quantum computing and the entanglement frontier (2012). arXiv preprint arXiv:1203.5813, 165.

44) 뉴스1(2022.01.12.), “과학기술정보통신부, ICT원천연구개발에 904억 투자···양자컴·차세대 반도체 개발”, <https://n.news.naver.com/mnews/article/421/0005840490>

45) 동아일보(2022.01.11.), “LG전자 “美 IBM과 양자컴퓨팅 기술 개발 협력””, <https://n.news.naver.com/mnews/article/020/0003404977>

46) 동아일보(2022.01.21.), “현대차, 양자컴퓨터로 차세대 배터리 개발 나선다”, <https://n.news.naver.com/mnews/article/020/0003407064>

47) 연합뉴스(2023.08.10.), “美 “투자가 위협 키워”···AI·반도체·양자컴퓨팅 對中투자 제한(종합)”, <https://n.news.naver.com/mnews/article/001/0014122863>

48) 뉴스1(2023.08.19.), “한미일 파트너십 새시대···안보·공급망, 인태·우주까지 협력”, <https://n.news.naver.com/mnews/article/421/0006997988>

49) 머니투데이(2025.01.15.), “마이크로소프트 “올해는 양자컴퓨팅 준비하는 해” 선언”, <https://n.news.naver.com/mnews/article/008/0005141675>

Cryptography, PQC)* 기반 솔루션인 ‘알파키’를 출시⁵⁰⁾했으며, 부산시는 마이크로소프트·아이온큐와 지역 IT산업 투자 업무협약을 체결⁵¹⁾하여 산업계와 지자체 차원의 움직임이 집중 보도되며 가장 많은 언급량을 기록

* Peter Shor(1994)가 제안한 알고리즘으로, 양자컴퓨터 환경에서도 안전하게 사용할 수 있는 새로운 공개키 암호체계를 의미⁵²⁾

- 2025년 3월에는 삼성전자의 2025년형 비스포크 AI 가전 내 양자내성암호(PQC) 도입⁵³⁾, 엔비디아 젠슨 황 CEO의 “양자컴퓨터 상용화 20년” 발언으로 인한 시장 충격에 대한 사과 및 엔비디아 가속 양자 연구센터(NVAQC) 설립 발표⁵⁴⁾ 등으로 언론 관심이 증가

[그림 2-12] 양자컴퓨팅 뉴스 언급량 추이



자료: 썬트렌드 분석자료

50) 머니투데이(2025.01.30.), “LG유플러스, PQC 기술 활용한 통합 계정관리 솔루션 ‘알파키’ 출시”, <https://n.news.naver.com/mnews/article/008/0005146778>

51) 파이낸셜뉴스(2025.01.12.), “美 MS·아이온큐와 투자협약… 부산 ‘제2 시애틀’로 만든다”, <https://n.news.naver.com/mnews/article/014/0005294383>

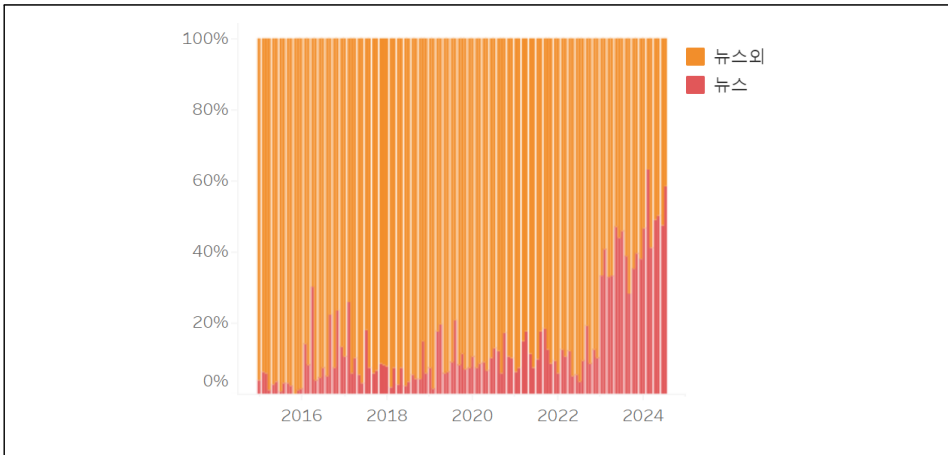
52) KISA 암호이용활성화 홈페이지(<https://seed.kisa.or.kr/kisa/ngc/pqc.do>)

53) SBS Biz(2025.03.30.), “AI·보안·편의성 더했다…삼성전자, 비스포크AI 가전 신제품 공개”, <https://n.news.naver.com/mnews/article/374/0000432256>

54) 세계일보(2025.03.21.), “‘개발 20년 걸릴 것’ 젠슨 황, 양자컴 기업에 사과…“상장사 있는 줄 몰랐다””, <https://n.news.naver.com/mnews/article/022/0004020981>

- (매체별 언급량) 양자컴퓨팅 키워드는 뉴스 이외의 매체(커뮤니티, 인스타그램, 블로그, X)에서의 언급이 상대적으로 높은 비중을 차지
 - 2018~2021년까지는 주로 뉴스 이외의 매체에서 관련 담론이 형성되었으나, 2022년 이후 정부 정책 발표 및 기업 협력 소식이 이어지며 뉴스 비중이 점차 확대
 - 2024~2025년에는 시장·투자자의 관심이 폭발적으로 증가하여 뉴스 외 매체의 비중이 압도적으로 높아진 것으로 나타나, 양자컴퓨팅이 연구 중심 담론에서 산업·투자 담론으로 이동
 - ※ 양자컴퓨팅 ETF 급등, 아마존의 아이온큐 투자 발표 이후 관련주 강세, 엔비디아 젠슨 황 발언에 따른 양자컴퓨팅 관련주 급등 등 투자 이슈가 뉴스 외 매체에서 다수 언급됨
 - 다만, 본고에서는 정확성과 시의성이 높은 뉴스 매체 중심으로 연관어 분석을 수행하여, 주요 정책·기술·산업 이슈를 도출

[그림 2-13] 양자컴퓨팅 매체별 언급량 및 비중 추이



자료: 썬트렌드 분석자료

4. 양자컴퓨팅 연관어 분석 결과

〈표 2-13〉 연도별 양자컴퓨팅 관련 주요·사업자 관련 키워드

구분	2018년	2019년	2020년	2021년	2022년	2023년	2024년	2025년 1월~8월 ³⁾
주요 키워드 ¹⁾	연구 (144건)	연구 (220건)	연구 (221건)	연구 (342건)	연구 (470건)	기업 (869건)	연구 (1,026건)	기업 (1,788건)
	인공지능 (122건)	기술개발 (137건)	교수 (135건)	인공지능 (240건)	인공지능 (359건)	인공지능 (820건)	기업 (993건)	산업 (1,405건)
	기업 (99건)	산업 (122건)	인공지능 (121건)	산업 (195건)	전략 (312건)	연구 (765건)	산업 (952건)	인공지능 (1,360건)
	산업 (91건)	인공지능 (99건)	과제 (97건)	전략 (195건)	산업 (308건)	미국 (750건)	인공지능 (905건)	연구 (1,342건)
	미국 (72건)	미국 (88건)	반도체 (96건)	반도체 (180건)	반도체 (293건)	반도체 (731건)	미국 (819건)	미국 (1,145건)
	반도체 (69건)	반도체 (84건)	미국 (87건)	미국 (169건)	미국 (244건)	투자 (572건)	반도체 (644건)	투자 (826건)
	연구개발 (69건)	투자 (81건)	학습 (83건)	투자 (157건)	투자 (222건)	산업 (568건)	투자 (536건)	전략 (664건)
	전략 (58건)	기업 (79건)	기술개발 (81건)	기술개발 (140건)	기술개발 (219건)	중국 (511건)	전략 (508건)	반도체 (644건)
	기초 (57건)	교수 (68건)	KAIST (71건)	기업 (136건)	기업 (219건)	안보 (438건)	중국 (472건)	시장 (643건)
	투자 (56건)	프로젝트 (54건)	산업 (67건)	중국 (127건)	시스템 (211건)	전략 (390건)	시스템 (439건)	주가 (552건)
	기술개발 (53건)	전략 (53건)	기업 (65건)	연구개발 (119건)	연구개발 (174건)	협력 (308건)	협력 (434건)	기술개발 (467건)
	미래 (47건)	연구자 (52건)	발전 (58건)	일본 (99건)	중국 (135건)	기술개발 (296건)	안보 (405건)	증시 (418건)
	계획 (46건)	기초 (51건)	기초 (55건)	협력 (94건)	안보 (134건)	경쟁 (258건)	연구개발 (379건)	혁신 (393건)
	성장 (38건)	연구개발 (48건)	논문 (53건)	경쟁 (91건)	우주 (129건)	시장 (244건)	기술개발 (335건)	가능성 (349건)
	중국 (34건)	개념 (45건)	전략 (52건)	정책 (78건)	로봇 (117건)	펀드 (227건)	시장 (289건)	종목 (344건)

구분	2018년	2019년	2020년	2021년	2022년	2023년	2024년	2025년 1월-8월 ³⁾
사업자 관련 키워드 ²⁾	IBM (51건)	IBM (100건)	삼성 (99건)	IBM (81건)	IBM (132건)	IBM (200건)	IBM (265건)	엔비디아 (775건)
	마이크로 소프트 (46건)	구글 (82건)	삼성전자 (77건)	LG (56건)	LG (78건)	SK (136건)	구글 (233건)	아이온 (아이온큐) ⁴⁾ (703건)
	구글 (44건)	아마존 (49건)	IBM (52건)	구글 (39건)	삼성 (77건)	LG (113건)	엔비디아 (166건)	마이크로 소프트 (557건)
	인텔 (28건)	마이크로 소프트 (44건)	마이크로 소프트 (42건)	LGU+ (30건)	마이크로 소프트 (62건)	구글 (109건)	아이온 (아이온큐) (144건)	구글 (503건)
	화웨이 (22건)	삼성전자 (37건)	구글 (22건)	메타 (29건)	메타 (68건)	SKT (87건)	아마존 (138건)	IBM (466건)

- 주: 1) 연도별 양자컴퓨팅 관련 뉴스 기사 중, 언급 건수 상위 50개 키워드 중 주요 키워드를 선정하여 표기
- 2) 연도별 양자컴퓨팅 관련 뉴스 기사 중, 언급 건수 상위 5개 사업자를 표기. '마이크로 소프트웨어'는 '마이크로소프트'와 'MS' 언급 건수 합산 기준
- 3) 2025년 1월-8월 기사 기준으로 분석한 자료로, 이전 연도와 건수를 직접 비교 시 유의
- 4) '아이온'은 미국의 양자컴퓨팅 및 소프트웨어 기업인 '아이온큐(IonQ)'를 지칭하는 것으로 판단하였음. '아이온'이 상위 키워드로 언급된 2024년~2025년 양자컴퓨팅 연관에 분석에서 '아이온'을 포함어로 분석한 결과, 아이온큐의 ETF 상장, 추가 급등, 기술 발표, 국내 대학 협력 등의 뉴스가 언급되는 것으로 확인되어 '아이온큐'로 해석, 사업자 관련 키워드에 포함하였음

- (담론의 변화) 양자컴퓨팅과 관련한 담론은 연구 → 산업·정책 → 투자·시장 중심으로 변화해가는 양상을 보임
 - 분석 기간 초기에는 “연구”, “연구개발”, “기술개발”, “인공지능”, “기초” 등의 키워드가 주요 키워드로 나타났으며, “교수”, “학습”, “KAIST” 등 학술 관련 키워드도 함께 등장하여 양자컴퓨팅이 사회적으로 차세대 기술로 인식되었으나 산업적 적용이나 정책 지원의 논의는 본격화되지 않은 단계로 인식
 - 분석 기간 중기에는 “기업”, “산업”, “전략”, “협력” 등의 언급량이 증가하고, “반도체”, “시스템”, “경쟁”, “우주”, “로봇” 등이 주요 키워드로 언급되며 산업·정책 관련 담론이 형성

- 분석 기간 후기에는 “시장”, “투자”, “주가”, “증시”, “펀드” 등의 키워드가 상위권으로 진입하며, 산업화와 더불어 투자·시장 관련 담론이 얽힌 종합적 담론으로 발전
 - ※ 연구 관련 키워드는 전체 분석 기간에 걸쳐 지속적으로 주요 키워드로 등장하고 있어, 기술적 진보와 연구개발이 여전히 담론의 핵심 축으로 기능하고 있는 것으로 판단

[연구] (2018년) 144건 → (2020년) 221건 → (2022년) 470건 → (2024년) 1,026건
→ (2025년) 1,788건

- (사업자 관련 키워드) 양자컴퓨팅 사업자 관련 키워드는 시기별로 성격이 달라지며 점차 다층화되는 양상을 보임

- 분석 기간 초기에는 “IBM”, “구글”, “마이크로소프트” 등 글로벌 빅테크기업이 주로 언급되며, 해당 기업들의 양자우위 실험, 양자칩 개발, 클라우드 기반 시뮬레이터 등 기술 가능성과 연구 성과를 중심으로 이루어진 것으로 보임

※ (양자우위 실험) 2019년 구글의 54qbit 시카모어 프로세서를 통한 양자 우위 (Quantum Supremacy) 달성⁵⁵⁾, (양자칩 개발) 2021년 IBM의 127qbit 이글(Eagle) 프로세서 공개⁵⁶⁾, (클라우드 기반 시뮬레이터) 2021년 마이크로소프트의 클라우드 기반 양자컴퓨팅 개발 환경 애저 퀀텀(Azure Quantum) 프리뷰 버전 공개⁵⁷⁾

- 산업·정책 관련 담론이 형성되었던 분석 기간 중기에는 국내 기업인 “삼성”, “LG”, “SK”, “SKT” 등이 본격적으로 뉴스에 등장하며 산업 생태계가 확장되는 양상을 보이며, 주로 글로벌 빅테크기업과의 협력, 정부 로드맵과 같은 산업화·정책 추진과 관련된 내용이 포함

※ (글로벌 빅테크 협력) 2022년 1월 LG-IBM 양자컴퓨팅 개발 협력⁵⁸⁾, 2022년

55) 조선비즈(2019.10.31.), “구글 “10년 뒤엔 양자컴퓨터…슈퍼컴퓨터로 1만년 할 일을 200초만에””, <https://n.news.naver.com/mnews/article/366/0000446509>

56) MoneyS(2021.11.16.), “IBM, ‘양자컴퓨팅 시대’ 선도한다… 127큐비트 ‘이글’ 공개”, <https://n.news.naver.com/mnews/article/417/0000754665>

57) ZDNET Korea(2021.02.02.), “MS, 양자컴퓨팅 개발환경 누구에게나 제공”, <https://n.news.naver.com/mnews/article/092/0002212585>

58) 세계일보(2022.01.11.), “LG, IBM과 양자컴퓨팅 개발 협력”, <https://n.news.naver.com/mnews/article/022/0003656164>

8월 삼성SDS AWS·MS 등 글로벌 13개 기업과 함께 양자내성암호 전환 프로젝트 참여⁵⁹⁾, (정부 로드맵) 2022년 과기정통부 ‘양자정보주관’ 개최, 아이온큐·IBM 등 글로벌 기업과 LG전자·현대기아차 참여, 통신3사의 양자암호 상용화 발표 등⁶⁰⁾ - 후기에는 “엔비디아”, “아이온큐”와 같은 글로벌 기업이 투자 테마주·ETF 편입 종목으로 급부상하여 관련 기사량이 증가한 것으로 나타남

※ 2024년 말 국내 최초 양자컴퓨팅 ETF KOSEF 미국양자컴퓨팅이 상장 5분만에 완판되며, 7거래일만에 순자산 500억 원 돌파⁶¹⁾, 2025년 엔비디아 쉐슨 황 CEO의 양자컴퓨팅 상용화 임박 발언으로 양자컴퓨팅 테마 ETF 및 아이온큐 등 종목 급증⁶²⁾

5. 결론

- 본고에서는 차세대 핵심 기술로 주목받고 있는 양자컴퓨팅 관련 이슈가 언론에서 어떻게 언급되고 있는지를 분석하고, 관련 담론의 변화에 대해 살펴보고자 함
- 언론 언급량 분석 결과, 양자컴퓨팅의 뉴스 언급량은 2018년 346건에서 2024년 2,536건 언급되는 등 뉴스 언급량이 급격히 증가하는 추세를 보이고 있음
 - 양자컴퓨팅 관련 뉴스 언급량이 증가한 시점들은 글로벌 기업의 기술발표, 정부의 R&D 투자, 국제 협력 등의 이슈에 기인한 것으로 보임
 - SNS 빅데이터 분석 서비스를 제공하는 썬트렌드 분석 결과, 양자컴퓨팅 키워드는 뉴스 이외의 매체(커뮤니티, 인스타그램, 블로그, X)에서의 언급이 상대적으로 높은 비중을 차지

59) 서울경제(2022.08.25.). “삼성SDS, AWS·MS 우수 기업과 ‘양자암호’ 기술 표준 만든다”, <https://n.news.naver.com/mnews/article/011/0004091499?rc=N&ntype=RANKING>

60) 뉴스1(2022.06.27.), ““양자 시대 연다”...과기정통부 ‘2022 양자정보주관’ 진행”, <https://n.news.naver.com/mnews/article/421/0006180376?rc=N&ntype=RANKING>

61) 머니투데이(2024.12.27.), “KOSEF 미국양자컴퓨팅 ETF, 7거래일만에 순자산 500억 돌파”, <https://n.news.naver.com/mnews/article/008/0005133795>

62) 머니투데이(2025.06.12.), “셸슨 황 한 마디에 또다시 들쭉...양자컴퓨팅 ETF 주목도 ↑”, <https://n.news.naver.com/mnews/article/008/0005206753>

- 연관이 분석 결과, 양자컴퓨팅과 관련한 담론은 연구 → 산업·정책 → 투자·시장 중심으로 변화해가고, 사업자 키워드도 시기별로 다층화되는 양상을 보임
 - 분석 기간 초기에는 “연구”, “기술개발”, “교수”, “학습” 등 학술 키워드가 주를 이루어 차세대 기술로 인식되었으나 산업·정책 논의는 제한적이었으며, 사업자 키워드는 “IBM”, “구글”, “마이크로소프트” 등 글로벌 빅테크 중심으로 기술 성과와 가능성에 초점
 - 중기에는 “기업”, “산업”, “전략”, “경쟁” 등의 키워드가 등장하며 산업·정책 관련 담론이 형성되고, “삼성”, “LG”, “SK” 등 국내 기업이 등장하며 생태계 확장이 시작
 - 후기에는 “시장”, “투자”, “주가”, “증시” 등이 상위권에 진입하며 산업화와 더불어 투자·시장 담론으로 발전하고, “엔비디아”, “아이온” 등 글로벌 기업이 투자 테마로 급부상
- 양자컴퓨팅은 연구 중심의 기술에서 산업·정책 담론과 더불어 투자 분야로 확산되며, 국가 경쟁력 확보의 핵심 영역으로 부상하고 있음
 - 정부는 2023년 「대한민국 양자과학기술 전략⁶³⁾」을 통해 양자컴퓨터·양자통신·양자센서 3대 분야에 대해 7대 추진 방향을 제시하고, 「양자과학기술 및 양자산업 육성에 관한 법률⁶⁴⁾」 제정을 통해 5년 단위 종합계획 수립·양자전략위원회 설치·양자팍 구축 지원 등을 법제화
 - 또한, 2024년 4월에는 양자과학기술의 도약을 위한 4대 추진전략을 담은 「퀀텀 이니셔티브」를 확정⁶⁵⁾하고, 2025년 3월 「퀀텀 이니셔티브」의 본격 추진을 통해 신속한 성과 창출·산업화를 위해 10대 추진과제 도출⁶⁶⁾
 - 향후 양자컴퓨팅은 AI·보안·신약개발·신소재 탐색 등과 융합해 산업·사회 전반에 새로운 가치를 창출할 핵심 기술임에 따라, 지속적인 정책 지원과 산업 생태계 조성을 병행하여 글로벌 경쟁에서 선도적 위치를 확보할 필요가 있음

63) 과학기술정보통신부(2023.06.), “대한민국 양자과학기술 전략”.

64) 법제처 국가법령정보센터, 「양자과학기술 및 양자산업 육성에 관한 법률」, 법률 제19990호, 제정 2023.10.31, 시행 2024.11.1.

65) 과학기술정보통신부 보도자료(2024.04.25.), “양자과학기술 대도약, 디지털을 넘어 “퀀텀의 시대”로”.

66) 과학기술정보통신부 보도자료(2025.03.12.), “양자기술, 연구개발을 넘어 산업화로”.

제 5 절 개방형 AI 모델 동향 분석⁶⁷⁾

1. 개방형 AI 모델의 정의

- (개요) 본 절에서는 개방형 AI 모델의 발전 과정과 동향을 정리하고, 관련 텍스트 분석을 통해 개방형 AI와 관련된 담론의 변화를 살펴보고자 함
 - (정의) 개방형 AI 모델은 오픈소스(Open Source) 또는 오픈웨이트(Open Weight) 형태로 공개된 인공지능 모델을 의미하며, 폐쇄형 모델과 달리 구조·가중치·데이터 등을 활용해 자유롭게 커스터마이징과 파인튜닝이 가능
 - 누구나 활용·수정이 가능하며, 폐쇄형 모델 대비 투명성과 접근성이 높음
 - 데이터 보안, 비용 절감, 커스터마이징 측면에서 강점을 지니며, 온프레미스(On-Premise) AI 구축을 촉진
- * Open Source Initiative의 최신 정의에 따르면 ‘오픈소스 AI’는 코드·가중치·데이터(또는 데이터 정보)까지 접근/수정/공유 자유가 담보돼야 하며 오픈웨이트만 공개한 모델과는 차이가 존재

2. 개방형 AI 모델의 발전 과정

- 최근 개방형 AI 모델은 오픈소스 모델과 오픈웨이트 모델을 포괄하는 개념으로 사용되고 있으며, 폐쇄형 AI 모델과 대비되는 용어로 활용
 - 오픈소스 모델은 모델의 구조, 파라미터, 훈련 코드, 데이터셋 정보, 라이선스 문서 등이 공개되어 누구나 사용 및 수정할 수 있는 모델들을 의미
 - 오픈웨이트 모델은 학습된 가중치(파라미터)만 공개하는 모델이며, 파인튜닝이나 추론은 가능하나 원본과 동일한 재학습은 불가능
 - 폐쇄형 모델의 경우 API나 서비스 호출을 통해 모델을 활용할 수 있으나 모델의 구조에 직접 손댈 수 없는 반면, 개방형 AI 모델은 파인튜닝과 모델에 따라 재학습이 가능한 것이 차이

67) 분석 내용은 손영선(2025), “개방형 AI 모델 동향 분석”으로 ICT통계분석 25-04호에 발표되었음

- 개방형 AI 모델은 직접 다운로드하여 로컬에서 작동하기에 기본적으로 외부 전송이 필요 없으며 설정에 따라 완전 오프라인 운용이 가능하나, 폐쇄형 모델의 경우 외부 서버로 데이터 전송이 필요
- 다만 로컬에서의 제어를 위해 강력한 컴퓨팅 파워가 필요하며, 모델 파인튜닝과 보안 유지 등을 위해 관련 지식이 필요

〈표 2-14〉 개방형 AI 모델과 폐쇄형 AI 모델 비교

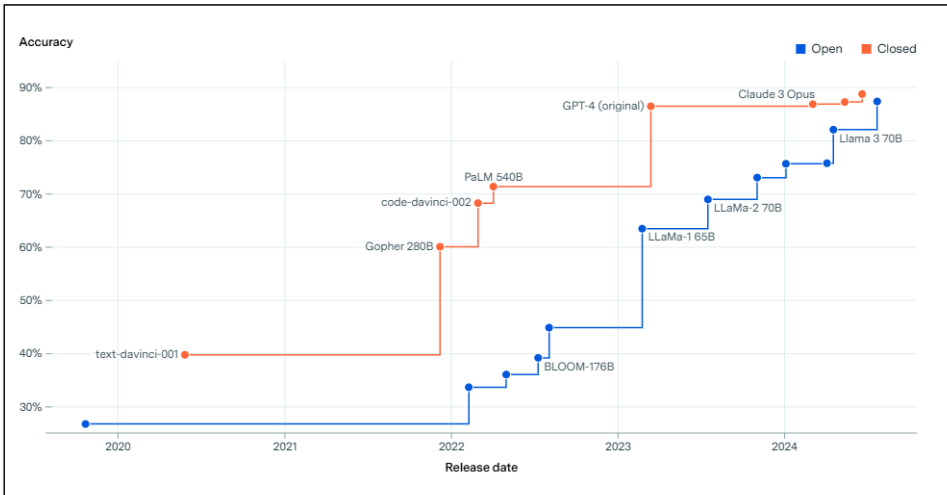
구분	개방형 AI 모델	폐쇄형 AI 모델
접근 방식	모델을 직접 다운로드하여 로컬 환경에서 사용	서비스 또는 API 등 클라우드 기반 접근
커스터마이징 가능여부	파인튜닝 가능 (오픈소스의 경우 재학습 가능)	프롬프트 또는 API 옵션 수준의 제한적인 커스터마이징 가능
정보 공개 여부	전체 구조 및 작동 원리 공개	내부 구조 및 가중치 비공개
데이터관리	로컬 서버에서 직접 제어	기본적으로 외부 서버 전송 필요

자료: 연구진 작성

- 초기 개방형 AI 모델은 오픈소스 기반으로 출발했으나, 2020년 OpenAI의 GPT-3를 계기로 폐쇄형 형태로 전환되는 추세를 보임
 - 2017년 구글에서 Transformer 아키텍처를 공개하면서 LLM 연구가 전환점을 맞이함
 - 2018년 6월에 OpenAI는 Transformer 아키텍처를 기반으로 한 첫 번째 GPT 모델인 GPT-1를 공개
 - 이후 BERT, T5 등 다양한 오픈소스 모델들이 공개되었음
 - OpenAI는 2019년에 GPT-2를 오픈소스로 공개 후, 2020년에 GPT-3를 발표하면서 API 전용으로만 모델을 제공하여 폐쇄형으로 전환
 - OpenAI는 폐쇄형으로의 전환이 상업화 및 경쟁 우위, 안전성 우려 등을 복합적으로 고려한 결과라고 밝힘
 - GPT 모델의 상업화와 함께 Anthropic, Cohere, 구글 등 주요 빅테크 기업들도 상업적인 관점에서 폐쇄형 모델 제공 비중을 확대하는 추세를 보임

- 2020년 이후 개방형 AI 모델에서도 오픈소스/오픈웨이트 방식으로 다양한 모델들이 공개되었지만, 모델 간 성능 격차가 컸음
 - BLOOM-176B, OPT 66B 등 새로운 모델이 공개되었으나 GPT-3와 GPT-3.5 등 대표적인 폐쇄형 모델에 비해 성능 차이가 존재하였음
 - 주요 벤치마크(MMLU, GPQA, GSM8k, BBH) 기준으로 개방형 AI 모델과 폐쇄형 AI 모델간 가장 큰 차이는 25개월까지 벌어진 적이 있었으며, 2018-2024년 7월 기간 평균적으로 5~22개월의 격차가 존재
 - ※ MMLU는 지식/추론, GPQA는 고난이도 논리, GSM8k는 수학, BBH는 다양한 Reasoning에 대한 벤치마크를 의미

[그림 2-14] 개방형 AI 모델과 폐쇄형 AI 모델간 MMLU 벤치마크 비교



출처: Epoch AI(2024.11.04.), "How far behind are open models?"

- 2023년 Meta에서 발표한 LLaMA는 개방형 AI 모델의 혁신을 일으킨 발화점이 됨
 - 2023년 발표된 LLaMA 1은 매개변수의 크기 대비 뛰어난 성능을 보였으며, 소형 모델의 가능성을 입증
 - 2023년 발표된 LLaMA 1은 매개변수의 크기 대비 뛰어난 성능을 보였으며, 소형

모델의 가능성을 입증

- RMSNorm 사용, 프리-노멀라이제이션(Pre-normalization) 구조 채택, SwiGLU 활성화 함수, 회전식 위치 임베딩 등의 최적화 기술을 사용하여 모델의 성능을 향상
- 본래 Meta는 제한적인 조건 하에서 신청자에게만 LLaMA의 모델 가중치와 코드를 제공하였으나, 배포 직후 LLaMA의 전체 가중치가 인터넷에서 유출되는 사건이 발생
- LLaMA의 의도치 않은 공개 이후 전 세계적으로 LLaMA를 기반으로 한 실험, 튜닝, 파생 모델(스탠퍼드의 Alpaca, UC 버클리의 Vicuna 등)이 폭발적으로 증가하며 개방형 AI 생태계 혁신이 촉발
- 누구나 로컬 환경에서 고성능, 가벼운 아키텍처, 공개 데이터 셋을 활용한 투명성 등의 장점을 지닌 개방형 모델을 활용할 수 있게 되었으며, 개방형 AI 생태계가 확산
- Meta 또한 이후 LLaMA 2(2023년 7월), LLaMA 3(2024년 4월) 등의 후속모델에서는 코드와 가중치를 공개하는 식으로 개방형 AI 모델로 방향을 전환
- 이미지 생성 분야의 경우, 2022년 Stability AI에서 발표한 Stable Diffusion이 개방형 텍스트-이미지 생성 모델로 뛰어난 성능을 보여 관련 분야의 혁신을 주도

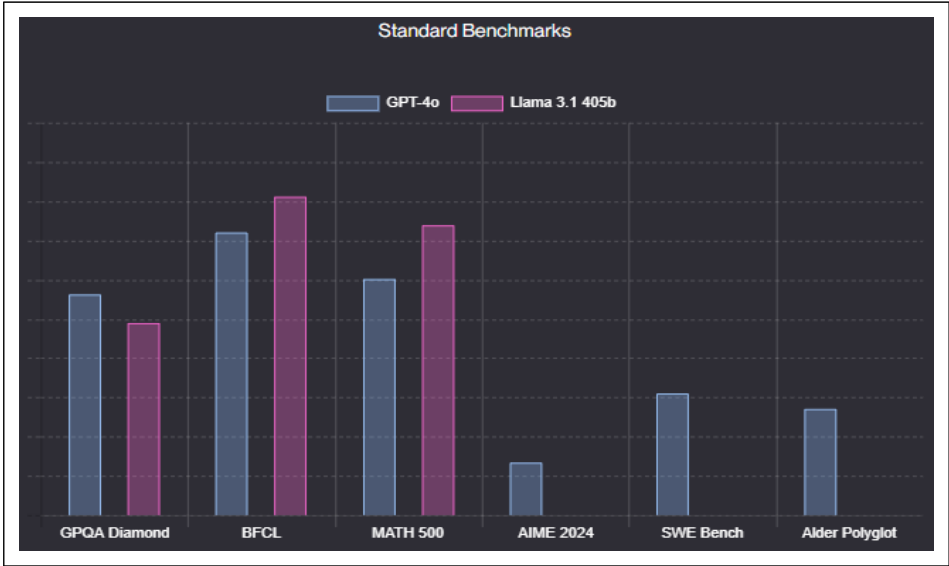
〈표 2-15〉 LLaMA 1의 벤치마크 결과

		BoolQ	PIQA	SIQA	HellaSwag	WinoGrande	ARC-e	ARC-c	OBQA
GPT-3	175B	60.5	81.0	-	78.9	70.2	68.8	51.4	57.6
Gopher	280B	79.3	81.8	50.6	79.2	70.1	-	-	-
Chinchilla	70B	83.7	81.8	51.3	80.8	74.9	-	-	-
PaLM	62B	84.8	80.5	-	79.7	77.0	75.2	52.5	50.4
PaLM-cont	62B	83.9	81.4	-	80.6	77.0	-	-	-
PaLM	540B	88.0	82.3	-	83.4	81.1	76.6	53.0	53.4
LLaMA	7B	76.5	79.8	48.9	76.1	70.1	72.8	47.6	57.2
	13B	78.1	80.1	50.4	79.2	73.0	74.8	52.7	56.4
	33B	83.1	82.3	50.4	82.8	76.0	80.0	57.8	58.6
	65B	85.3	82.8	52.3	84.2	77.0	78.9	56.0	60.2

출처: Touvron, Hugo, et al.(2023)

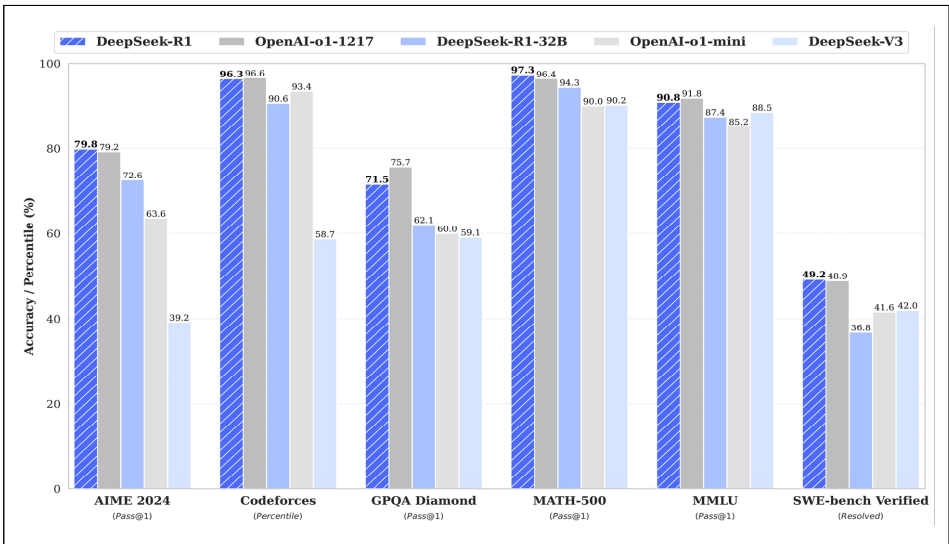
- 이후 개방형 AI 모델의 성능은 빠르게 향상되어 폐쇄형 AI 모델과의 격차를 줄이고 있음
 - Ollama, LM Studio, vLLM 등 로컬 실행 도구들과 Hugging Face와 같은 플랫폼이 급속히 확산되면서 개방형 모델 생태계가 확장
 - 스탠퍼드의 Alpaca, UC 버클리의 Vicuna 등 LLaMA를 파인튜닝하여 우수한 성능을 보이는 뛰어난 파생 모델들이 등장하였음
 - LLaMA 이후 여러 단체와 스타트업이 자체 LLM 개발에 뛰어들었으며 Mistral AI의 Mistral, TI의 Falcon, Alibaba의 Qwen과 같은 우수한 모델들이 공개
 - 구글(Gemma), 마이크로소프트(Phi) 등 빅테크 기업들도 개방형 모델의 개발 및 공개에 참여
 - 2024년 7월에 발표된 LLaMA 3.1 405B는 여러 주요 벤치마크에서 폐쇄형 AI 모델과의 격차를 1년 이내로 단축하면서 가장 극적으로 좁힌 사례로 소개
 - 2025년 발표된 DeepSeek-R1은 비용 대비 높은 효율성과 우수한 벤치마크 점수를 통해 개방형 AI 모델의 새로운 가능성을 제시
 - 한국에서도 업스테이지의 SOLAR 10.7B(2023년 12월), LG AI 연구원의 엑사원 4.0(2025년 7월), 네이버의 하이퍼클로바 X SEED(2025년 4월) 등을 공개하며 개방형 AI 모델 개발에 동참

[그림 2-15] LLaMA 3 70B와 GPT-4o 벤치마크 비교



출처: Vellum(2024.05.08.), "Llama 3 70B vs GPT-4: Comparison analysis"

[그림 2-16] DeepSeek과 OpenAI-o1 벤치마크 비교



출처: DeepSeek-AI, "DeepSeek-R1"

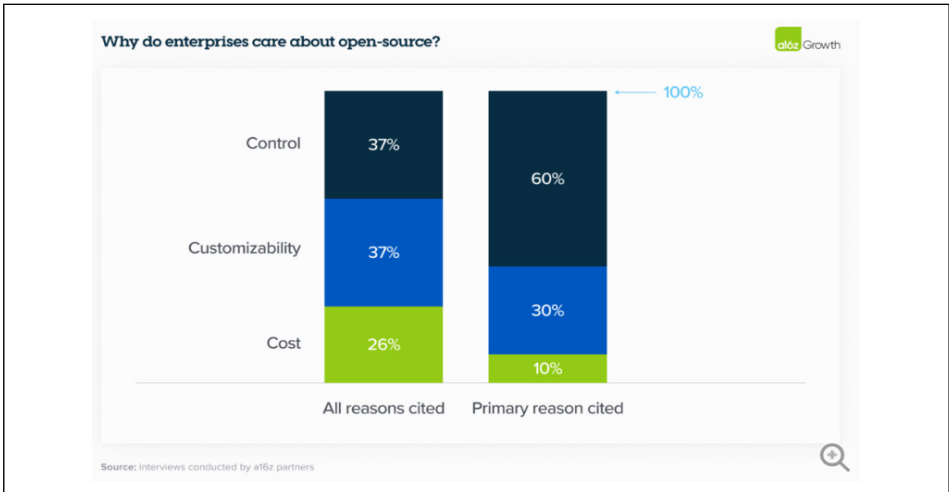
- 2025년 8월 OpenAI에서 오픈웨이트형 개방형 AI 모델인 gpt-oss를 공개한 것은 개방형 AI 모델 생태계에 상징적 사건이 되었음
 - GPT-3부터 상업용 폐쇄형 AI 모델로 방향을 선회한 OpenAI는 2025년에 모델의 가중치 등을 공개한 gpt-oss를 발표
 - 비록 gpt-oss가 파인 튜닝은 허용하나 내부 구조는 완전히 공개하지 않는 절충적 방식을 내세웠지만, 커스터마이징이 가능하다는 점에 의의가 있음
- 생성형AI 서비스를 제공하는 기업들은 생태계 주도권 확보와 표준화 경쟁 등의 목적으로 개방형 AI 모델이 가지는 가치를 높게 평가
 - 개방형 AI 모델을 통한 커스터마이징에 기반한 생태계가 향후 새로운 AI 에이전트와 워크플로우를 만들 가능성이 높으며, 이 과정에서 자사의 모델이 널리 활용될 수록 주도권을 가질 가능성이 높음
 - 수많은 사용자의 활용 사례를 통해 원본 모델에 대한 다양한 피드백을 빠르게 접할 수 있으며, 이를 통해 기술 혁신의 속도를 높일 수 있음
 - AI 민주화의 가치를 통한 긍정적 이미지 제고와 소수 빅테크의 AI 파워 독점에 대한 우려를 해소하는 도구로 작용
 - 다만 모델 가중치가 공개되어 악의적 파인튜닝이 가능하고, 보안 취약점 분석이 용이해져 악용될 소지가 있으며 라이선스 위반 위험 등의 법적 리스크도 존재

3. 개방형 AI 모델과 온프레미스 추세

- 개방형 AI 모델의 성능 향상과 확산은 온프레미스, 즉 사내 AI 구축이라는 추세의 바탕이 되고 있음
- 온프레미스 구축은 데이터 보안, 비용 효율성, 커스터마이징 등의 목적으로 활용 되는 경우가 많음
 - 기업은 외부로 유출되지 않아야 하는 민감한 데이터를 보유하고 있는 경우가 많은데, 온프레미스 형태를 활용할 경우 데이터를 보호하면서 AI 시스템을 구축 할 수 있음
 - 특히 규제가 강한 산업(금융, 의료 등)에서는 기밀 정보 보호를 위해 사내 AI 시스템 활용을 선호

- 초기 구축에 많은 시간과 비용이 소요될 수 있으나, 구축 이후에는 유료 API 호출 없이 AI를 사용할 수 있기에 장기적으로 비용 감소에 기여할 수 있음
- 기업이 필요로 하는 전문 분야 지식에 어울리는 도메인 특화 AI를 튜닝할 수 있기에, 기업의 요구사항에 적합한 전문 AI를 만들 수 있음
- 다만 온프레미스 구축을 위한 기술의 난이도, 초기 인프라 투자, 사내 데이터의 품질 이슈 등은 장벽이 될 수 있음
- o a16z의 2024년 기업 AI 리더 대상 설문조사 결과는 기업들의 오픈소스 모델 선호도가 높아지고 있음을 상징적으로 보여주고 있음
 - 오픈소스 모델 사용 확대에 관심을 보이는 리더는 60%에 달했으며, 46%는 오픈소스 모델을 선호한다고 응답
 - 2023년에는 폐쇄형 모델 80% - 개방형 모델 20% 활용에서 2024년에는 50% - 50% 분할을 목표로 하는 기업들이 증가
 - 오픈소스 모델을 사용하는 주된 이유로는 제어권과 커스터마이징, 비용 순으로 나타남
 - ※ a16z survey는 실리콘밸리의 벤처캐피털인 Andreessen Horowitz(a16z)가 기획하고 진행하는 기업 AI 관련 설문조사

[그림 2-17] 기업이 오픈소스 모델을 사용하는 이유



출처: a16z(2024.03.21.), “16 Changes to the Way Enterprises Are Building and Buying Generative AI”

- 온프레미스 구축 필요성이 증가함에 따라 현장에서도 자체적인 AI 시스템을 구축하는 사례들이 증가하고 있음
 - Bloomberg는 2023년부터 자체 개발한 금융 특화 LLM을 BloombergGPT라는 이름으로 활용하고 있으며, Goldman Sachs도 2025년 6월부터 자체 AI 시스템을 배포
 - 이들은 데이터 보호를 위해서 자체 LLM 개발이 필요했다고 강조하고 있으나, 순수 자체 개발(Bloomberg) / 멀티모델 활용(Goldman Sachs)과 같이 접근 전략에는 차이가 있음
 - 삼성전자 반도체 사업부는 2025년 4월 메타의 LLaMA 4를 도입하여 반도체 설계 및 제조에 활용
 - 삼성SDS는 2025년 7월 IBM의 오픈소스 모델 그라니트 3.1의 도입 검토를 위해 내부 테스트를 진행
 - 미래에셋증권은 2024년 9월 네이버클라우드의 하이퍼클로바X 대시를 기반으로 한 금융권 전용 소형언어모델을 온프레미스 형태로 구축하여 활용

4. 개방형 AI 키워드 텍스트 분석

- (분석 개요) 개방형 AI에 대한 동향 변화를 파악하기 위해 관련 뉴스 기사를 수집하여 텍스트 분석을 실시하였음
 - 개방형 AI 관련 기사 수집을 위해 개방형 AI, 오픈소스LLM, 오픈웨이트LLM, mistral, DeepSeek, 메타 LLaMA, 구글 gemma, 알리바바 qwen, 네이버 하이퍼클로바X, LG 엑사원을 키워드로 활용
 - 개방형AI, 오픈소스LLM, 오픈웨이트LLM은 개방형 AI와 관련된 전반적인 키워드 도출을 위해 활용
 - mistral, DeepSeek, 메타 LLaMA, 구글 gemma, 알리바바 qwen은 인지도 및 활용도가 높은 대표적인 개방형 AI 모델로 포함하였고, 일부 키워드는 관련 없는 기사 수집을 막기 위해 기업명을 함께 검색하였음
 - 국내 개방형 AI 모델로 하이퍼클로바X와 LG 엑사원을 포함

- 2023년 1월 ~ 2025년 8월까지 관련도가 높은 기사 중심으로 수집하였고, 총 13,849개의 기사로 분석을 진행

※ 2023년은 LLaMA의 공개로 개방형 AI 모델의 혁신이 가속화된 시점

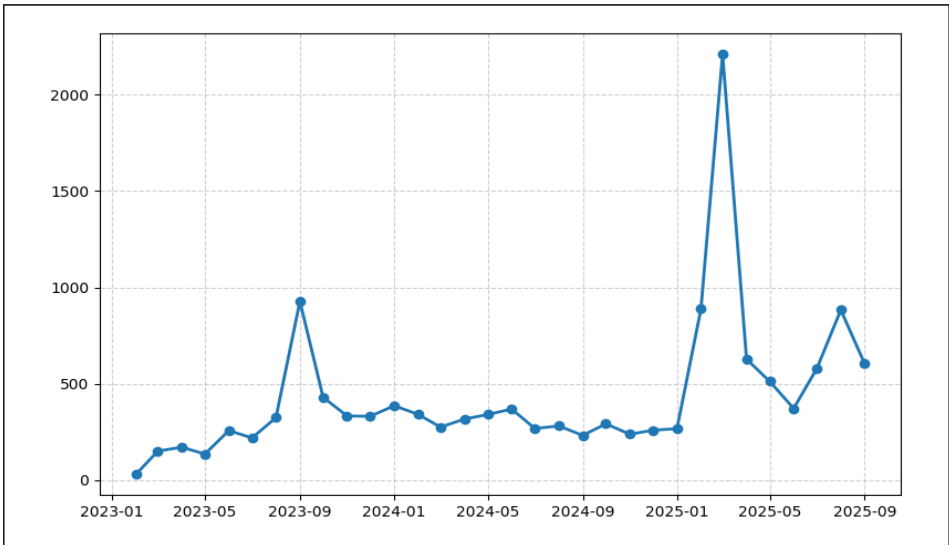
○ (언급량) 월별 기사 추이를 보면, 특정 시점(2023년 9월, 2025년 3월)에 기사 수가 크게 증가한 것이 특징이며 전반적으로 상승 추세를 보임

- 2023년 초반 점진적인 상승을 보이던 기사 수는 2023년 8월에 큰 폭으로 증가하였으며, 하이퍼클로바X의 공개 등 국내 모델과 관련된 기사가 증가한 것이 영향을 주었음

- 2024년에는 큰 변동 없이 유사한 수준을 유지하였으나, 2025년 초 다시 큰 폭의 상승을 보임

- 특히 2025년 2월에 기사 수가 크게 증가하였는데, DeepSeek 모델이 발표되며 관련 기사가 급증한 것이 영향을 주었음

[그림 2-18] 월별 기사수 추이



자료: 연구진 분석자료

- (상위 키워드) 각 연도별로 등장 빈도가 높은 상위 키워드를 살펴보면 다음과 같음
 - 국내 기사를 주로 수집하였기에 전반적으로 네이버, 카카오, LG 등 개방형 AI와 관련된 국내 주요 기업들의 등장 빈도가 높았음
 - 국내 모델에서는 하이퍼클로바X의 언급 빈도가 높은 편에 속함
 - 메타, 구글과 같이 대표적인 AI 관련 기업들도 많이 언급되었으며, 제품으로는 GPT에 대한 언급량이 가장 많았음
 - 2025년의 경우 딥시크에 대한 언급 빈도가 가장 높았으며, 중국과 미국이 상위 키워드로 등장하였음. 이는 딥시크 출현 이후 미중의 AI 기술 패권 경쟁이 심화된 상황을 보여주는 결과라 판단됨
 - 이를 잘 보여주는 것이 2025년에 “반도체”가 상위 키워드로 등장한 것이며, 딥시크 출현 이후 고성능 학습용 반도체가 미중 패권경쟁 구도에서 핵심 키워드로 떠오른 상황을 잘 표현하고 있음
 - 데이터, 클라우드, 플랫폼같은 키워드의 등장 빈도도 매우 높았으며, 이는 AI 활용에 있어 이러한 요소들이 핵심 인프라임을 보여주는 결과
 - 2023년~2024년 기간에는 주로 기업, 사업, 시장과 같이 비즈니스 측면에서의 단어들 많이 등장한 반면, 2025년에는 정부, 지원과 같은 단어들의 등장 빈도가 증가하였음. 이는 2025년 들어 정부 차원에서 AI에 관심이 높아진 상황을 반영하였다고 볼 수 있음
 - 전략이라는 단어 또한 2025년에 등장 빈도가 증가하였는데, 앞서 해석과 연결하면 정부 차원에서의 전략을 의미하는 단어로 보는 것이 적절할 것으로 판단됨

〈표 2-16〉 연도별 등장 빈도 상위 키워드

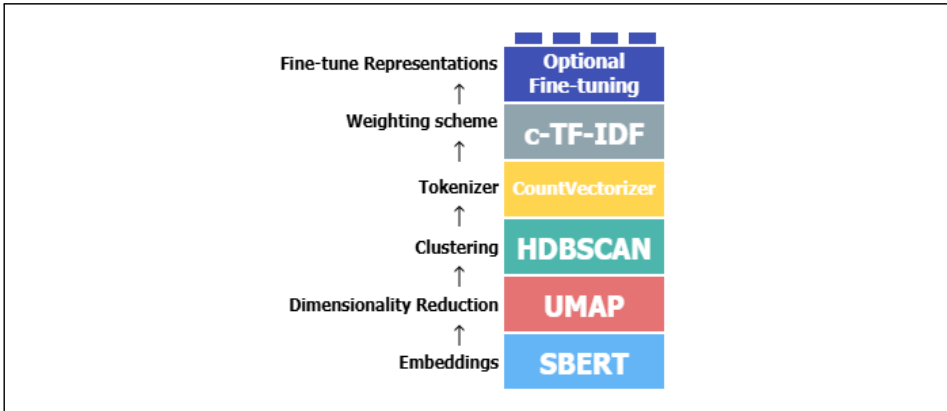
2023년	2024년	2025년
네이버	네이버	딥시크
서비스	서비스	중국
기업	기업	기업
GPT	데이터	데이터
데이터	LG	미국
X & 하이퍼	클라우드	LG
시장	사업	서비스
플랫폼	시장	네이버
LG	플랫폼	정부
카카오	X & 하이퍼	시장
사업	GPT	산업
구글	구글	투자
투자	카카오	전략
출시	메타	반도체

주: 연도별로 상위 키워드 중 의미가 큰 단어들을 정리하였음. AI, 인공지능, 모델, 오픈소스, LLM 등도 상위 키워드에 속하나 해당 키워드들은 개방형 AI를 의미하는 단어들로 판단, 주는 의미가 크지 않다고 판단하여 결과에는 포함하지 않았음. 2025년은 8월까지 수집된 기사의 분석 결과임

자료: 연구진 분석자료

- (토픽 분석) 뉴스기사 텍스트를 기반으로 주요 주제와 흐름을 식별하고 분석하기 위해 BERTopic 분석을 수행하였음
 - BERTopic은 텍스트 데이터 내 주제 구조를 파악하기 위해 사용되는 방법
 - 기존 토픽 분석에 자주 사용된 LDA 기반 모델보다 문맥 이해 능력이 우수한 것으로 알려져 있음
 - 임베딩 - 차원 축소 - 클러스터화 - 키워드 추출 - 후보 정제의 절차로 단순 빈도가 아닌 의미 맥락에 기반하여 추출한 핵심 키워드 후보군을 제공

[그림 2-19] BERTopic 알고리즘



주: 본문 분석에서는 Clustering 단계에서 KMeans를 활용
출처: Grootendorst, "The Algorithm - BERTopic"

- 토픽 분석 결과 제품(하이퍼클로바X, 엑사원, LLaMA) 관련 토픽이 등장하였으며, 답시크와 관련된 토픽과 정책 관련 토픽이 주를 이루었음
 - 하이퍼클로바X 토픽에는 하이퍼클로바X, 네이버, 네이버클라우드, 서비스 등의 단어가 포함되어 있음
 - 네이버와 카카오가 하나로 연결된 토픽도 존재하였는데, 두 기업이 국내 IT 업계의 대표 기업만큼 뉴스에서도 함께 언급되는 빈도가 많았을 것으로 추측됨
 - 엑사원 토픽에서는 LG, LG 유플러스, 엑사원이 포함되어 있었으며, KT 또한 함께 언급되어 통신사 관련 토픽으로 해석할 수도 있음
 - LLaMA 토픽에서는 메타, 저커버그, CEO와 같은 키워드가 포함되어었으며, MS 관련 단어로 연결되었음
 - 답시크발 미중 기술 패권을 함의하는 토픽도 등장하였는데, 여기에는 미국, 중국, 엔비디아, 반도체, 관세 등이 등장하여 답시크에 따른 기술 충격과 이에 따른 국가간 경쟁을 보여주고 있음
 - 더불어 답시크와 개인정보가 연결된 토픽이 등장하였는데, 이는 답시크 활용과 개인정보 노출 관련 이슈와 연결된 토픽으로 판단됨. 여기에는 답시크, 접속, 개인정보, 차단, 보안, 중국 등이 포함되어 있음

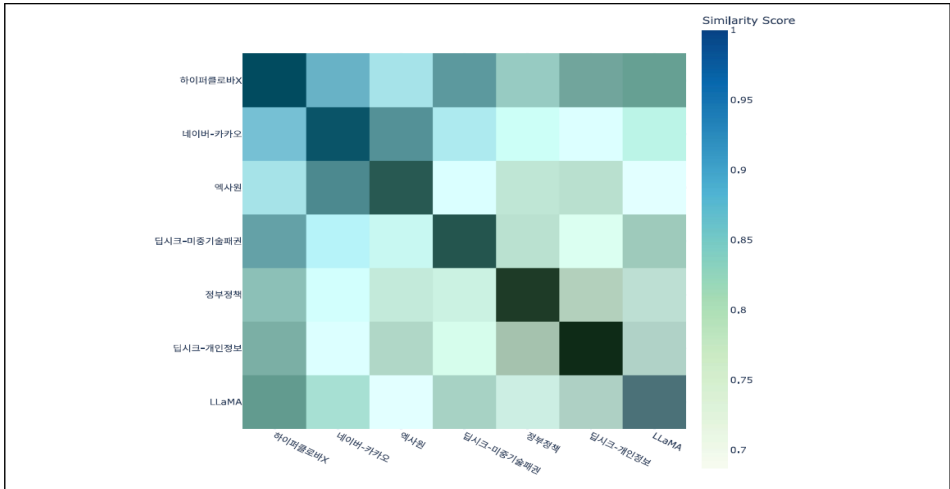
- 글로벌, 장관, 국가, 인공지능이 묶인 토픽도 존재하며, 이는 AI와 관련된 국가 차원에서의 관심도 증대를 보여주는 토픽으로 볼 수 있음
- 전반적으로 개방형 AI 기사에서는 각 제품군 - 기업군간 유사한 토픽으로 묶이는 경향이 많았으며, 답시크 출시에 따른 파급효과와 국가 차원에서의 AI 관심도 증대를 보여주는 토픽들이 등장하였음

〈표 2-17〉 토픽별 주요 키워드

하이퍼클로바X	네이버-카카오	엑사원	답시크-미중기술패권	정부정책	답시크-개인정보	LLaMA
네이버	네이버	엑사원	미국	글로벌	답시크	메타
하이퍼클로바X	검색	LG	엔비디아	기술	접속	저커버그
서비스	카카오	LG유플러스	반도체	장관	개인정보	라마
네이버클라우드	서비스	모델	중국	산업	중국	MS
생성형	대비	사업	관세	국가	차단	CEO
		기술	답시크		보안	
		KT				

자료: 연구진 분석자료

〔그림 2-20〕 토픽별 유사도 행렬



주: 색깔이 진할수록 두 토픽간 유사도가 높음을 의미

자료: 연구진 분석자료

5. 결론

- 최근의 AI 생태계는 폐쇄형 모델 중심의 경쟁에서 개방형 모델 기반의 협력·확산 체계로 확장 중
 - GPT-3 이후 시작된 폐쇄형 모델의 상업화 흐름이 2023년 LLaMA를 기점으로 새 국면을 맞이
 - 개방형 모델은 기술 자율성과 투명성, 데이터 보안, 비용 효율성을 바탕으로 AI 주권 확보 및 산업 내 자립 역량 강화의 핵심 축으로 부상
- 개방형 AI 모델의 발전은 단순한 기술 공개를 넘어 AI 생태계 구조 자체를 재편
 - 오픈소스 및 오픈웨이트 기반의 모델들은 누구나 접근 가능한 혁신 인프라로 기능
 - 글로벌 기업뿐 아니라 스타트업·학계·국가 연구기관까지 AI 개발 참여의 문턱을 낮춤
 - 이러한 개방성은 기술 민주화뿐 아니라 AI 기술의 표준화 경쟁 구도로 이어짐
- 개방형 AI 모델의 확산은 온프레미스(On-Premise) AI 구축을 촉진
 - 데이터 보호·규제 대응·도메인 특화 모델 구축의 요구가 맞물리며, 기업이 내부적으로 모델을 운용·관리하는 방향으로 전환
 - 폐쇄형 API 의존을 줄이고, 장기적 비용 절감과 내부 역량 축적을 가능하게 함
- 글로벌 주요 기업들은 개방형 AI를 전략적 자산으로 인식하고 경쟁적으로 생태계 확장 중
 - Meta, Google, Microsoft 등 빅테크 기업들도 개방형 모델을 통해 시장 영향력 확대
 - 한국 또한 하이퍼클로바X, 엑사원, SOLAR 등을 통해 자체 기술 역량을 축적
 - 개방형 생태계 참여를 통한 국가 AI 기술 자립도 제고가 중요한 과제로 부상
- 텍스트 분석 결과, 개방형 AI 담론은 기술 중심에서 정책·산업 전략 중심으로 이동
 - 2025년 이후 정부, 전략, 반도체, 미중 경쟁 등 거시적 키워드가 급부상
 - 이는 개방형 AI 모델이 더 이상 연구·기술 이슈가 아닌, 국가 차원의 산업정책 및 기술패권 논의의 핵심의제로 자리 잡았음을 시사
 - 딥시크 출시 이후로 다양한 토픽(미중기술패권, 개인정보 등)이 추가로 등장하며, 딥시크 파급효과에 따른 AI 생태계의 담론 변화가 뚜렷하게 나타나는 것도 특징

- 개방형 모델은 향후 다양한 측면에서 중요도가 높아질 것으로 전망
 - AI 주권(Sovereign AI)의 실현 수단으로서의 주목도 상승
 - 개방형 AI 생태계 조성을 위한 공공 데이터 개방, AI 인프라 확충, 윤리·보안 가이드라인에 대한 관심도가 커질 것으로 보임
 - 향후 AI 산업의 경쟁력은 단일 모델의 우위뿐만 아니라, 개방성과 협력 네트워크를 통한 집단적 혁신 역량에서 결정될 것으로 전망

제 3 장 ICT 뉴스경기지수 구축 및 고도화

제 1 절 개 요

- 본 장에서는 비정형 데이터인 뉴스 기사를 활용하여 구축한 ICT 뉴스경기지수를 소개하고 주요 결과를 설명
- 비정형 데이터를 통해 경제 현황을 확인하고 단기 전망에 활용하려는 연구는 다양하게 진행되고 있음
 - Baker et al.(2016), Scott and Varian(2014)가 대표적인 연구며 Baker et al.(2016)은 경제정책 관련 기사를 통해 정책의 불확실성을 확인하였고, Scott and Varian(2014)은 실업률과 물가의 초단기 변동을 측정하였음
- 최근 연구들의 경우, 비정형 데이터 중 뉴스 기사의 가능성에 주목하여 경제환경을 측정하고자 시도하고 있음
 - IMF 및 각 국가의 중앙은행은 대중의 경제 상황 인식을 확인하기 위한 방법으로 뉴스 기사에 기반한 심리지수를 만들고 있으며, 한국은행도 2021년부터 뉴스심리지수를 공개 중에 있음(서범석 외, 2022)
 - 뉴스 기사는 지속적으로 업데이트가 이루어지므로 시의성이 높으며 자료의 양이 방대하기에 활용 가치가 높다고 평가받고 있음
 - 구축한 지수 결과를 공개하는 것 뿐만 아니라, 이 값을 경제전망을 위한 하나의 변수로 사용하고자 하는 연구도 다양하게 진행 중에 있음
- 이런 관점에서, ICT 관련 뉴스 기사를 통해 ICT 산업의 경기를 시의성 있게 파악할 수 있는 지표를 작성한다면 ICT 산업의 현황 분석과 전망에도 많은 기여를 할 것으로 보임
- 본 연구에서는 ICT 산업 관련 뉴스를 수집하고 ICT 산업의 경기 상황을 확인할 수 있는 뉴스경기지수를 작성하였으며, 이를 소개하고자 함

제 2 절 ICT 뉴스경기지수 작성 방법론⁶⁸⁾

1. 키워드 선정 및 자료수집

- ICT 뉴스경기지수 작성을 위해서는, ICT 관련 뉴스를 정의할 수 있는 적절한 기준을 마련하는 것이 선행되어야 함
- 본 연구에서는 ICT 관련 뉴스를 대표할 수 있는 키워드를 선정하고, 해당 키워드가 포함된 기사를 ICT 관련 뉴스로 분류하였음
 - ICT 산업의 전반적인 현황을 직접 나타내는 키워드(ICT 기업, ICT 산업 등)를 기초로 하여 과학기술정보통신부, Gartner 등 공신력 있는 기관과 기업에서 발표하는 ICT 트렌드를 참고하여 키워드 후보들을 선정
- Google Trends 및 네이버 Datalab을 통해 ICT 키워드 후보를 ① 과거부터 현재까지 꾸준히 언급된 키워드, ② 최근에 나타난 키워드로 분류하였음
 - ‘과거부터 현재까지 꾸준히 언급된 ICT 키워드’는 2010년대 이후~2020년대까지 일정 수준 이상의 검색량을 유지하는 키워드를 의미⁶⁹⁾
 - 네이버 Trend의 ‘ICT’ 평균 검색량(67.5)과 평균 분산(212.5)을 기준으로 ‘ICT’ 평균 검색량이 비슷한 수준이거나 관측된 기간 내 일정한 수준을 유지하는 키워드, 그리고 ‘ICT’의 평균 분산과 비슷한 수준이거나 더 낮은 키워드를 선정하였음
 - ‘최근에 나타난 ICT 키워드’는 2010년대에는 검색량이 미미하였으나 2020년대에 검색량이 증가한 키워드를 의미⁷⁰⁾

68) 본고에서는 손녕선 외(2023)에서 연구한 방법론 및 모델을 활용하여 ICT 뉴스경기지수를 작성한 과정 및 주요 결과를 소개하였음. 모델에 대한 검증 과정에 대한 설명은 손녕선 외(2023)을 참고.

69) 과학기술정보통신부(2020), 「2020년 10대 전략 기술 트렌드」, 정보통신기획평가원, ICT 동향정보, <https://www.iitp.kr/kr/1/knowledge/organScrapList.it>, ZDNET Korea(2023. 2. 20.), “ICT 연구개발에 올 1.6조 투입, AI 양자 등 6대 분야 주력”.

70) Sciencetimes(2022. 12. 26.), “2023년 ICT 이슈를 전망하다 : 디지털 전환을 넘어 디지털 대도약 시대”, Samsung SDS(2023. 1. 25.), “미리 예측해보는 2023년 IT 기술 트렌드”.

- 네이버 Trend의 'ICT' 검색량 평균(67.5), 분산(212.5)을 기준으로 'ICT' 검색량 평균보다 낮거나 평균 분산보다 높은 키워드를 선정
- 최근에 등장한 키워드들은 검색량 평균은 상대적으로 낮을 것이나, 변동성은 클 것으로 예측
- ICT 산업을 전반적으로 대표하는 8개 키워드와 과거부터 꾸준히 언급된 ICT 키워드 5개, 최근에 나타난 ICT 키워드 5개를 추가하여 총 18개 키워드를 ICT 키워드로 선정

〈표 3-1〉 키워드 리스트

ICT 산업 관련	과거부터 꾸준히 언급된 키워드	최근에 나타난 키워드
ICT 산업	ICT 네트워크	ICT 6G
ICT 기업	ICT 보안	ICT 메타버스
ICT 수출	ICT 로봇	ICT 모빌리티
ICT 품목	ICT 우주	ICT 플랫폼엔지니어링
IT 산업	ICT 플랫폼	ICT 디지털안전
IT 기업	-	-
IT 수출		
IT 품목		

자료: 연구진 작성

- 선정한 키워드를 중심으로 관련 기사를 수집하기 위해 네이버 뉴스를 활용
 - 뉴스 기사와 댓글 트렌드를 분석하기 위하여 2008년 1월 1일부터 2025년 11월 30일까지의 기간 동안 네이버에 게재된 뉴스 기사 전체를 수집하여 분석에 활용
 - 보고서에서는 분석의 안정성을 위해서 기사 수가 충분히 제공되는 2015년부터의 결과를 제시

2. ICT 뉴스경기지수 작성 방법

- ICT 뉴스경기지수는 뉴스기사에 대한 감성분석을 기반으로 하며, 긍정/부정 키워드에 따라 지수의 값이 움직이는 형태를 지님
- 해당 기간 수집된 기사들 중 긍정 관련 기사가 많은 경우 뉴스경기지수의 값은 높아지고, 부정적인 기사가 많을 경우 감소하는 형태를 지님
- 구체적인 작성 방법은 다음과 같음
 - 뉴스기사를 문장 단위로 나눈 후, 각 문장별로 긍정 및 부정 키워드 단어 수를 비교
 - 단어 수가 더 많은 쪽으로 해당 문장을 긍정 혹은 부정으로 분류
 - 분류한 기사를 지수화하기 위한 방법은 한국은행에서 작성하는 뉴스심리지수(NSI) 계산 방법을 참고하였음(서범석 외(2022))

$$\text{뉴스경기지수} = \left(\frac{X - \bar{X}}{S} \right) \times 10 + 100$$

$$X = \frac{\text{기간내긍정기사(문장)수} - \text{기간내부정기사(문장)수}}{\text{기간내긍정기사(문장)수} + \text{기간내부정기사(문장)수}}$$

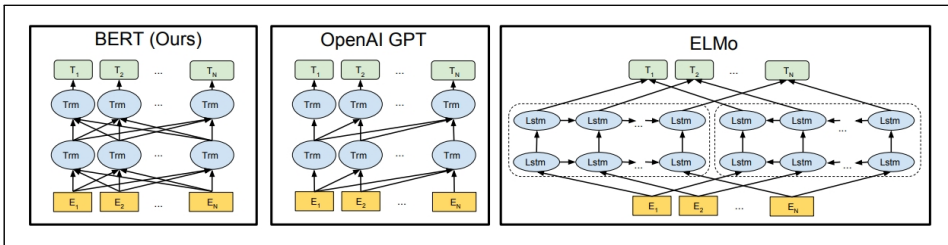
$$\bar{X} = \text{표준화구간평균}, S = \text{표준화구간표준편차}$$

- 기간은 7일을 설정하였으며, 기간 내 문장 수는 당일 기준으로 최근 7일의 긍정 및 부정 문장수의 평균값을 의미
- 위 식을 통해 계산한 ICT 뉴스경기지수 값이 100을 초과하는 경우 긍정적인 내용이 많았음을 의미하고, 반대로 100 미만인 경우 부정적인 내용이 더 많았음을 의미
- 표준화 구간은 2008년 - 2023년 4월 30일로 설정하였음
- 이상의 방법을 통해 일별 단위 ICT 뉴스경기지수를 계산하였고, 일별 결과를 월 단위로 평균하여 월별 ICT 뉴스경기지수를 작성하였음
 - 일별 단위 ICT 뉴스경기지수는 변동성이 큰 편이며, 기사 수가 부족한 날짜에서는 값을 측정하기 어려운 한계점이 있기에 월별 단위 ICT 뉴스경기지수를 기본 단위로 활용하였음

3. 문장 분류: KoBERT

- ICT 뉴스경기지수 산출에 필요한 문장의 긍정/부정 라벨링을 위해서 KoBERT 모델을 활용하였음
 - BERT는 2018년 구글이 공개한 NLP 기법 중 하나로, 레이블이 붙어 있지 않은 대용량 자연어 데이터를 학습하여 문맥을 이해하고 단어의 의미를 파악할 수 있는 사전 학습 모델
 - 대량의 텍스트를 사전 훈련(pre-trained)해서 성능을 높였고 단어의 확률 예측시 전체 문장을 파악 가능하도록 양방향을 사용한다는 점에서 강점이 있음

[그림 3-1] BERT와 타 자연어 처리 모델 비교



출처: 인공지능신문(2019.01.03.), “인공지능(AI) 언어모델 BERT(버트)는 무엇인가”

- KoBERT는 SK Telecom에서 개발한 모델이며, BERT의 한국어 성능의 한계를 극복하기 위해 수백만 개의 한국어 문장을 학습한 모델임
- KoBERT 학습을 위해 사전 기반 라벨링을 진행하였음
 - 사전 기반 라벨링은 외부의 단어 사전(lexicon)에서 긍정, 중립, 부정 등으로 라벨링 된 단어들을 이용해 키워드에 대한 라벨링을 수행하는 방식
- 사전 기반 라벨링을 위해서 한국은행의 뉴스심리지수에서 측정된 부문별 긍부정 키워드(서범석 외(2022)), 썬트렌드(<https://some.co.kr/>)에서 ICT 산업, ICT 기업, ICT 수출 관련 기사에 대한 긍부정 분석에서 등정한 단어, KNU 감성사전, 한국은행 사전(김현중 외(2019), 전종준 외(2020)), KOSELF를 조합하여 활용하였음

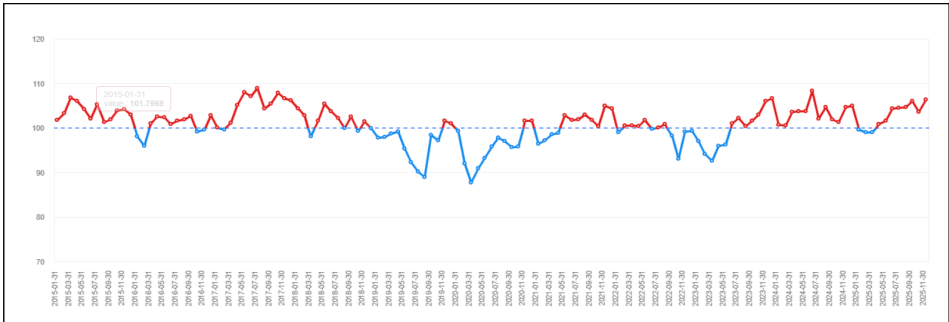
- KNU 한국어 감성 사전은 인간의 보편적인 기본 감정 표현을 나타내는 긍정어 및 부정어로 구성되었으며, 각 도메인의 감성 사전 구축에 활용되는 기초 자료 목적으로 개발되었음
- 김현중 외(2019)에서는 속성이 정의된 온라인 뉴스 기사와 Im et al.(2018)에 소개된 어휘사전 생성 알고리즘을 활용하게 자체 감성사전을 구축
- 전종준 외(2020)에서는 각 심리 주제(기업, 소비자, 기타)별로 감성이 있다고 표시된 기사 및 문장을 선택하여 감성단어를 추출
- KOSELF는 기업 재무분석에 특화된 한국어 금융 사전으로, 한경 컨센서스에서 제공하는 2016년부터 2018년까지 발행된 약 2만 개의 애널리스트 보고서 자료를 활용하여 사전의 유효성을 검증
- KoBERT 학습에는 2008년 1월 1일부터 2023년 4월 30일까지의 ICT 기사를 활용하였음
 - 수집한 기사에 대해 토큰화(Tokenization), 표제어 추출(Lemmatization), 품사 태깅(POS), 불용어 제거(Stopword) 전처리를 진행
 - 토큰화는 텍스트를 특정 단위로 나누는 과정이며, 한국어의 경우 형태소 단위로 토큰화를 수행하는 경우가 많음
 - 표제어 추출은 어근을 추적하여 토큰의 개수를 줄이고 정리하는 과정을 의미
 - 품사 태깅은 각 토큰들에 품사를 매칭하고, 이후 분석에 사용하고자 하는 품사들만 추출하는 과정을 의미
 - 불용어 제거는 불필요한 토큰들을 제거하는 과정이며, 보통 분석에 큰 의미를 주지 않는 조사나 접속사를 제거하게 됨
- 사전 기반 라벨링을 수행한 약 430만개의 문장을 70%의 training set, 30%의 test set으로 구분하여 학습을 진행
 - Binary Cross Entropy를 손실 함수로, AdamW를 optimizer로 설정하고 학습을 진행하였고 training set에서는 98.6%, test set에서는 98.2%의 정확도를 보임
 - 전체 예측 수 중 옳게 예측한 비율을 나타내는 Accuracy Score는 98.28%를 기록함

제3절 ICT 뉴스경기지수 구축 결과

- 2015년 1월부터 2025년 11월까지의 뉴스 기사를 활용하여 ICT 뉴스경기지수를 구축한 결과는 다음과 같음
 - 2015년부터 2018년까지는 ICT 관련 뉴스에 긍정적인 이슈가 많았던 것으로 나타났음
 - 반대로 2019년 이후부터는 부정적 이슈가 긍정적 이슈를 앞선 것을 확인할 수 있으며, 특히 2019년 8월과 2020년 3월에 최저점을 기록하였음
 - 2021년 이후 다시 상승 추세를 보이던 ICT 뉴스경기지수는 2022년 9월 이후로 다시 하락하는 모습을 보임
 - 2023년 6월 이후 긍정적인 기사가 늘어나는 것을 확인할 수 있으며, 2024년의 경우 꾸준히 100 이상의 값을 기록하고 있음
 - 2025년 1월~3월에 감소세를 보였으나 4월 이후부터 긍정적인 이슈가 늘어나며 100 이상의 값을 꾸준히 유지
 - 이를 통해 2025년에는 1분기를 제외하고 ICT 산업 전반적으로 긍정적인 신호가 많았음을 추측할 수 있음
 - 2015년~2018년까지 ICT 산업의 생산과 수출이 증가하였다는 점, 2019년과 2020년에는 역성장이 발생했다는 점 등을 고려하면⁷¹⁾, ICT 뉴스경기지수의 추세가 그 당시의 ICT 산업의 경기를 잘 반영하고 있는 것으로 판단
 - 2020년 상반기의 급격한 하락세는 코로나로 인한 영향이 반영된 것으로 보임
 - ICT 뉴스경기지수가 BSI에 선행하는 모습을 확인한 손영선 외(2023)의 결과를 고려하면, ICT 뉴스경기지수를 통한 경기 측정이 의미 있는 함의점을 제공하는 것으로 판단

71) 생산액은 과학기술정보통신부의 ICT 실태조사를, 수출액은 과학기술정보통신부·정보통신기획평가원의 ICT 수출입동향을 참고하였다.

[그림 3-2] ICT 뉴스경기지수 결과



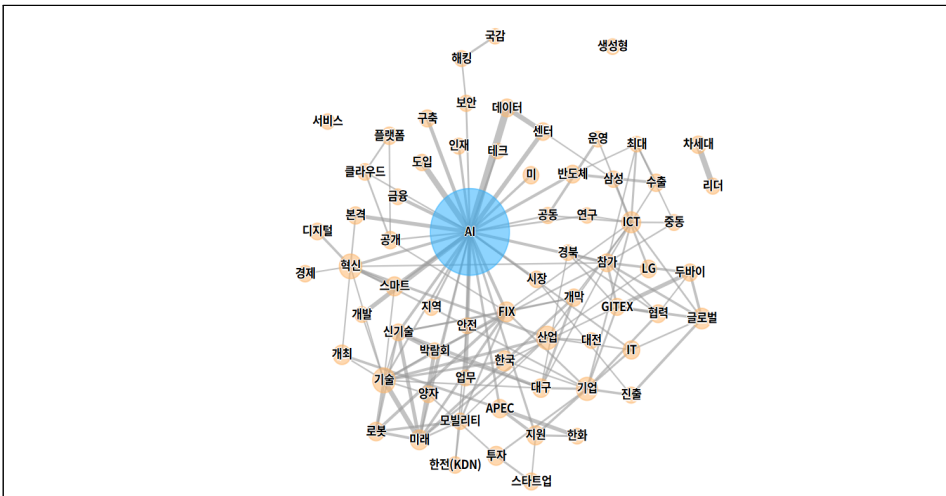
주: 2015년 1월부터 2025년 11월까지의 뉴스기사를 활용하여 작성한 ICT 뉴스경기지수 결과. 지수값이 100이 넘으면 긍정적인 기사가 많았음을 의미하며, 100 미만이면 부정적 기사가 많았음을 의미.

출처: ITSTAT, <https://www.itstat.go.kr/itstat/kor/economic/economicList.html>

- 뉴스경기지수의 활용도를 높이기 위해, ICT 뉴스기사를 활용하여 네트워크 분석을 실시
 - 네트워크 분석은 관계 분석에 사용되는 도구로 사람, 기업, 웹사이트, 단어, 문장 등 다양한 대상에 활용 가능
 - 객체들 간의 관계를 이해하고, 객체들간 연결되어 있는 패턴 등을 분석하여 함의점을 제공
 - 뉴스기사를 대상으로 한 네트워크 분석은 해당 시점 ICT 관련 기사들 중 어떤 키워드의 중요도가 높은지와 이 키워드들이 어떠한 관계를 맺고 있는지에 대한 정보를 제공
 - 이를 통해, 해당 시점 ICT 산업 내에서 중요하게 여겨지는 주제가 무엇인지를 파악할 수 있으며 이 주제들이 어떤 방식으로 상호작용하고 있는지를 파악하여 함의점을 제공할 수 있음
- 본고에서는 뉴스기사의 제목을 사용하여 네트워크 분석을 수행하였음
 - 네트워크 내 노드(단어)의 중요도를 측정하는 방식으로는 구글에서 개발한 PageRank를 사용
 - PageRank에서는 다른 객체로부터 링크를 많이 받을수록 중요도가 높은 객체로 간주되며, 특히 중요한 객체로부터 링크를 받을수록 해당 객체의 중요도가 높아지는 구조를 가지고 있음

- PageRank 기준으로 상위 60개 단어들을 중심으로 네트워크 분석 결과를 작성하였으며, 이는 웹페이지에서 네트워크 분석 결과에 대한 가독성을 높이기 위함임⁷²⁾
- 2025년 10월의 ICT 뉴스 기사를 활용하여 네트워크 분석을 수행한 결과는 다음과 같음
- 특히 AI 노드가 가장 크고 연결된 예지가 많은 것으로 보아 다양한 주제의 기사들이 AI와 연관성이 높은 것을 볼 수 있었으며 이를 통해 AI가 ICT 트렌드의 핵심 역할을 맡고 있음을 파악할 수 있음

[그림 3-3] 네트워크 분석 결과: 2025년 10월



주: 2025년 10월 ICT 관련 기사를 사용하여 네트워크 분석을 수행한 결과임. 단어를 둘러싼 원이 클수록 해당 단어가 네트워크에서 중요도가 높음을 의미하며, 단어를 연결한 예지가 두꺼울수록 해당 단어 간 관계가 중요하다는 것을 의미.

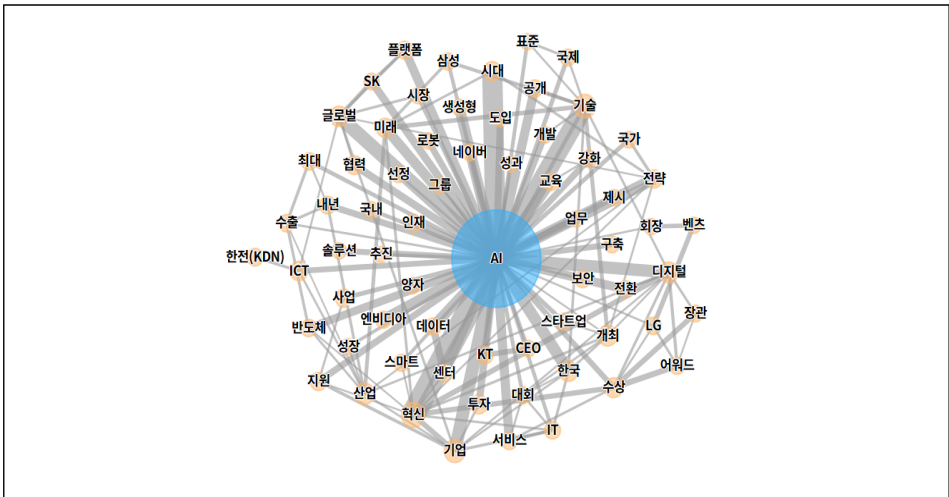
출처: ITSTAT, <https://www.itstat.go.kr/itstat/kor/economic/economicList.html>

- 박람회, 개막, 참가 등의 단어들이 등장하여 해당 기간 동안 ICT 관련 행사들이 진행되었음을 알 수 있으며, 우리나라에서 개최된 제33회 APEC도 키워드로 등장하게 됨

72) 최초 분석 결과를 검토하고, 이 과정에서 의미가 없는 단어들이 상위에 위치할 경우 이를 제외하고 다음 순위의 단어를 올리는 작업을 수행하였다.

- 보안, 해킹, 국감 단어가 등장하며 AI와 관련한 중요한 이슈로 인식되고 있음을 알 수 있음
- 2025년 11월의 ICT 뉴스기사를 활용하여 네트워크 분석을 수행한 결과는 다음과 같음
 - 10월과 마찬가지로 AI가 중요도가 가장 높은 키워드로 등장하였으며, AI에 대한 관심이 꾸준히 유지되고 있음을 확인 가능
 - 그 외 기술, 혁신, 시대의 등장 빈도가 높았으며 예지 중요도가 높은 것으로 나타나 AI와 밀접한 관계를 지님을 보여줌
 - 국제, 표준, 국가, 전략 등의 키워드가 등장한 것으로 보아 AI 관련 제도적 논의와 대응이 대내외적으로 이루어지고 있음을 확인 가능

[그림 3-4] 네트워크 분석 결과: 2025년 11월



주: 2025년 11월 ICT 관련 기사를 사용하여 네트워크 분석을 수행한 결과임. 단어를 둘러싼 원이 클수록 해당 단어가 네트워크에서 중요도가 높음을 의미하며, 단어를 연결한 에지가 두꺼울수록 해당 단어 간 관계가 중요하다는 것을 의미

출처: ITSTAT, <https://www.itstat.go.kr/itstat/kor/economic/economicList.html>

제 4 절 국내 최신 트렌드를 반영한 ICT 키워드 발굴

1. 개요

- 본 절에서는 최신 기술 발전과 산업 트렌드를 반영하여 ICT 관련 주요 키워드를 발굴하기 위한 연구 결과를 제시
 - 2023년에 최초로 ICT 트렌드와 관련된 키워드를 수집한 후 ICT 뉴스경기지수 수집을 위한 기반으로 활용해 왔음
 - 본 연구에서는 기존에 활용되었던 키워드가 여전히 ICT 산업의 변화 흐름을 적절히 반영하고 있는지를 점검하고, 키워드의 활용 가능성에 대한 정량적·정성적 분석을 통해 그 유효성을 재검토하였음
 - 더불어 최근의 기술 진보, 산업 구조의 변화, 사회적 관심사 등을 종합적으로 고려하여 기존 키워드를 보완하거나 대체할 수 있는 신규 키워드를 발굴
 - 발굴한 키워드는 ICT 뉴스경기지수 산출을 위한 수집 키워드 최신화에 활용
- ICT 뉴스경기지수의 정책적 활용성과 민간 동향 파악 기능을 지속적으로 유지·강화하기 위해서는 산업 및 기술별 변화를 반영한 키워드 체계의 주기적인 재점검과 최신화가 필수적임
 - 특히 기존 키워드의 유효성을 평가하고, 새로운 트렌드를 반영한 키워드를 추가 발굴함으로써 ICT 동향 분석의 최신성을 확보할 수 있을 것으로 기대

2. 연구범위 및 방법론

가. 연구범위

- ICT 관련 키워드의 최신화 및 대표성 제고를 위해 다음과 같은 범위 내에서 연구를 진행
- 첫째, 2023년에 발굴한 주요 ICT 키워드의 변화 양상을 분석하여 현시점에서 과거 ICT 키워드들이 여전히 유효한지 검토
 - Google Trends를 통해 검색량 수준 및 트렌드 변동성을 정량적으로 분석하였으며, 검색 기간은 최대 2004년 1월부터 2025년 7월까지 활용

- 둘째, 2025년 신규 키워드 선정을 위해 ICT 산업 전반과 관련 기술 분야를 포괄하되, 다수 기관의 관심이 집중된 주요 세부 영역을 중심으로 키워드를 구성
 - 산업의 구조적 변화 및 정책 이슈의 흐름을 반영하고자, 최근 수년간의 기술 동향, 보도자료, 공신력 있는 기관의 트렌드 보고서를 종합적으로 검토
- 셋째, 키워드는 단순히 최신성만을 기준으로 선정하지 않고, 각 기술 및 산업 분야를 대표할 수 있는 상위 개념 수준의 키워드를 중심으로 구성하였으며, 이를 통해 ICT 분야 전반에 대한 포괄성과 지속 가능성을 확보하고자 하였으며 지나치게 협소하거나 일시적인 기술 편중을 지양하였음
 - 생성형 AI와 같이 주요 분야를 안정적으로 대표할 수 있는 키워드를 중심으로 선정하여 지수 산출의 일관성과 해석의 안정성을 높이고자 함
 - 선정된 키워드에 대해서 Google Trends의 5개년 데이터(2020년 8월 ~ 2025년 7월)를 사용하여 정량적 검증을 수행
- 넷째, 분석의 초점은 국내 ICT 산업에 맞추었으며, ICT 뉴스경기지수와 직접적으로 연계 가능한 국내 정보 출처 기반의 키워드를 중심으로 검토를 진행
 - 뉴스 기반 정성 지표인 ICT 뉴스경기지수와 기업 체감 기반 정량 지표인 BSI 간의 상관관계 및 시차 구조를 비교함으로써, ICT 뉴스경기지수가 ICT 산업의 경기 상황을 적시에 반영하거나 선행할 수 있는지를 재검토

나. 키워드 최신화 방법론

- 본 연구에서는 키워드별 검색 관심도의 트렌드 변화를 분석하기 위해 Google Trends 데이터를 활용
 - Google Trends는 전 세계 사용자들의 검색 데이터를 기반으로, 특정 키워드 변화를 시계열로 파악할 수 있는 유용한 도구로 여겨짐
 - 특히, 시간의 흐름에 따른 검색 관심도 추이를 통해 특정 기술이나 이슈에 대한 대중의 인식 변화를 분석하는 데 효과적임
 - 네이버 트렌드는 2016년부터의 데이터만 제공하므로, '트렌드가 지난 키워드'처럼 2010년대 이전부터 장기 흐름을 검토하는 데 한계가 있음
 - 반면, Google Trends는 2004년 이후 데이터를 제공하여 분석 대상 키워드를

장기간에 걸쳐 추적할 수 있으며, 이에 따라 모든 키워드에 대해 일관된 시계열 분석이 가능

- Google Trends에서 제공하는 관심도 수치는 절대 검색량이 아닌, 표준화된 상대 지수로서 0에서 100 사이의 값으로 표시되기에 주의가 필요
 - 해당 지수는 분석 대상 키워드 내에서 검색량이 가장 높았던 시점을 100으로 설정하고, 나머지 시점은 이에 비례하여 상대적으로 산출
 - 이러한 특성으로 인해 단일 키워드의 지수만으로는 해당 키워드의 실제 대중적 관심 수준을 정확히 판단하기 어려움
 - 예를 들어, 특정 키워드가 최근 일시적으로 관심을 끌었을 경우, 전체적인 관심도는 낮음에도 불구하고 지수상 최고치(100)로 나타날 수 있음
- 이에 따라 본 연구에서는 특정 키워드 트렌드의 실질적인 위상과 상대적 관심도 수준을 보다 명확히 파악하기 위해, 기준 키워드를 'ICT'로 설정하고 이를 비교군으로 활용하여 상대 분석을 수행
 - 이를 통해 단순한 검색량 분석을 넘어, 정책적으로 의미 있는 핵심 키워드 여부를 판단하고자 하였음
- 기존에 선정된 키워드는 관심도 추세에 따라 세 가지 유형으로 분류하여 분석
 - 첫째, 트렌드가 지난 키워드는 2010년대 이전에 활발히 언급되었으나 이후 관심도가 지속적으로 감소한 키워드를 의미하며, 이러한 키워드에 대해서는 최근 다시 주목을 받고 있는지를 검토
 - 둘째, 꾸준히 언급되는 키워드는 2010년부터 2023년까지 비교적 안정적인 검색 관심도를 유지해 온 키워드로, 2023년 이후 하락세로 전환되었는지를 중점적으로 평가
 - 셋째, 최근에 나타난 키워드는 2020년대 이후 급격히 증가한 키워드를 의미하며, 이들이 단기 유행에 불과한지 또는 장기적인 지속성을 보이는지를 중심으로 평가
- 신규 키워드 후보군의 도출을 위해 과학기술정보통신부, 한국인터넷진흥원, 삼성 SDS 등 국내 주요 공공기관 및 민간 리서치 기관이 발표한 ICT 관련 보고서 및 보도자료를 기반으로 2025년 ICT 주요 트렌드를 정리

- 이 과정에서 확보된 키워드 후보군은 2020년대 이후의 Google Trends 데이터를 통해 검증하였으며, 특히 2023년 이후 새롭게 관심도가 급증한 키워드인지 여부를 중점적으로 분석
- ICT 뉴스경기지수와 기업경기실사지수(BSI)의 비교를 통해 뉴스 기반 정성 지표와 기업 체감 기반 정량 지표 간의 상관성과 시차 구조를 분석
 - 이는 ICT 산업 경기 진단 및 예측에 있어 ICT 뉴스경기지수의 활용 가능성을 평가하는 데 중요한 의의를 지님
 - 지표 간 비교를 통해 정성 기반 뉴스 지표의 선행성과 보완 지표로서의 활용 가능성이 키워드 최신화 이후에도 유지되고 있는지를 실증적으로 평가하고자 함

3. 기존 키워드의 변화 양상 분석

- 먼저 2023년 연구에서 선정된 ICT 주요 키워드에 대해 Google Trends 기반으로 최신 동향과 변화 양상을 확인하고자 함
 - 2023년에는 ICT를 대표하는 포괄적인 키워드로 ICT 기업, ICT 산업, ICT 품목, ICT 수출, IT 기업, IT 산업, IT 품목, IT 수출을 선정
 - 더불어 다양한 ICT 트렌드를 반영할 수 있는 추가 키워드 선정을 위해 Google Trends의 데이터(2004년 1월~2023년 4월)를 활용하여 트렌드가 지난 키워드, 과거부터 꾸준히 언급된 키워드, 최근 새롭게 등장한 키워드를 구분하여 확장된 키워드 목록을 구성
 - 이번 연구에서는 2020년 8월 ~ 2025년 7월까지의 기간으로 확장하여 각 키워드의 최신 트렌드를 확인하고 여전히 유효한 키워드인지를 검증
- 2023년 기준으로 ICT 뉴스경기지수에 포함된 확장 키워드는 다음과 같음

〈표 3-2〉 키워드 리스트

트렌드가 지난 키워드	꾸준히 언급되는 키워드	최근에 나타난 키워드
유비쿼터스	네트워크	6G
자동화	보안	메타버스
소프트웨어	로봇	모빌리티
LTE	우주	플랫폼 엔지니어링
어플리케이션	플랫폼	디지털 안전

자료: 연구진 작성

가. 트렌드가 지난 키워드

- 트렌드가 지난 키워드는 2010년대 이전에는 많은 관심을 받았지만 이후에 큰 관심을 받지 못하는 키워드를 의미
- 트렌드가 지난 키워드 중에서는 ‘자동화’ 키워드를 제외한 키워드들에서는 유의미한 변화가 없이 정체된 양상을 보였으므로 본 분석에서는 트렌드의 변화가 있었던 ‘자동화’ 키워드에 대해서 집중적으로 검토를 수행

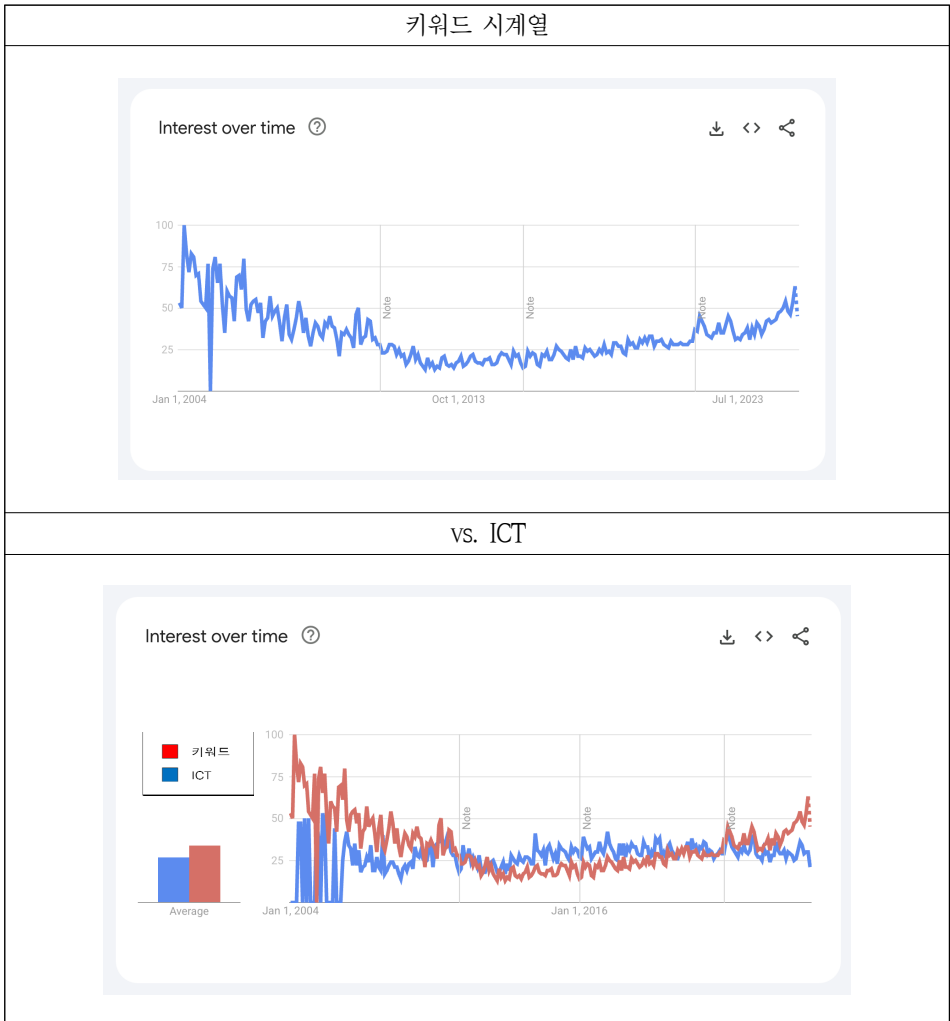
〈표 3-3〉 트렌드가 지난 키워드 변화 양상

키워드	추세 지속 여부
유비쿼터스	하강
자동화	상승
소프트웨어	하강
LTE	하강
어플리케이션	하강

자료: 연구진 작성

1) 키워드: 자동화

[그림 3-5] '자동화' 키워드 시계열 분석 결과



자료: 연구진 작성

- 최근 ‘자동화’ 키워드는 인공지능과 스마트팩토리의 확산과 더불어 다시금 주목을 받고 있음
 - 글로벌 제조업체들은 AI, 머신러닝, 사물인터넷 기반의 스마트 제조 기술에 대한 투자를 확대하고 있으며, 약 95%의 기업이 향후 5년 이내에 AI 기반 자동화를 도입할 계획을 밝히고 있음
 - 이러한 흐름은 단순한 기계적 자동화를 넘어, 예측 정비, 품질 검사, 로봇 셀 제어 등으로 대표되는 지능형 자동화로 확장되는 양상을 보이고 있음
 - 특히 AI가 결합된 자동화는 노동력 부족 문제를 해소하고 생산성과 품질을 동시에 향상시킬 수 있는 핵심 수단으로 부각되고 있으며, 이에 따라 자동화에 대한 사회적 관심과 기술적 수요는 지속적으로 확대될 것으로 예상됨
- 시계열 분석 결과 또한 2012년 이후로 관심도가 지속적으로 상승하는 모습을 보여, 이러한 흐름을 뒷받침하고 있음
 - 기준 키워드인 ‘ICT’와 비교했을 때, 2024년부터는 ‘자동화’ 키워드가 더 높은 관심도를 기록하는 것으로 분석되었음
 - 이러한 결과를 종합적으로 고려할 때, ‘자동화’는 다시 주목받는 트렌드 키워드로 분류할 수 있음

나. 꾸준히 언급되는 키워드

- 꾸준히 언급되는 키워드는 2010년부터 2023년까지 비교적 안정적인 검색 관심도를 유지해 온 키워드로, 꾸준히 언급되는 키워드는 모든 키워드가 관심도에 큰 하락 없이 꾸준한 관심을 받아온 모습을 보임

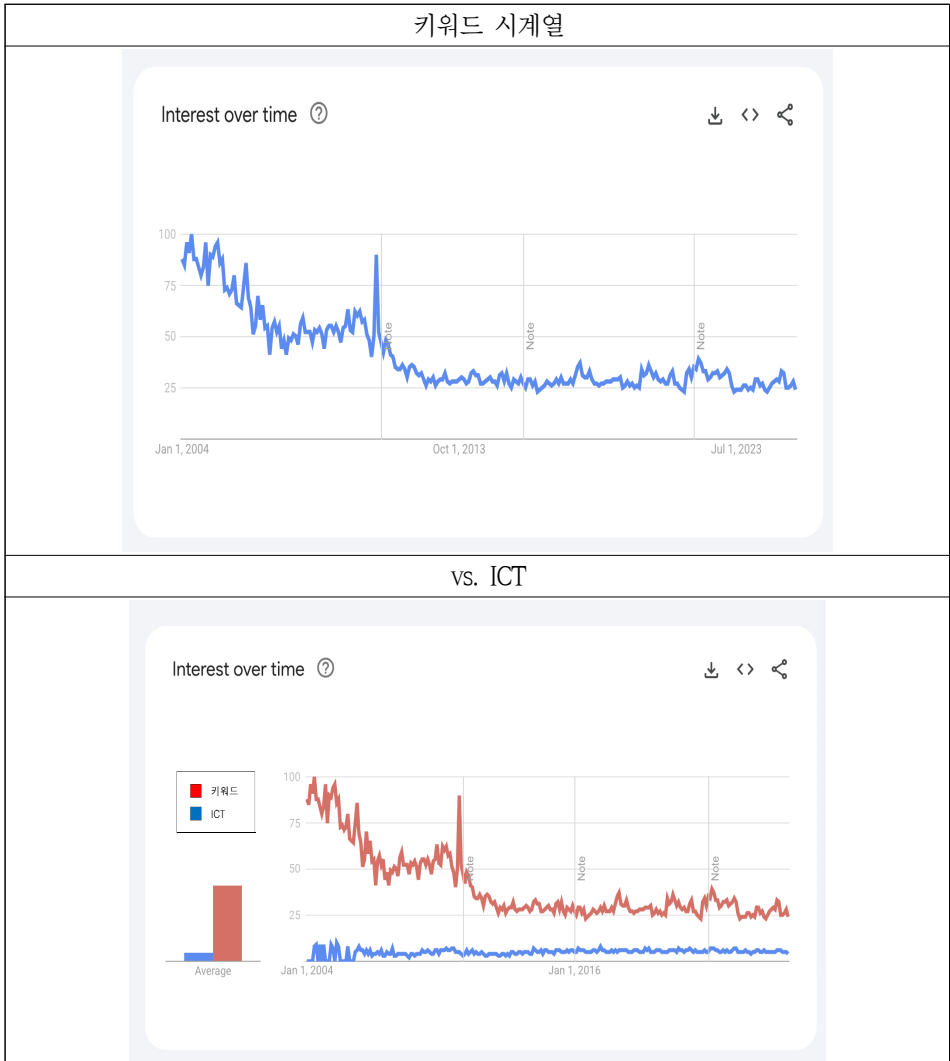
〈표 3-4〉 꾸준히 언급되는 키워드 변화 양상

키워드	추세 지속 여부	키워드 유지 여부
네트워크	수평	유지
보안	상승	유지
로봇	수평	유지
우주	상승	유지
플랫폼	상승	유지

자료: 연구진 작성

1) 키워드: 네트워크

[그림 3-6] '네트워크' 키워드 시계열 분석 결과



자료: 연구진 작성

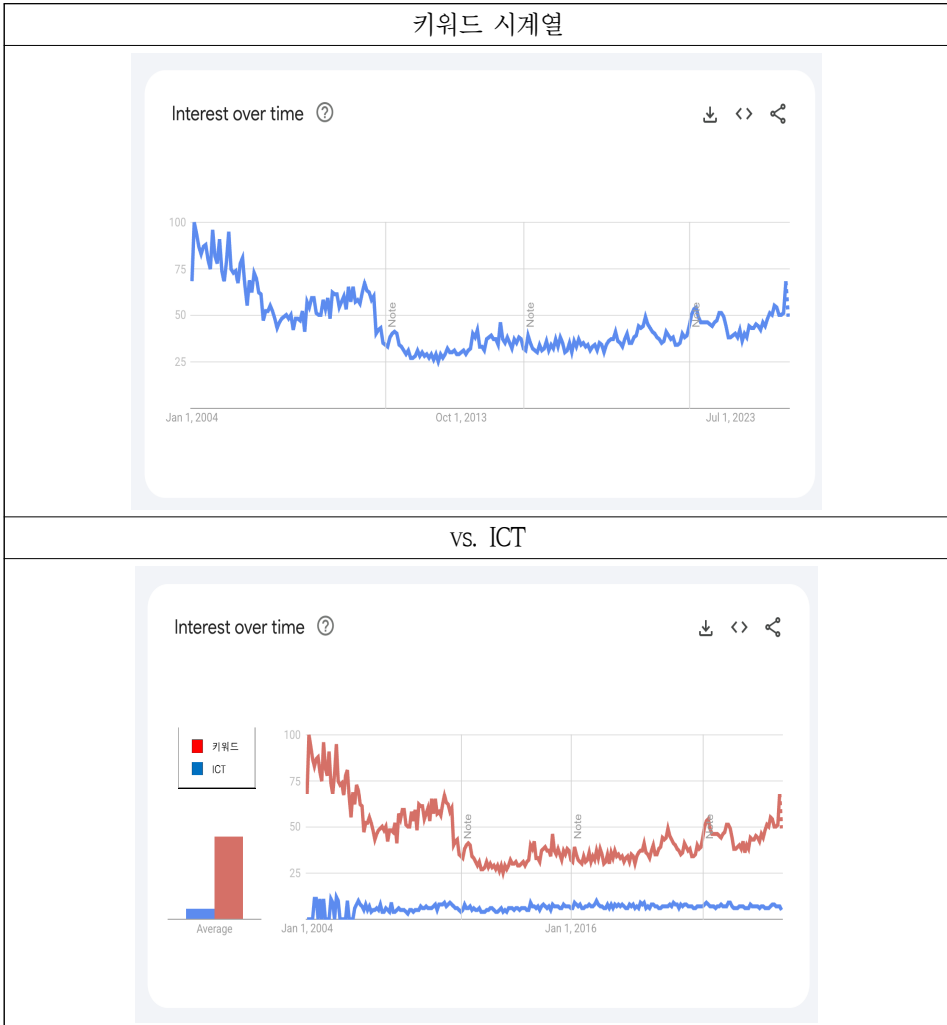
- '네트워크' 키워드는 2010년대 이후부터 일정 수준 이상의 검색량을 지속적으로 유지하고 있음
 - 특히 기존 키워드인 'ICT'와 비교하였을 때에도 전반적으로 높은 검색량을 기록

하고 있는 것으로 나타났음

- 2023년 이후에도 유의미한 관심도 변화 없이 안정적인 추세를 이어가고 있어, 현재까지도 꾸준히 언급되는 키워드로 분류할 수 있음

2) 키워드: 보안

[그림 3-7] '보안' 키워드 시계열 분석 결과

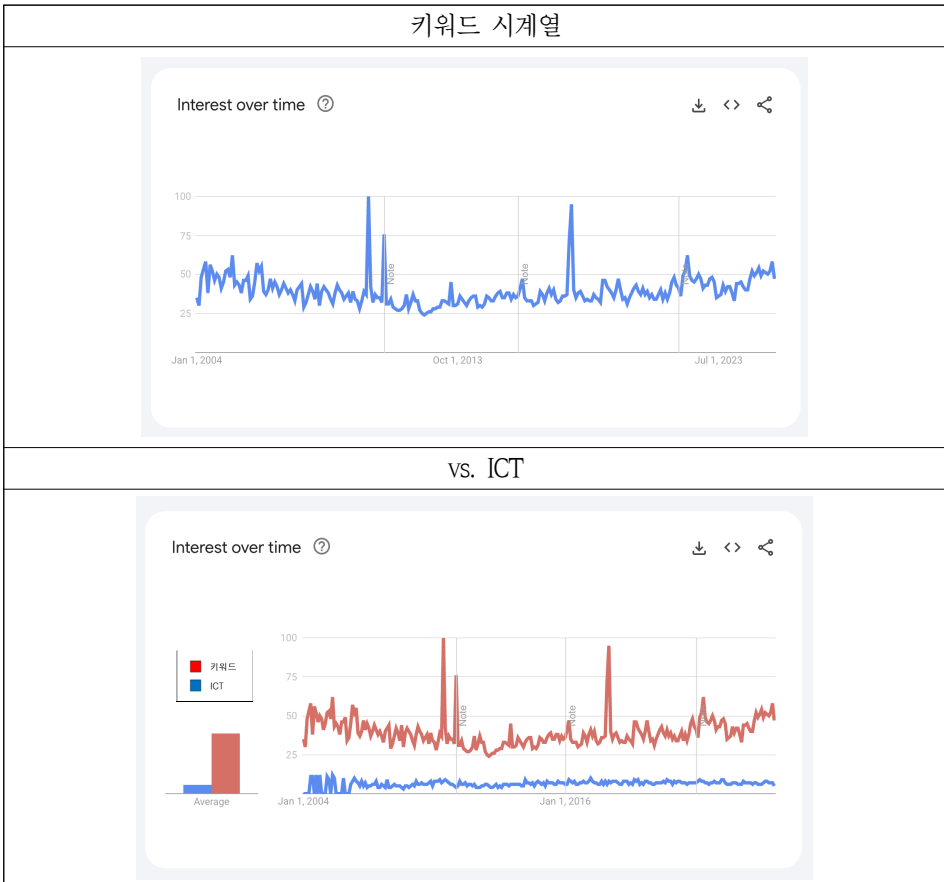


자료: 연구진 작성

- ‘보안’ 키워드는 2010년부터 2023년까지 일정 수준 이상의 관심도를 지속적으로 유지한 것으로 나타났음
 - 이후에는 관심도가 서서히 상승하는 추세를 보이며, 이는 반복적으로 발생한 해킹 사건과 사이버 공격 사례를 통해 사회 전반에 사이버 보안에 대한 경각심이 확산된 결과로 해석됨
 - 특히 2023년 이후에는 관심도가 더욱 상승하는 양상을 보여, ‘보안’ 키워드는 현재까지도 꾸준히 언급되는 키워드로 분류할 수 있음

3) 키워드: 로봇

[그림 3-8] ‘로봇’ 키워드 시계열 분석 결과

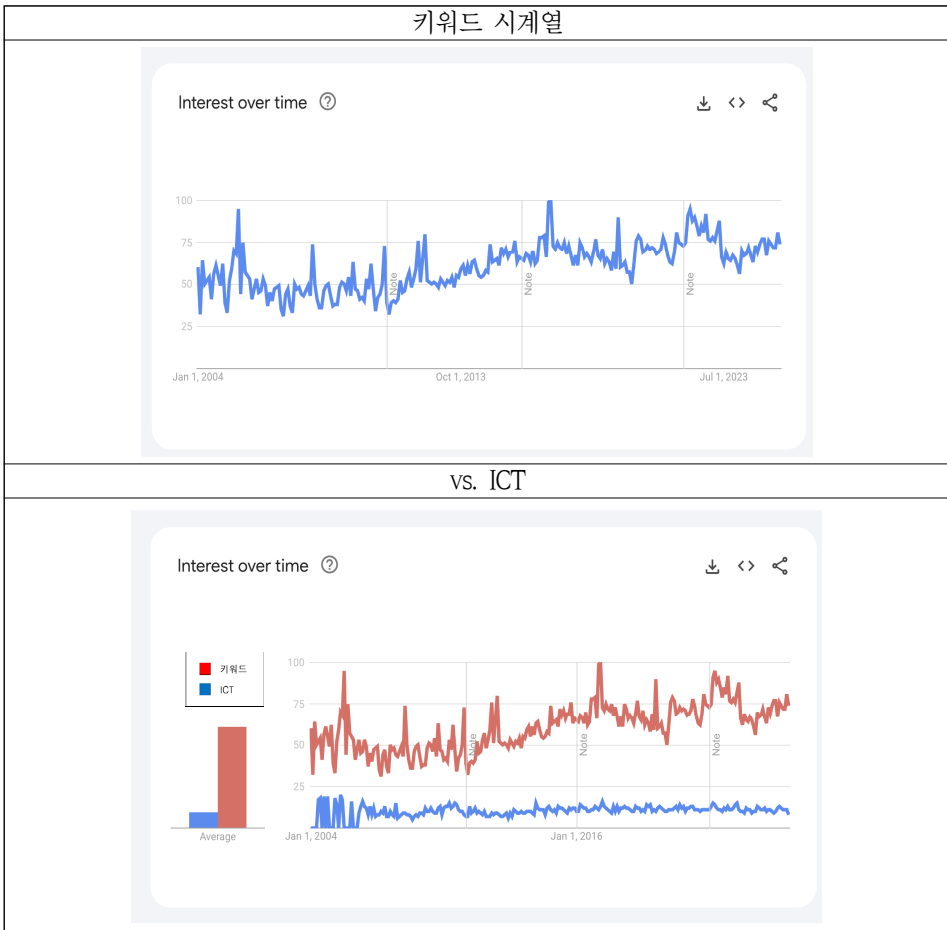


자료: 연구진 작성

- ‘로봇’ 키워드는 일시적인 관심도 급등 시기를 제외하면 전반적으로 안정적인 수평 추세를 보이며, 장기적인 관심도의 지속성이 확인됨
 - 최근에는 ‘Physics AI’와 같은 신기술 키워드의 등장, 테슬라의 로봇택시 공개 등 주요 이슈가 이어지면서 검색량이 다시금 상승하는 양상을 나타내고 있음
 - 이러한 흐름을 종합적으로 고려할 때, ‘로봇’ 역시 현재까지 꾸준히 언급되는 키워드로 분류할 수 있음

4) 키워드: 우주

[그림 3-9] ‘우주’ 키워드 시계열 분석 결과



자료: 연구진 작성

- ‘우주’ 키워드는 장기적으로 완만한 우상향 추세를 보이며, 꾸준한 관심의 지속성이 확인됨
 - 더불어 최근에는 SpaceX의 스타십 테스트 및 발사 성공, 발사 비용 획기적 절감, 민간 우주 투자 확대 등이 시선을 끌며 관련 키워드의 검색량 증가를 이끌고 있음
 - 특히 파이낸셜타임스는 우주산업을 “다음 산업혁명의 무대”로 평가하였음
 - 이러한 흐름을 종합적으로 고려할 때, ‘우주’ 키워드는 2023년 이후에도 꾸준히 언급되는 키워드로 분류될 수 있음

5) 키워드: 플랫폼

[그림 3-10] ‘플랫폼’ 키워드 시계열 분석 결과



자료: 연구진 작성

- ‘플랫폼’ 키워드는 2004년부터 꾸준히 언급되어 왔으며, 특히 2018년을 기점으로 관심이 급격히 증가한 것으로 나타남
 - 이는 당시 배달의민족, 쿠팡, 에어비앤비, 우버 등 국내외 디지털 플랫폼 기업의 급성장이 주목받고, 플랫폼 경제에 대한 논의가 본격화된 흐름과 맞물려 있음
 - 국내의 경우, 배달앱 시장의 급성장과 온라인 유통 플랫폼 간의 경쟁 심화 등 ‘플랫폼’이라는 용어가 다양한 맥락에서 폭넓게 사용되기 시작하였음
- 2023년 이후 관심이 다소 하락하는 양상이 관측되었으나, 여전히 최대 검색량의 약 60% 수준을 유지하고 있으며, 기준 키워드인 ‘ICT’보다도 높은 검색량을 기록하고 있음
 - 이러한 점을 고려할 때, ‘플랫폼’은 최근에도 일정 수준 이상의 주목도를 유지하고 있는 키워드로, 꾸준히 언급되는 키워드로 분류할 수 있음

다. 최근에 나타난 키워드

- ‘최근에 나타난 키워드’는 2020년대 이후 급격한 관심도 상승을 보인 키워드를 의미
 - 본 분석에서는 이러한 키워드가 2023년 이후에도 꾸준한 관심을 유지하고 있는지, 혹은 일시적인 유행에 그쳐 관심이 하락하였는지를 평가함
 - 추세가 지속적으로 상승하고 있는지 혹은 급등 후 하락하고 있는지를 바탕으로 최근에 나타난 키워드로 분류되었던 항목들을 다시 검토하여 키워드 유지 여부를 결정

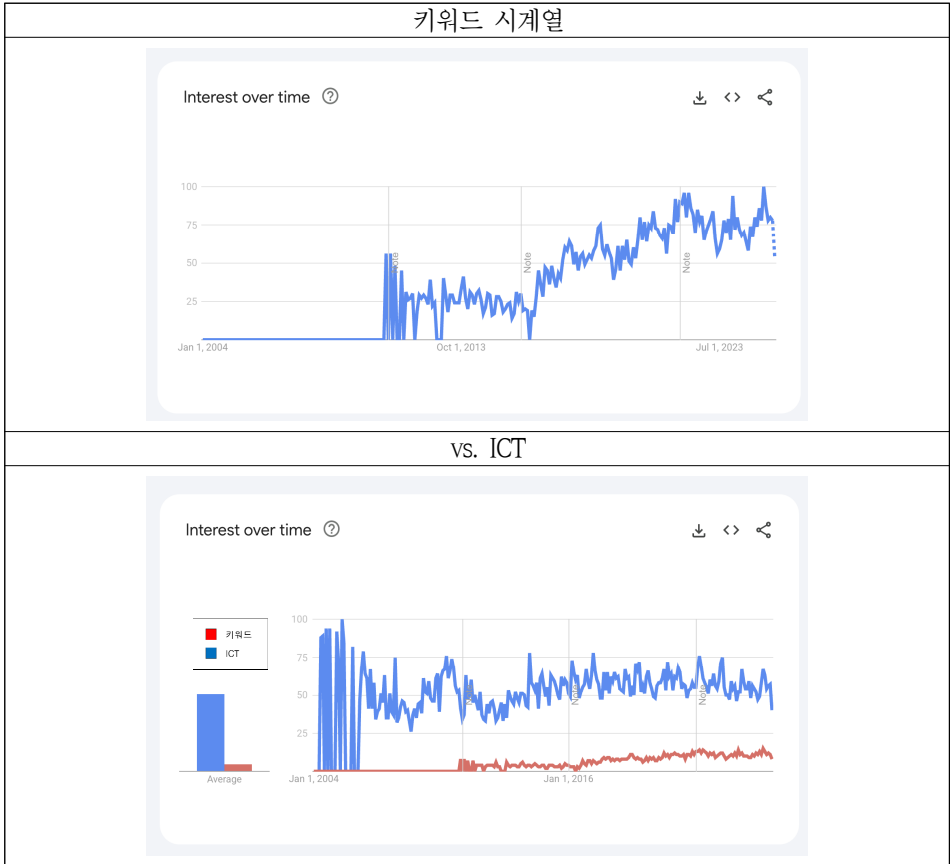
〈표 3-5〉 최근에 나타난 키워드 변화 양상

키워드	추세 지속 여부	키워드 유지 여부
6G	상승	유지
메타버스	하강	폐기
모빌리티	상승	유지
플랫폼 엔지니어링	하강	폐기
디지털 안전	상승	유지

자료: 연구진 작성

1) 키워드: 6G

[그림 3-11] '6G' 키워드 시계열 분석 결과

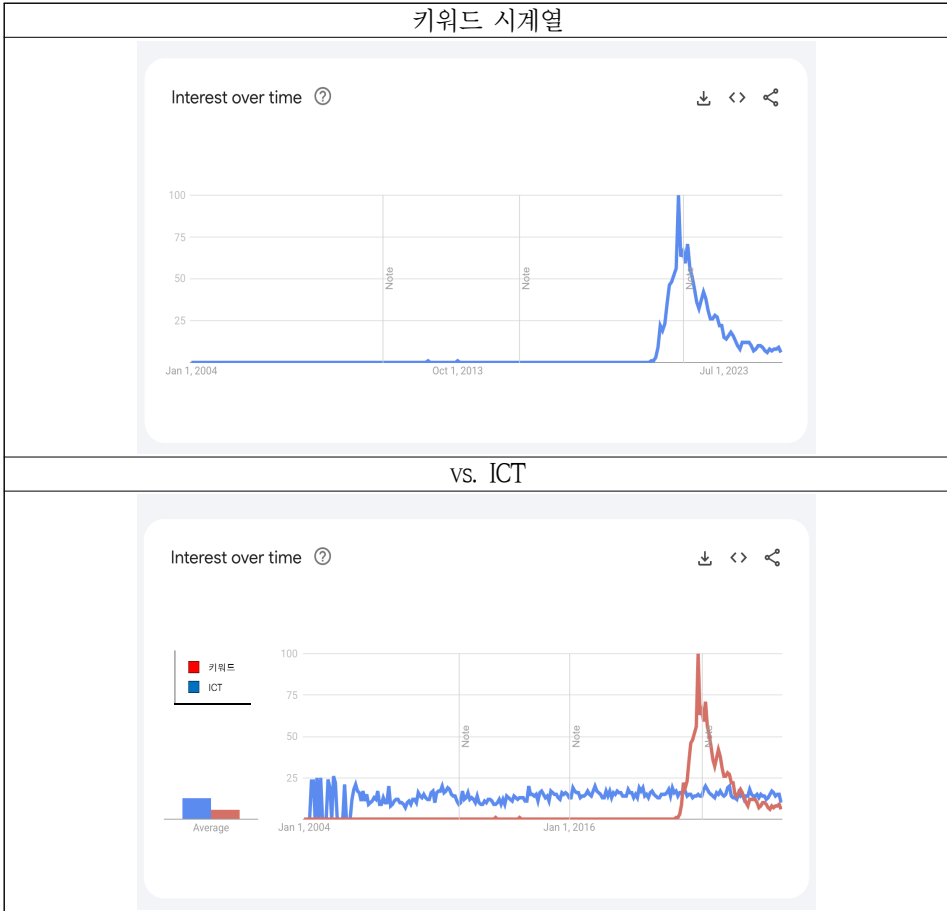


자료: 연구진 작성

- '6G' 키워드는 2023년 이후에도 지속적으로 상승해온 것으로 나타났음
 - － 이는 6G 관련 표준화 논의의 본격화, 국제적 기술 주도권 경쟁, 차세대 통신 인프라에 대한 투자 계획 발표 등과 같은 이슈들이 지속적으로 제기되고 있기 때문으로 해석됨
- 기준 키워드에 비해 검색량 자체는 낮은 편이나, 2020년대 이후 일정한 관심도를 유지하고 있다는 점에서 '6G'는 현재까지도 꾸준히 언급되고 있는 키워드로 분류할 수 있음

2) 키워드: 메타버스

[그림 3-12] '메타버스' 키워드 시계열 분석 결과



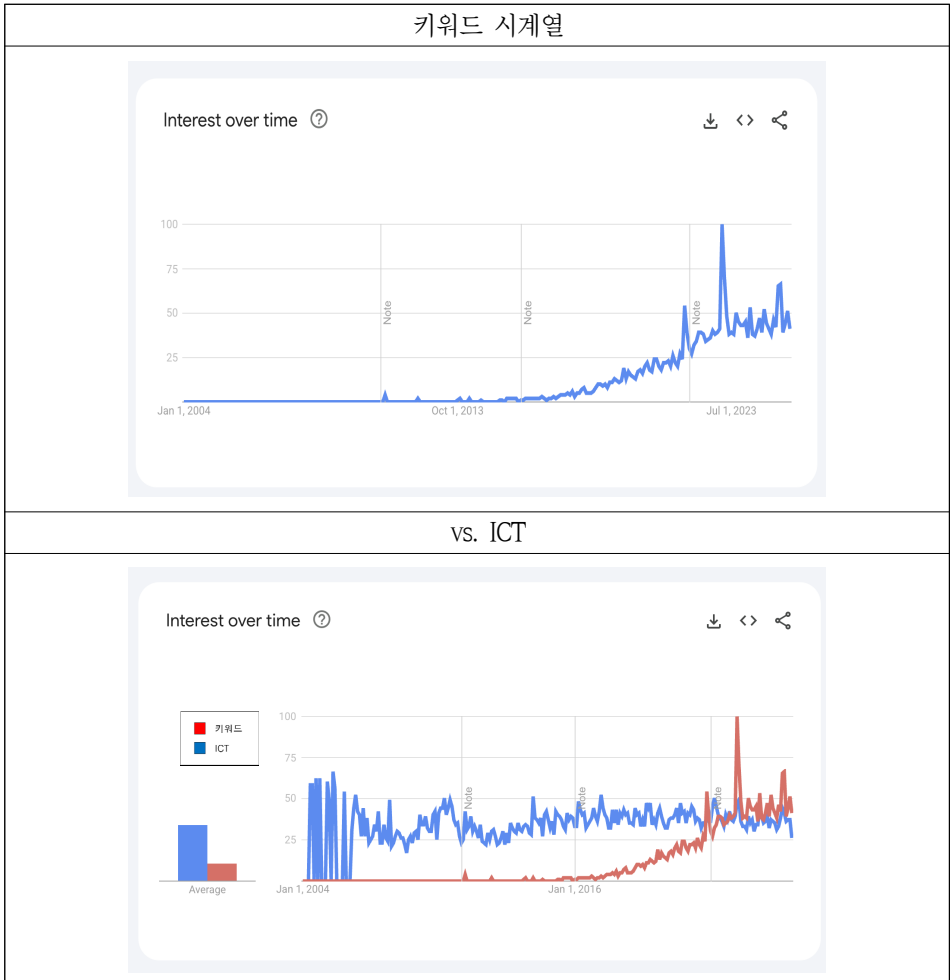
자료: 연구진 작성

- '메타버스' 키워드는 2021년을 기점으로 급격한 관심도 상승을 보였으나, 2022년 이후에는 관심도가 점차 하락하는 추세를 나타내고 있음
 - 이는 초기 기대와 달리 기술적 구현의 한계, 지속적인 이용 경험의 부재, 투자 대비 낮은 실효성 등으로 인해 대중의 기대에 부응하지 못한 결과로 해석됨
 - 과도하게 부풀려졌던 기대가 현실적인 제약과 맞물리며, COVID-19 이후 관심이 점차 식어간 것으로 판단됨

- 실제로 2021년 초에는 기준 키워드 대비 높은 검색량을 기록하였으나, 2023년 하반기에는 급격한 하락세를 보이며 기준 키워드와 유사한 수준으로 수렴
- 절대적인 검색량 자체는 여전히 일정 수준을 유지하고 있으나, 관심도 감소 폭이 크다는 점을 고려할 때 ‘메타버스’는 트렌드가 지난 키워드로 분류할 수 있음

3) 키워드: 모빌리티

[그림 3-13] ‘모빌리티’ 키워드 시계열 분석 결과

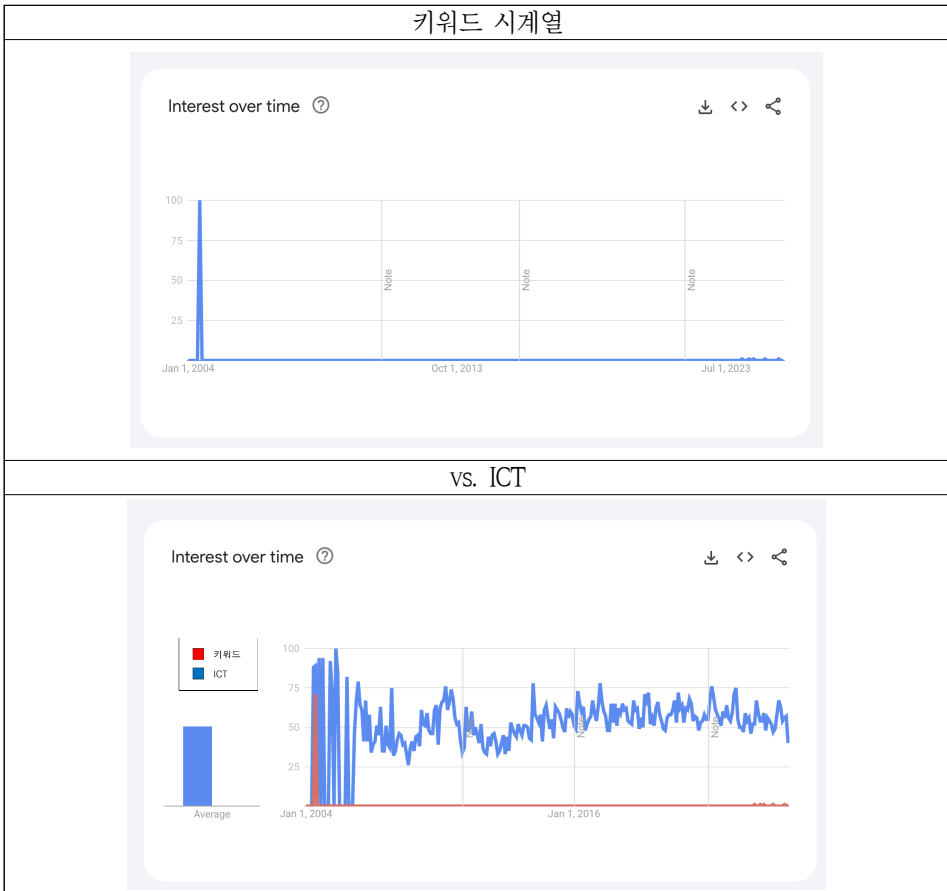


자료: 연구진 작성

- ‘모빌리티’ 키워드는 자율주행, 공유 차량, 전기차 등 차세대 교통 기술 전반에 대한 관심을 포괄하는 개념으로, 이러한 기술들이 미래 이동 수단의 혁신을 이끄는 핵심 분야로 주목받음에 따라 2020년대에 들어서며 관심도가 지속적으로 상승하는 추세를 보이고 있음
- 기준 키워드 대비 검색량 수준은 상대적으로도 다소 높으며, 안정적인 상승세가 유지되고 있다는 점에서 ‘모빌리티’는 현재까지도 꾸준히 언급되는 키워드로 분류할 수 있음

4) 키워드: 플랫폼 엔지니어링

[그림 3-14] ‘플랫폼 엔지니어링’ 키워드 시계열 분석 결과

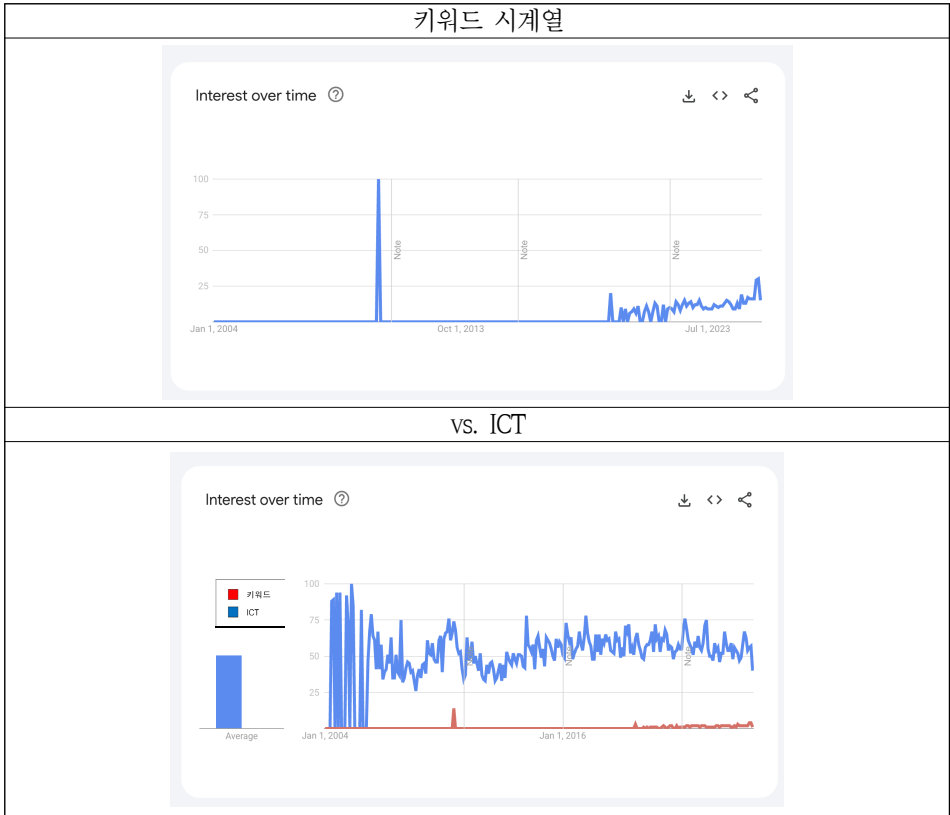


자료: 연구진 작성

- ‘플랫폼 엔지니어링’은 개발자들이 소프트웨어를 보다 효율적으로 개발·배포할 수 있도록 지원하는 내부 개발 플랫폼을 설계하고 구축하는 기술 및 조직적 접근 방식을 의미
 - 이는 인프라, 툴체인, 자동화 도구 등을 표준화하여 개발자의 생산성과 자율성을 동시에 제고하는 것을 목표로 함
- 다만, 실제 활용 과정에서는 ‘플랫폼 엔지니어링’이라는 용어보다 관련 개념이 다른 용어나 표현으로 대체되어 사용되는 경우가 많은 것으로 보임
 - 이로 인해 기존 키워드 대비 검색량은 낮은 수준에 머무르고 있으며, 최신 키워드로 분류하기에는 어렵다고 판단됨

5) 키워드: 디지털 안전

[그림 3-15] ‘디지털 안전’ 키워드 시계열 분석 결과



자료: 연구진 작성

- ‘디지털 안전(Digital Safety)’은 단순한 기술적 보안 조치를 넘어, 사용자가 온라인 환경에서 물리적·정신적·사회적으로 안전하게 활동할 수 있도록 보호하는 종합적인 개념을 의미
 - 디지털 안전은 익명성, 정보 비대칭, 확산성 등 디지털 환경의 특성으로 인해 발생하는 사이버 괴롭힘, 개인정보 침해, 온라인 범죄 등으로부터 사용자를 보호하기 위한 대응 개념으로 등장하였으며, 온라인 공간에서의 피해가 사회적 문제로 부각됨에 따라 이에 대한 제도적·기술적 보호의 필요성이 점차 확대되고 있음
- 기준 키워드 대비 검색량은 아직 낮은 수준이나, 관심도는 지속적으로 증가하는 추세를 보이고 있기에 ‘디지털 안전’은 꾸준히 언급되는 키워드로 분류할 수 있음

4. 최신 키워드 발굴

- 최신 ICT 키워드 발굴을 위해, 공신력 있는 국내 주요 기관의 보고서를 기반으로 2023년 이후 ICT 분야에서 새롭게 주목받고 있는 키워드를 분석하였음
 - 기간은 2020년 8월~ 2025년 7월을 고려하였으며, 이 과정에서 기존에 선정된 키워드들과의 중복 여부를 검토하여 의미가 유사하거나 포괄적으로 포함되는 항목에 대해서는 대표 키워드로 통합하는 방식으로 정리하였음
- 키워드 선정 시에는 해당 보고서들에 제시된 용어를 중심으로 키워드 후보군을 구성한 후, 시계열 분석을 통해 각 키워드의 관심도 추세를 확인하였음
 - 이를 바탕으로 ‘디지털 트윈’, ‘딥페이크’, ‘퀀텀 컴퓨팅’, ‘생성형 AI’, ‘드론’ 이상 5개 키워드를 최종 선정하였음
- 한편, 보고서들에서 공통적으로 4회 언급된 ‘사이버 보안’은 기존 키워드인 ‘보안’ 및 ‘디지털 안전’과 의미적으로 중복되는 것으로 판단되어 최종 후보군에서 제외

〈표 3-6〉 신규 키워드 후보 리스트: 출처

기관명	주제	키워드 목록
NIA 과학기술정보통신부	2025년 12대 디지털 트렌드	에이전트 AI
		AI 거버넌스
		반도체
		AI 로봇
		자율주행
		디지털 트윈
		디지털 창작
		보안
		포스트 5G
		엣지 컴퓨팅
		온디바이스 AI
		딥페이크
KISA 한국인터넷진흥원	동향분석	블록체인
		사이버 보안
		빅테크 규제
		생성형 AI
		플랫폼 수수료
IITP 정보통신기획평가원	ICT 동향정보	데이터 독점
		다큐먼트 AI
		AI 거버넌스
		AI 데이터 센터
		휴머노이드 로봇
		퀀텀 컴퓨팅
		드론
		사이버보안
		AI 저작권
		6G
삼성SDS	2024 ICT 전망 보고서	오픈소스
		사이버 보안
		생성형 AI
		기술 부채
	2025 DX 전망 보고서	IT 신뢰성
		고성능 스토리지
		AI 에이전트
		디지털 트랜스포메이션(DX)
		데브옵스(DevOps)
		AI 거버넌스
		데이터 부채
		딥페이크
		보안
		자동화
클라우드 마이그레이션		

자료: 연구진 작성

〈표 3-7〉 신규 키워드 후보 리스트: 등장횟수

순번	키워드명	등장횟수	Sources
1	생성형 AI	4	NIA, KISA, IITP, 삼성SDS
2	사이버 보안	4	NIA, KISA, IITP, 삼성SDS
3	딥페이크	2	NIA, 삼성SDS
4	로봇	2	NIA, IITP
5	자동화	1	삼성 SDS
6	퀀텀 컴퓨팅	1	IITP
7	블록체인	1	KISA
8	디지털 트윈	1	NIA
9	AI 저작권	1	IITP
10	6G	1	IITP
11	오픈소스	1	IITP
12	기술 부채	1	삼성 SDS
13	IT 신뢰성	1	삼성 SDS
14	고성능 스토리지	1	삼성 SDS
15	AI 에이전트	1	삼성 SDS
16	디지털 트랜스포메이션	1	삼성 SDS
17	데이터 부채	1	삼성 SDS
18	드론	1	IITP
19	AI 데이터 센터	1	IITP
20	데이터 독점	1	KISA
21	플랫폼 수수료	1	KISA
22	빅테크 규제	1	KISA
23	엣지 컴퓨팅	1	NIA
24	포스트 5G	1	NIA
25	디지털 창작	1	NIA
26	자율주행	1	NIA
27	반도체	1	NIA
28	클라우드 마이그레이션	1	삼성SDS

자료: 연구진 작성

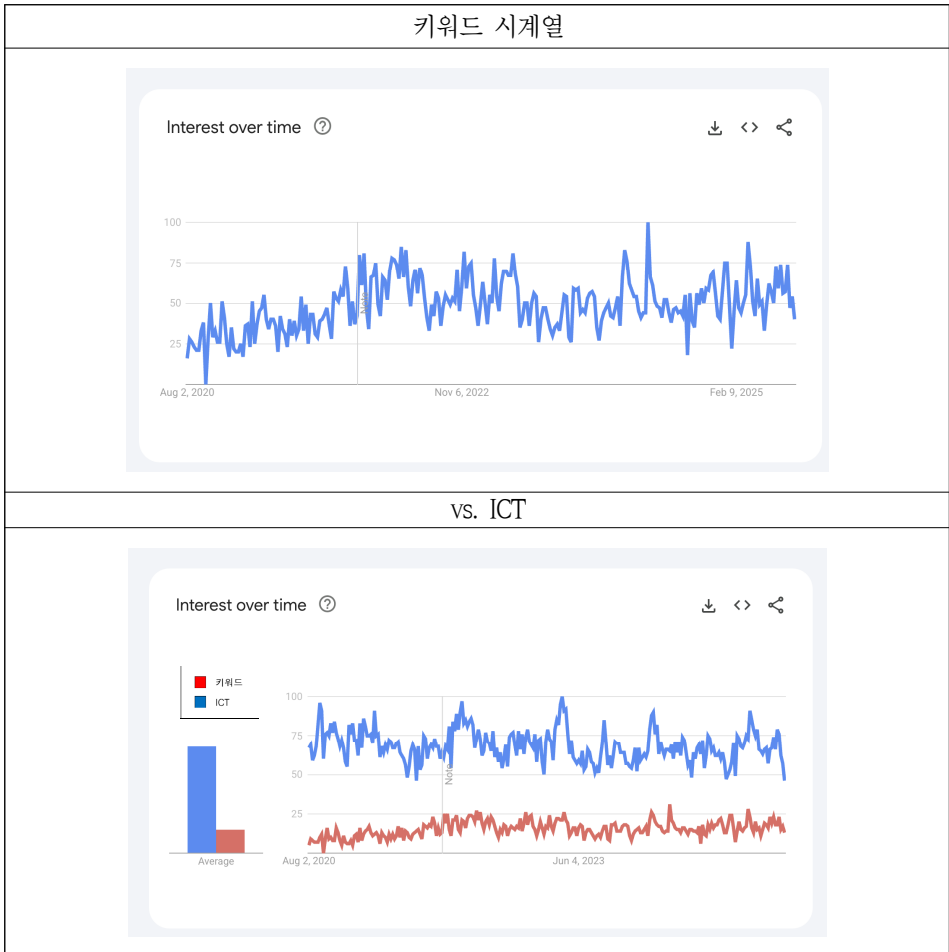
〈표 3-8〉 최종 선정 키워드 분류 결과

선정 키워드	추세 지속 여부
디지털 트윈	상승
딥페이크	상승
퀀텀 컴퓨팅	상승
생성형 AI	상승
드론	상승

자료: 연구진 작성

1) 키워드: 디지털 트윈

[그림 3-16] '디지털 트윈' 키워드 시계열 분석 결과



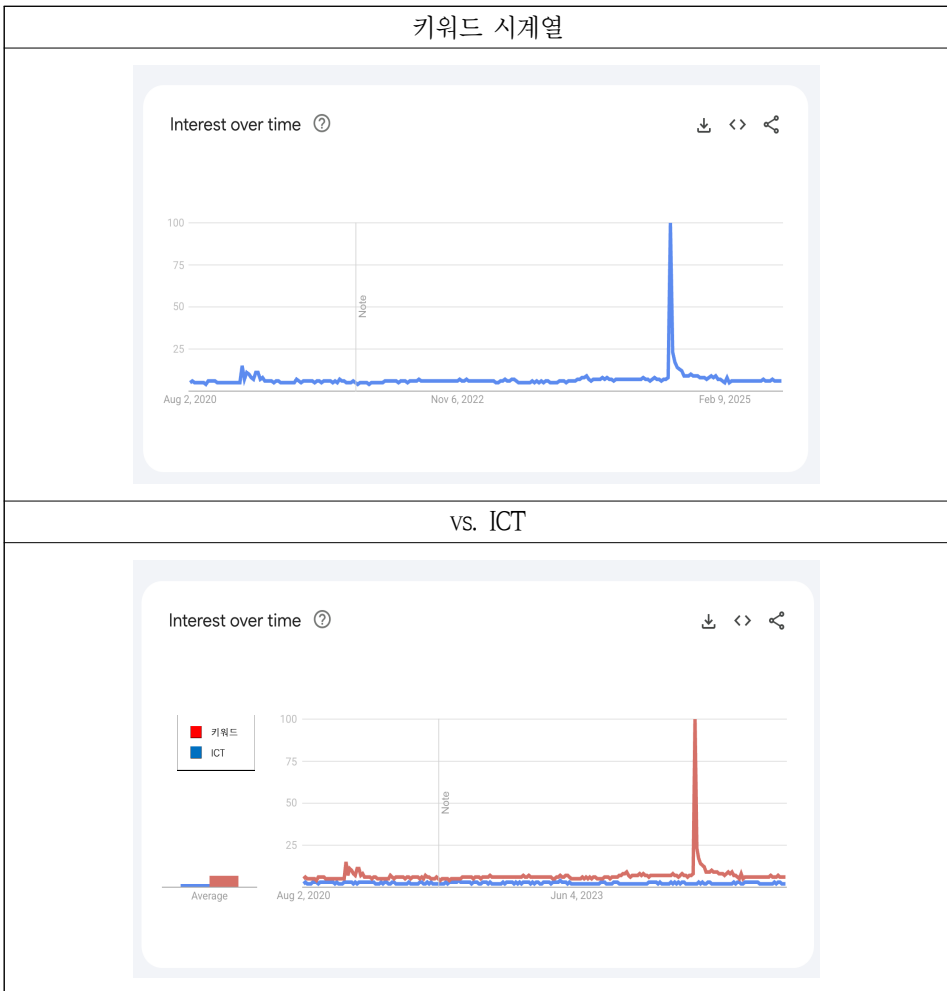
자료: 연구진 작성

- '디지털 트윈'은 현실에 존재하는 물리적 대상이나 시스템을 실시간 데이터 기반으로 가상 세계에 동일하게 복제한 디지털 모델을 의미
- 센서, IoT, 인공지능 등의 기술이 결합되어 물리 자산의 현재 상태를 모니터링하고, 시뮬레이션 및 예측 분석을 통해 유지보수의 최적화와 운영 효율성 향상 등 다양한 의사결정을 지원할 수 있음

- 디지털 트윈은 스마트 시티, 스마트 팩토리 등 여러 분야에서 핵심 요소로 인식되고 있어 이에 대한 관심이 점차 확대되고 있음
- o 해당 키워드에 대한 관심도가 지속적으로 상승하고 일정 수준 이상을 유지하고 있는 것으로 나타나, 본 키워드는 최근에 나타난 키워드로 분류할 수 있음

2) 키워드: 딥페이크

[그림 3-17] '딥페이크' 키워드 시계열 분석 결과

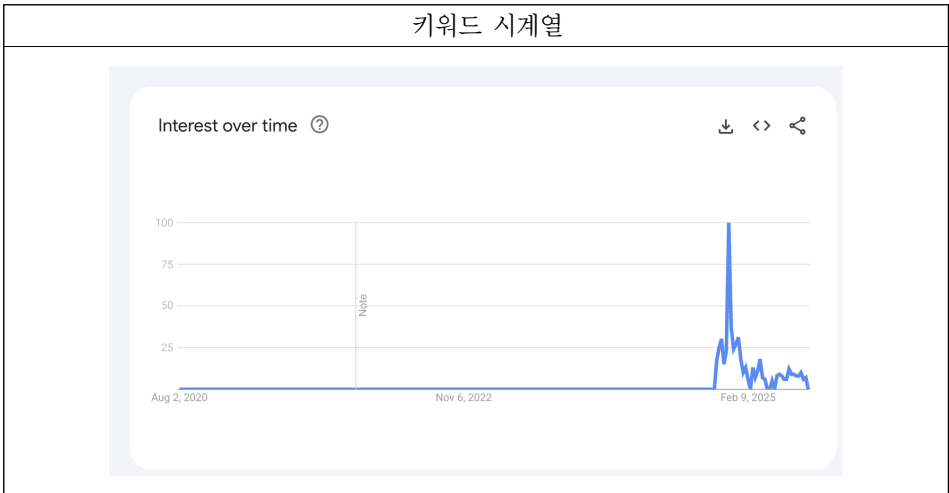


자료: 연구진 작성

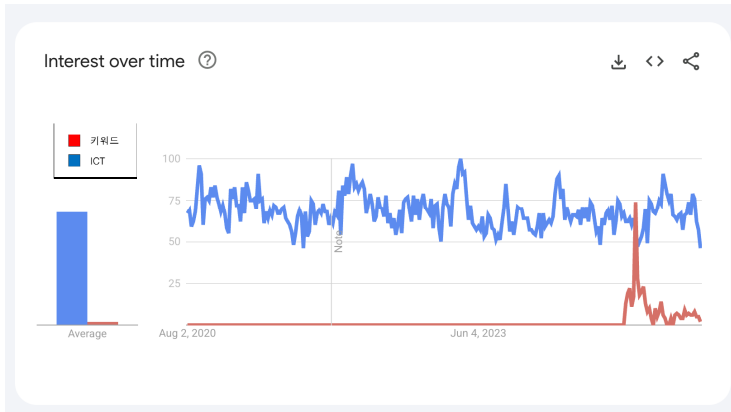
- ‘딥페이크’ 키워드는 최근 지속적으로 언급되어 온 키워드로, 해당 기술이 초래한 다양한 사회적 논란과 윤리적 쟁점이 관심을 꾸준히 유지시키는 주요 요인으로 작용하였음
 - 특히 관련 이슈나 사건이 발생했을 당시, 단기적으로 급격한 검색량 증가가 나타나는 특징을 보이고 있음
- 전체적인 추세는 2020년에 비해 검색량이 소폭 상승하며 일정하게 유지되고 있는 모습이 확인되었음
- 또한, 기준 키워드와 비교하였을 때도 지속적으로 높은 검색량을 기록하였기 때문에 이러한 점을 고려하여 ‘딥페이크’는 본 분석에서 최근에 나타난 키워드로 분류하였음

3) 키워드: 퀀텀 컴퓨팅

[그림 3-18] ‘퀀텀 컴퓨팅’ 키워드 시계열 분석 결과



vs. ICT



자료: 연구진 작성

- ‘퀀텀 컴퓨팅’은 양자역학의 특성을 활용하여 병렬적이고 고속의 계산을 가능하게 하는 차세대 컴퓨팅 패러다임
 - 현재는 연구 및 초기 실험 단계에 머물러 있으나 암호해독, 분자 시뮬레이션, 최적화 등 여러 분야에서 기존 고전 컴퓨팅을 능가할 수 있는 잠재력을 지니고 있어 높은 관심을 받고 있음
- 2024년 하반기에는 관련 기술 발표와 정책적 관심에 힘입어 검색량이 일시적으로 급증하였으나, 이후 상용화의 어려움과 현실적인 제약으로 인해 관심도가 빠르게 하락하는 양상을 보였음
- 그럼에도 불구하고, 전체적인 추세를 살펴보면 일정 수준 이상의 관심도가 지속되고 있으며, 활발한 연구가 진행 중인 분야인 점을 고려하여, 본 분석에서는 ‘퀀텀 컴퓨팅’을 최근에 나타난 키워드로 분류하였음

4) 키워드: 생성형 AI

[그림 3-19] '생성형 AI' 키워드 시계열 분석 결과

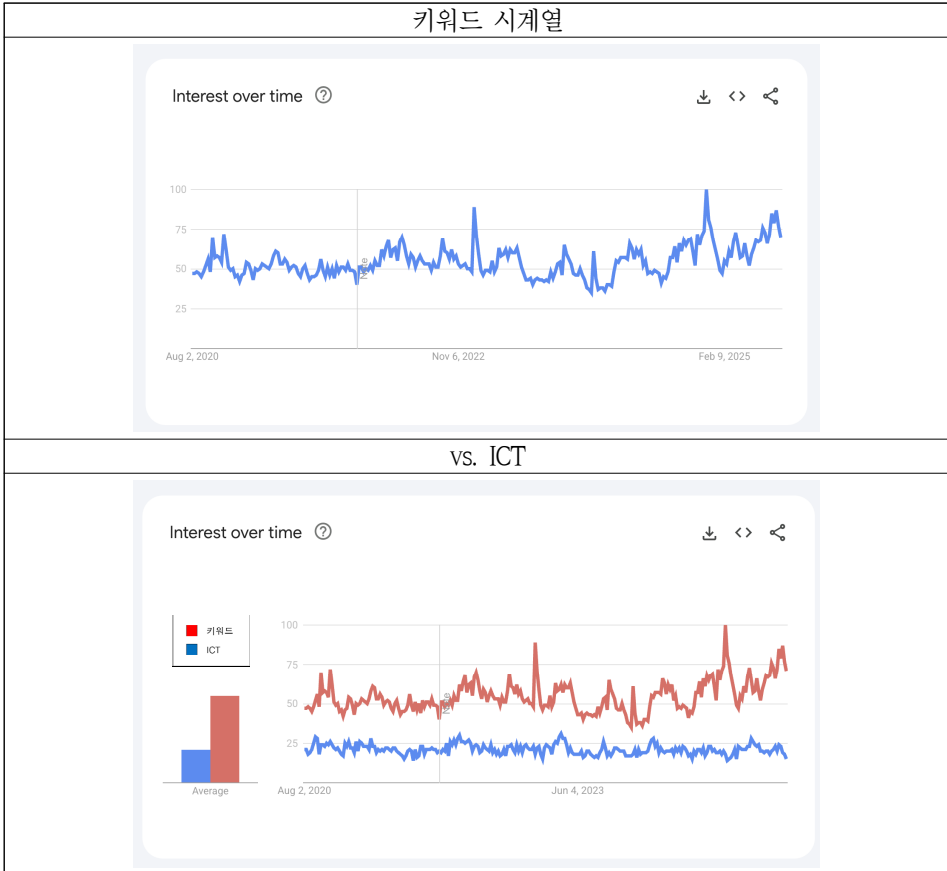


자료: 연구진 작성

- '생성형 AI' 키워드는 ChatGPT의 등장 이후 대중의 급격한 관심을 받기 시작한 것으로 나타났음
- 트렌드 분석 결과, 관심도가 지속적으로 상승하는 추세를 보이고 있으며, 기준 키워드와 유사한 수준의 검색량을 꾸준히 유지하고 있는 것으로 확인됨
 - 특히 '생성형 AI'라는 일반적 표현보다는, 실제로 대중이 활용하는 GPT, Gemini 등 구체적인 서비스 기반의 대규모 언어 모델(LLM) 키워드를 중심으로 보다 뚜렷한 관심도 증가와 높은 검색량이 나타나는 경향이 확인되었음

5) 키워드: 드론

[그림 3-20] '드론' 키워드 시계열 분석 결과



자료: 연구진 작성

- '드론' 키워드는 기술 발전과 다양한 분야에서의 응용 가능성에 대한 기대감으로 인해 과거부터 꾸준한 관심을 받아온 주제
 - 특히 최근에는 여러 국가 간 분쟁 및 전쟁 상황에서 드론이 정찰, 공격, 방어 등 전략적 무기로 활발히 활용되면서, 이에 대한 사회적 관심과 주목도가 한층 더 증가한 것으로 나타났음
- 특히, 기존 키워드보다 높은 검색량을 꾸준히 유지하고 있으며, 이러한 점을 종합적으로 고려할 때 '드론'은 최근에 나타난 키워드로 분류할 수 있음

5. 최종 선정 키워드

- 분석을 통해 기존 키워드 중 일부는 산업 대표성이나 관심도 측면에서 한계를 보여 제외 또는 보완 대상으로 판단되었으며, ‘자동화’, ‘보안’, ‘플랫폼’, ‘로봇’ 등은 여전히 유의미한 핵심 개념으로 유지되었음
- ‘생성형 AI’, ‘디지털 트윈’, ‘딥페이크’, ‘퀀텀 컴퓨팅’, ‘드론’ 등은 기술적 파급력과 정책적 중요성을 동시에 갖춘 키워드로, 향후 산업 및 정책 설정 과정에서 핵심적 역할을 할 가능성이 높음
 - 특히 ‘생성형 AI’와 ‘디지털 트윈’은 정부 주도의 산업 육성 및 규제 설계에서 우선적으로 고려되어야 할 영역으로 판단됨
- 이러한 점을 종합하여 선정한 ICT 관련 키워드는 다음과 같음

〈표 3-9〉 선정된 키워드 목록

순번	키워드	선정 여부
기존 키워드		
1	유비쿼터스	X
2	자동화	O
3	소프트웨어	X
4	LTE	X
5	어플리케이션	X
6	네트워크	O
7	보안	O
8	로봇	O
9	우주	O
10	플랫폼	O
11	6G	O
12	메타버스	X
13	모빌리티	O
14	플랫폼 엔지니어링	X
15	디지털 안전	O
최신 키워드		
1	디지털 트윈	O

순번	키워드	선정 여부
2	딥페이크	○
3	퀀텀 컴퓨팅	○
4	생성형 AI	○
5	드론	○

자료: 연구진 작성

- 기존에 트렌드가 지난 키워드로 분류된 ‘유비쿼터스’, ‘소프트웨어’, ‘LTE’, ‘어플리케이션’은 최신 ICT 트렌드를 반영하기에는 적절하지 않다고 판단되어 제외하였음
- 다만 ‘자동화’는 인공지능과 스마트팩토리 등의 확산으로 다시 관심을 받고 있는 것을 확인
- 2023년 기준 최근에 나타난 키워드로 분류된 ‘메타버스’와 ‘플랫폼 엔지니어링’의 경우 관심도가 뚜렷하게 하락하고 있기에 키워드 목록에서 제외하였음
- 2025년 기준으로 최신 트렌드를 반영하는 ‘디지털 트윈’, ‘딥페이크’, ‘퀀텀 컴퓨팅’, ‘생성형 AI’, ‘드론’을 신규 키워드로 포함
- 분석을 통해 기존 키워드 중 일부는 산업 대표성이나 관심도 측면에서 한계를 보여 제외 또는 보완 대상으로 판단되었으며, 신규 키워드를 고려하여 연구진의 논의를 거쳐 최종 20개의 키워드를 도출하였음

〈표 3-10〉 최종 키워드 목록

유지 키워드	제외 키워드	신규 키워드
ICT 산업	ICT 메타버스	디지털 트윈
ICT 기업	ICT 플랫폼 엔지니어링	퀀텀 컴퓨팅
ICT 수출	-	생성형 AI
ICT 품목		드론
IT 산업		
IT 기업		
IT 수출		
IT 품목		
ICT 우주		
ICT 플랫폼		

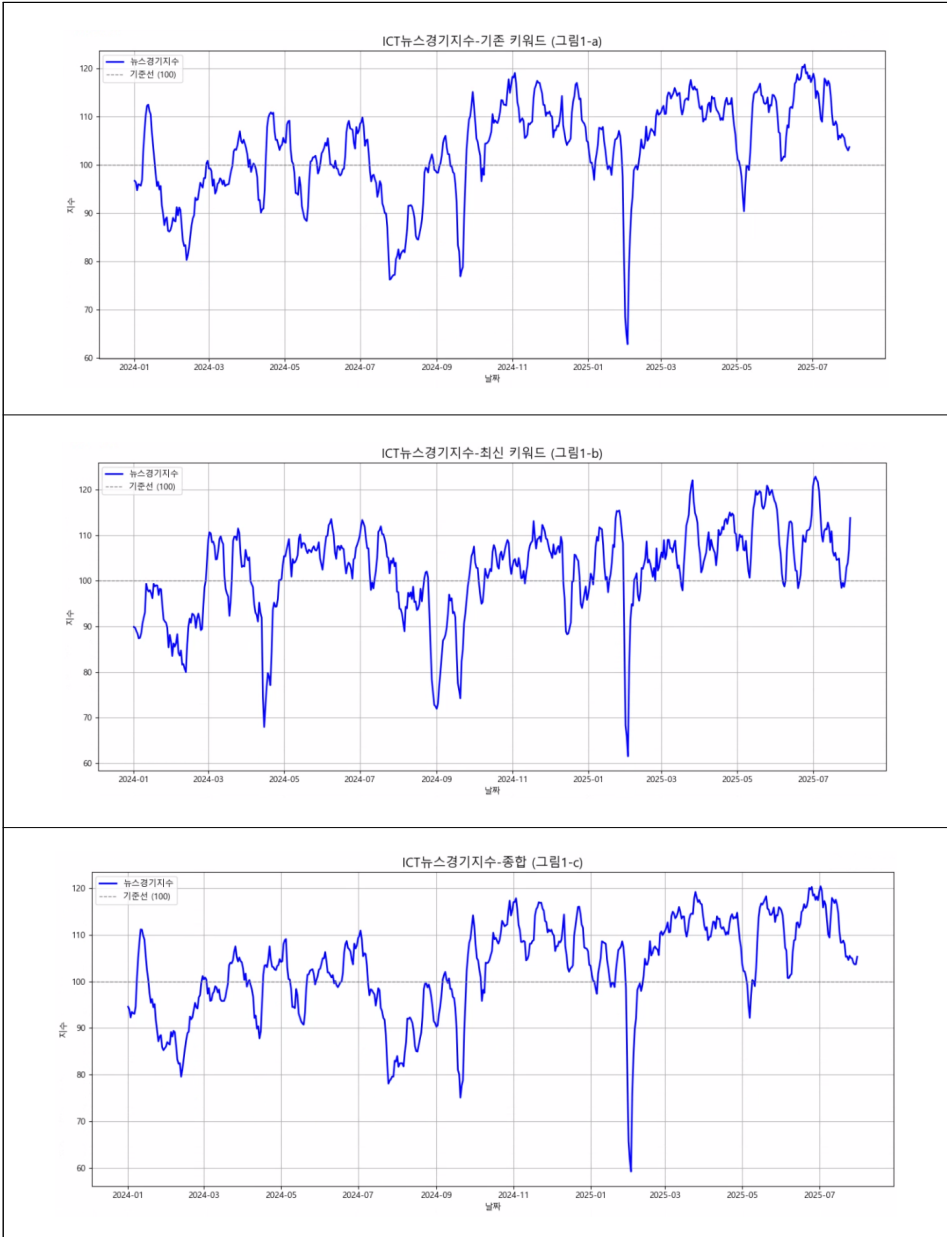
유지 키워드	제외 키워드	신규 키워드
ICT 모빌리티		
ICT 디지털 안전		
ICT 로봇		
ICT 보안		
ICT 6G		
ICT 네트워크		

자료: 연구진 작성

6. 신규 키워드 기반 ICT 뉴스경기지수와 BSI 간 비교

- 본 절에서는 5개의 신규 키워드를 반영한 ICT 뉴스경기지수와 기업경기지수(BSI) 간 상관관계 및 시차 구조를 실증적으로 분석함으로써, ICT 뉴스경기지수가 산업 경기 변화를 적시에 반영하거나 선행할 수 있는지를 검토하고자 함
- ICT 뉴스경기지수 구축을 위한 뉴스 기사 수집 기간은 2024년 1월부터 2025년 7월 까지로 설정하였으며, 동일 기간 동안 총 1,334,218건의 뉴스 기사가 수집되어 분석에 활용
- ICT 뉴스경기지수를 계산하는 구체적인 방법은 2절에 소개되어 있으며, 본 절에서는 분석 결과를 중심으로 소개하고자 함
- 도출한 일별 ICT 뉴스경기지수 결과를 소개하면 다음과 같음
 - 먼저, 2023년 연구에서 도출된 9개 키워드를 기반으로 ICT 뉴스경기지수를 산출하여 “기존 키워드”라는 이름으로 (그림 1-a)에 제시하였음
 - 이어, 본 연구에서 새롭게 도출한 5개의 최신 키워드를 활용하여 산출한 ICT 뉴스경기지수 시계열 그래프를 “최신 키워드”로 (그림 1-b)에 제시하였음
 - 마지막으로, 종합 그래프인 (그림 1-c)에서는 ICT를 대표하는 8개 키워드, 2023년 연구에서 도출된 9개 키워드, 그리고 이번 연구에서 도출한 5개 최신 키워드를 모두 포함하여 산출한 ICT 뉴스경기지수를 제시하였음

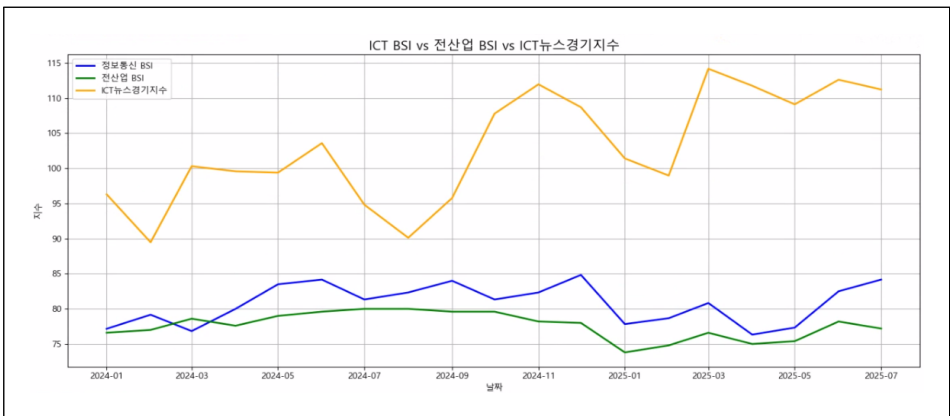
[그림 3-21] ICT 뉴스경기지수 (기존 키워드, 최신 키워드, 종합)



자료: 연구진 작성

- 분석 결과, 기존 키워드와 신규 키워드 모두 큰 틀에서의 추세는 유사하나 세부적인 흐름에서는 차이가 존재
 - 상대적으로 최신 키워드로 산출한 ICT 뉴스경기지수의 변동이 더 잦은 것으로 나타나는데, 이는 해당 키워드들에 대한 관심이 높은 것을 반영하고 있을 가능성이 있음
 - 다만 긍정/부정의 기준이 되는 100 이상/이하 구간 관점으로 보면 유사한 모습을 보인다고 판단됨
- 신규 키워드 반영을 통한 변화를 구체적으로 살펴보기 위해, 산출된 종합 일별 지수를 바탕으로 각 월의 평균값을 계산함으로써 월별 ICT 뉴스경기지수를 도출하였음
 - 이를 통해 ICT 산업 전반에 대한 시장의 정성적 인식 변화를 시간 단위로 파악할 수 있으며, 전통적 경기지표와의 비교 및 보완지표로서의 활용 가능성을 제시
- 월별 지수와 ICT BSI, 전산업 BSI를 함께 비교한 그림은 다음과 같음

[그림 3-22] ICT뉴스경기지수와 BSI 시각화



자료: 연구진 작성

- ICT 뉴스경기지수의 유효성을 검증하기 위해, 월별 ICT 뉴스경기지수와 기업경기실사지수(BSI) 간의 상관관계 분석을 수행하였음
 - 비교 대상은 전산업 BSI와 ICT 산업 BSI이며, 두 지표 모두 업황 전망치를 기준으로 분석

- ICT 산업 BSI는 산업별 세부 지수를 기반으로, 전자·영상·통신장비 제조업과 정보통신업에 해당하는 BSI의 단순 평균으로 정의
- 2024년 데이터를 활용하여 상관관계 분석을 수행하였으며, 상관계수 분석 결과는 다음과 같음

〈표 3-11〉 2024년 시차별 ICT와 BSI간 상관계수

선행시차 (개월)	ICT BSI 상관계수	전산업 BSI 상관계수
1	0.455	-0.270
2	0.494	0.134
3	0.126	0.515
4	-0.776	0.587

자료: 연구진 작성

- 전반적으로 살펴보면, 2024년 정보통신 BSI와 ICT 뉴스경기지수 간 상관계수는 선행 1~2개월 시점에서 가장 높게 나타났으며(0.4547, 0.4940), 이는 ICT 뉴스경기 지수가 ICT 산업의 경기 전망에 대해 약 1~2개월 선행하는 경향이 있음을 시사
- 반면, 전산업 BSI와의 상관계수는 선행 3~4개월 시점에서 더 높게 나타났는데(0.5140, 0.5878) 분석 결과 ICT BSI가 전산업 BSI를 1~2개월 선행하는 경향이 있는 것으로 나타났음(0.7780, 0.3564)
 - 이는 ICT 뉴스경기지수가 ICT 산업뿐만 아니라 산업 전반의 경기 심리를 선행적으로 포착할 수 있는 잠재적 지표임을 시사
- 2023년 기준 키워드 분석에서도 ICT 뉴스경기지수가 BSI 지수를 1~2개월 선행하는 것으로 분석된 바 있음
 - ※ 기존 결과: ICT BSI 1, 2개월 선행 상관계수는 각각 0.395와 0.348, 전산업 BSI 1, 2개월 선행 상관계수는 각각 0.513, 0.476임
- 따라서, 이번 키워드 최신화 이후에도 이러한 선행성이 유지되고 있는 것으로 나타나 동 지수가 여전히 유효한 선행 지표로서의 역할을 수행하고 있음을 확인
- 신규 키워드 추가를 통해 새로운 ICT 트렌드를 시의성 있게 반영할 수 있을 뿐만 아니라 ICT 뉴스경기지수의 유효성 또한 유지할 수 있을 것으로 판단

제 4 장 ICT 키워드 도출 자동화 체계 구축

제 1 절 개 요

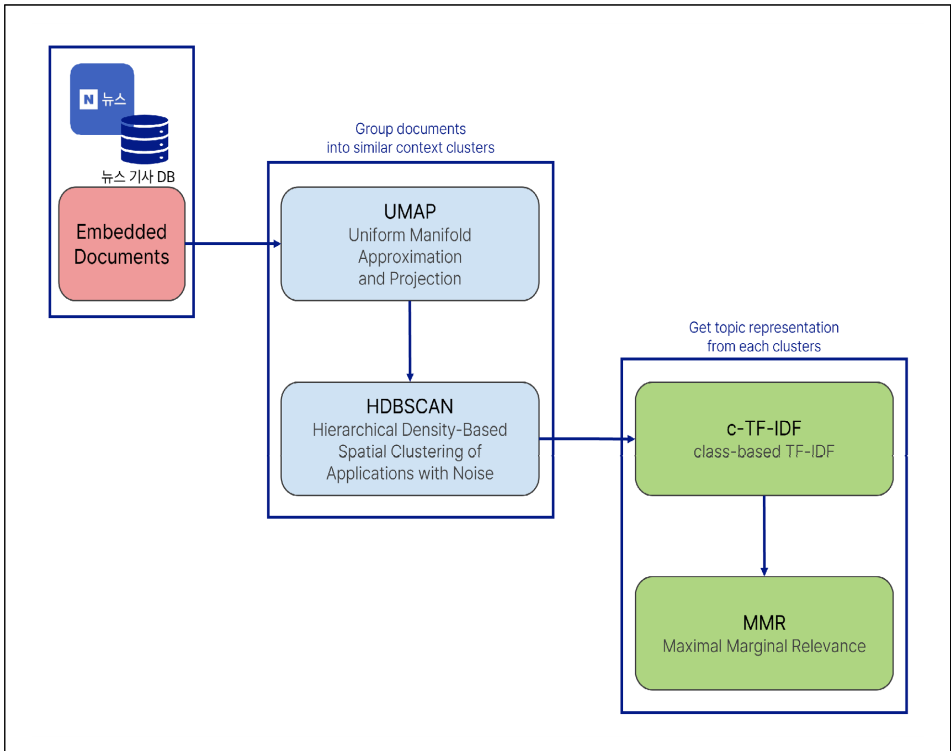
- 본 장에서는 뉴스 기사 수집 - 딥러닝 기반 임베딩/군집화 기법 분석을 통해 ICT 분야의 주요 키워드 발굴을 시의성 있게 발굴할 수 있는 프레임워크 개발 방안을 제시
 - 이전에는 ICT 키워드 도출을 위해 국내외 주요 기관 자료 및 Google Trends를 활용하여 키워드를 도출하고 분석하는 방법을 사용하였음
 - 이러한 방법은 공신력 있는 자료를 바탕으로 신뢰성 있는 결과를 도출할 수 있다는 장점이 있으나 시의성이 낮은 문제가 있으며, 아직 주목을 받지 못하였으나 향후 핵심 트렌드로 부상할 수 있는 주제들을 잡기가 어렵다는 한계점이 있음
 - ※ 예를 들어, 2025년 Gartner의 주요 기술 트렌드 보고서에서는 기술 민첩성, 기술 플랫폼 등이 유망 기술로 지속적으로 언급되었으나, 국내 주요 언론의 뉴스 기사에서는 이들 키워드가 거의 사용되지 않거나, 매우 제한된 전문 기술 기사에서만 등장하였음(Gartner, 2025)
 - ※ 반면, '생성형 AI'는 2022년 하반기부터 뉴스와 온라인 커뮤니티 등에서 폭발적인 관심을 끌며 급부상했지만, McKinsey나 Deloitte의 2022년 중반 이전 보고서에서는 이를 별도 기술로 분리하여 조명하지 않았던 사례도 존재 (McKinsey Tech Trends, 2022)
 - ICT 산업 및 기술 환경이 빠르게 변화하는 상황에서, 민간의 관심 흐름을 적시에 반영하는 키워드를 도출하기 위해서는 뉴스 데이터를 직접 분석하고 키워드를 상시적으로 점검할 수 있는 체계가 필요
 - 특히 기사 본문을 임베딩/군집화하여 뉴스에서 자생적으로 형성되는 주제 구조를 파악하고, 이를 기반으로 키워드 후보군을 자동으로 도출할 수 있다. 이를 통해

- 민간의 관심 흐름을 보다 잘 반영하는 키워드 체계를 구축할 수 있을 것으로 기대
- 이러한 관점에서 본 연구는 뉴스 데이터에 기반한 딥러닝/LLM 기반 키워드 자동화 프레임워크를 구축하고 ICT 트렌드를 반영한 키워드를 안정적으로 자동 산출 및 관리할 수 있는 시스템을 마련하는 것을 목적으로 함
 - 네이버 뉴스 IT 섹션을 대상으로 기사 수집-본문 정제-임베딩-군집화-키워드 추출에 이르는 전 과정을 하나의 모듈형 파이프라인으로 설계/구현하여, 분석 기간의 키워드 후보군이 자동 산출되는 환경을 구축
 - BERTopic⁷³⁾으로 도출된 군집별 후보 키워드를 LLM API와 연계하여 유사 및 중복 키워드를 상위 개념의 ICT 키워드로 재표현하고, 이를 통해 정책 및 산업 분석에 활용하기 적합한 형태의 최종 키워드 목록을 자동으로 생성
 - 재표현된 ICT 키워드별로 월별 문서 수를 기반으로 관심도 시계열을 구축하고, 트렌드의 변화를 살펴봄
 - 본 연구에서 전통적 방법론(예: 단순 빈도 기반 군집화 및 키워드 추출 방식) 대신 딥러닝 기반 접근을 통해 키워드 추출 프레임워크를 구축하고자 하는 이유는 다음과 같음
 - 전통적 방법론은 단어 빈도 공간에서 거리/확률을 사용하므로 공유 일반어(예: 시장, 성장, 투자) 영향으로 군집의 경계가 흐려지기 쉽고, 초기/희소 키워드는 기존 군집으로 흡수/희석되는 경향이 있음
 - 반면 딥러닝 방식은 [그림 4-1]과 같이 문서를 의미 임베딩으로 표현한 뒤, UMAP(Uniform Manifold Approximation and Projection)+HDBSCAN(Hierarchical Density-Based Spatial Clustering of Applications with Noise, 밀도 기반, 군집 수 미지정)으로 군집화하므로, 유사한 의미를 가지는 기사들은 임베딩 공간에서 서로 가깝게 위치하도록 설계됨

73) BERTopic은 본 연구에서 사용한 딥러닝 방법론으로, 최신 자연어처리(NLP) 기법을 통해 문서 집합에서 주제를 자동으로 추출하고, 각 주제를 구성하는 대표 키워드를 도출하는 토픽 모델링 기법이다. 특히 의미 기반 클러스터링과 class-based TF-IDF 방식을 결합하여 해석력 높은 주제 분류와 키워드 추출이 가능하다는 점에서 기존의 전통적 토픽 모델링보다 우수한 성능을 보인다.

- 이를 통해, ICT와 거시경제처럼 의미가 다른 집단에 공유 일반어가 존재해도 자연스럽게 분리할 수 있으며, 응집력 있는 소수 문서도 별도 군집으로 조기 분리 가능하고, 주제성이 약한 문서는 노이즈 (Noise)로 분리되어 왜곡을 줄일 수 있음
- 더불어 전통적 방법론은 빈도 축적 기반임으로, 등장 초기에는 임계치 미만으로 포착이 지연되기 쉬움
- 반면 딥러닝 방식은 소수 기사라도 의미적으로 응집되어 있으면 별도 군집으로 띄고, c-TF-IDF가 해당 군집의 차별 단어를 즉시 상위로 끌어올려 조기 신호 감지에 유리

[그림 4-1] BERTopic 구조

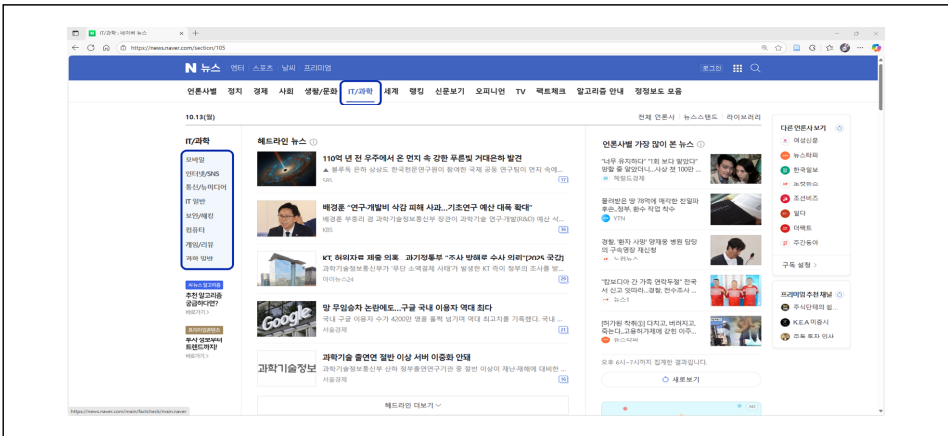


자료: 연구진 작성

제 2 절 국내 ICT 관련 뉴스 기반 키워드 자동화 체계 구축 방법론

- 본 연구는 네이버 뉴스 'IT/과학' 섹션을 활용하여 국내 ICT 관련 뉴스 기반 키워드 자동화 프레임워크를 구축하는 것을 목적으로 수행되었으며, 다음과 같은 범위 내에서 연구를 진행하였음
 - 네이버 뉴스의 'IT/과학' 섹션에는 [그림 4-2]와 같이 8개의 서브 카테고리 존재함을 확인할 수 있는데, BERTopic 기반 분석의 계산 비용과 토픽 탐지 성능을 고려하여 전체 'IT/과학' 섹션의 기사를 한 번에 처리하는 대신 서브 카테고리별로 기사를 분리해 수집
 - 이는 제한된 컴퓨팅 자원 하에서 처리 시간을 단축하는 동시에, 아직 문서 수는 적지만 의미 있는 신규 키워드를 개별 토픽으로 분리 및 탐지하기 위한 조치임
 - 또한, 주로 개별 게임·제품 리뷰에 치우쳐 관련성이 상대적으로 낮은 '게임/리뷰' 카테고리는 분석 대상에서 제외하였음
 - 분석 기간은 2024년 9월부터 2025년 8월까지 1년간 수집된 기사로 설정하여, 동일 기준의 데이터에 대해 BERTopic 기반 자동화 절차를 적용하였음

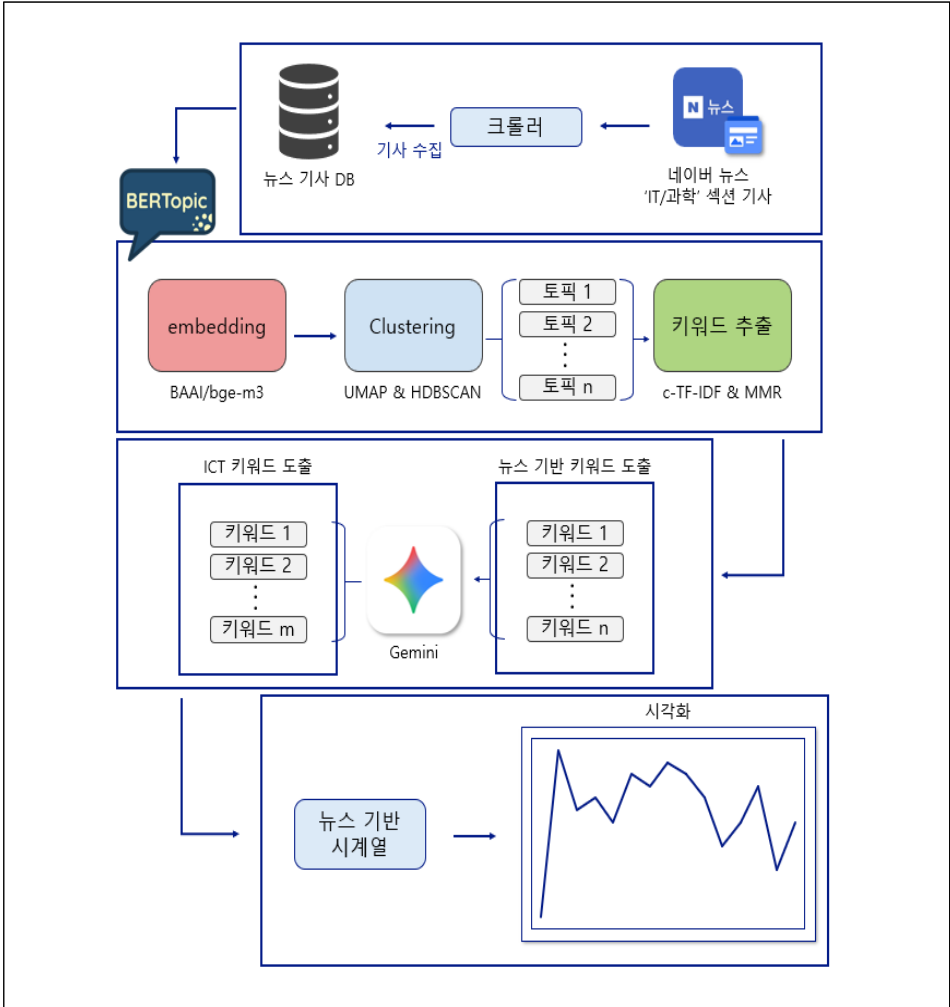
[그림 4-2] 네이버 뉴스 IT/과학 섹션 내 서브 카테고리(모바일, 인터넷/SNS, 통신/뉴미디어, IT일반, 보안/해킹, 컴퓨터, 게임/리뷰, 과학 일반)



출처: 네이버 뉴스, <https://news.naver.com/>

- 본 연구에서 구축한 뉴스 기반 키워드 도출 자동화 프레임워크는 [그림 4-3]과 같이 크게 네 개의 계층으로 구성

[그림 4-3] 뉴스 기반 키워드 도출 자동화 프레임워크



자료: 연구진 작성

1. 크롤링 자동화

- 먼저 크롤링 자동화의 경우, 일 단위 기사 수집을 수행하였으며 연구 수행을 위해 2024년 9월부터 2025년 8월까지의 IT/과학 카테고리의 기사 전체를 한 번에 수집하였음
 - 수집 후 분석 편의성과 키워드 산출 단위를 고려하여 월별 데이터셋으로 재구성하였음
 - 월 단위로 분할한 이유는 이후 키워드 선정 과정에서 각 월을 기준으로 1차 후보 키워드를 선별하기 위함임
 - 이러한 절차를 통해 총 333,065건의 기사 데이터가 구축되었으며, 이를 기반으로 후속 임베딩/군집화/키워드 추출 등의 모델 개발을 수행하였음

2. 뉴스기반 키워드 추출

- 이 단계에서는 BERTopic 프레임워크를 중심으로 기사 전처리, 기사 수준 임베딩 생성, 토픽 군집화, 토픽별 키워드 산출, 재랭킹 및 대표어 선정의 단계 절차를 순차적으로 수행하였음
- 뉴스 기반 키워드 추출은 전체 문서를 한 번에 입력하는 대신, 월 단위(12개월)와 서브 카테고리 단위(7개 서브 카테고리)로 나누어 수행하였음
 - 이를 통해 한 해 동안 최대 84개 (12개월×7개 서브 카테고리)의 뉴스 기반 키워드 묶음을 도출하도록 설계
 - 다만 보안/해킹 카테고리의 경우 문서 수가 너무 적어 키워드가 도출되지 않아 실제로는 72개의 뉴스 기반 키워드 묶음이 도출되었음
- 뉴스 기반 키워드 추출에 사용하는 모델은 개별 문서를 임베딩한 후, 이들을 군집화하고 각 군집에서 대표 키워드를 추출하는 방식으로 작동함
 - 이때 전체 33만여 개 문서를 한 번에 입력할 경우, 매우 소수의 문서만을 다루는 희소 토픽이 대규모 일반 토픽에 흡수될 가능성이 큼
 - 예를 들어 ‘양자컴퓨팅’을 다루는 소수의 문서가 ‘일반적인 컴퓨터/하드웨어’ 토픽에 흡수될 경우, 최종 키워드 목록에서는 ‘양자컴퓨팅’ 관련 키워드가 전혀

- 등장하지 않을 수 있음
- 서브 카테고리는 IT 일반, 모바일, 인터넷/SNS, 통신/뉴미디어, 보안/해킹, 컴퓨터, 게임/리뷰, 과학 일반의 8개로 구성
- 이 가운데 게임/리뷰를 제외한 7개 카테고리에 대해 각각 별도로 키워드 추출을 수행함으로써, 각 서브 카테고리의 특성을 반영한 보다 세분화된 키워드 집합을 얻을 수 있도록 설계함
- 키워드는 시점 및 카테고리 간에 중복되어 선정될 수 있음
- 뉴스 기사를 의미 단위로 임베딩 하기 전에 규칙 기반 전처리 과정을 거쳤으며, 이 과정을 통해 같은 내용이라도 표현 방식 차이로 인해 모델이 불필요하게 다르게 인식하는 상황을 줄이고자 하였음
- 전처리 후 사전 학습된 딥러닝 임베딩 모델을 사용하여 기사 본문을 의미 공간의 벡터로 변환(임베딩)하였음
 - 모델 개발 시에는 임베딩 모델로 BAAI에서 개발한 범용 임베딩 모델인 "BAAI/bge-m3"를 사용
 - ※ bge-m3는 다국어 환경을 지원하도록 설계된 임베딩 모델로, 최대 8,192 토큰까지의 긴 입력을 처리할 수 있어 긴 문서도 한 번에 임베딩할 수 있는 장점을 가짐 (Zilliz(2025))
 - 긴 입력 허용 범위 덕분에 기사를 여러 조각으로 분할하지 않고도 문서 전체를 임베딩할 수 있어, 문서 내 맥락 정보를 보다 잘 보존하면서 임베딩 파이프라인을 단순화할 수 있음
 - 또한 bge-m3는 다국어 검색 및 장문 문서 검색 벤치마크에서 우수한 성능을 보여, 본 연구와 같이 뉴스 데이터를 대상으로 한 토픽 분석 작업에 적합하다고 판단
- 이후 임베딩된 문서를 군집화하여 서로 유사한 문서들로 구성된 토픽 집합을 도출하였음
 - 고차원 임베딩 벡터를 UMAP (Uniform Manifold Approximation and Projection)을 통해 2~10차원 정도로 축소하였음
 - UMAP은 데이터의 국소 구조를 잘 보존하면서 군집 친화적인 저차원 표현을 생성하는 것으로 알려져 있어 차원 축소에 적합함

- 이러한 차원 축소 과정을 통해 고차원 공간에서 발생하기 쉬운 계산 비용 증가와 희소성 문제를 완화하고, 잡음을 제거하여 안정적인 군집 구조를 형성할 수 있음
- 이후 축소된 벡터에 대해 HDBSCAN을 적용하여 자동으로 군집 수를 결정하고, 밀도 기반으로 토픽을 형성
- HDBSCAN은 잡음을 별도 클래스로 분리하고, 밀도가 낮은 영역의 포인트를 잡음으로 처리함으로써 보다 안정적인 토픽 구조를 제공함
- 토픽 집합을 도출한 후, 각 토픽을 대표하는 단어를 각각 10개씩 선정
 - 각 토픽을 하나의 가상 문서로 간주하고, 토픽 내 기사들을 통합한 뒤, c-TF-IDF(class-based TF-IDF)를 계산하여 해당 토픽을 다른 토픽과 잘 구분해주는 단어를 추출
 - ※ c-TF-IDF(class-based TF-IDF)는 각 토픽에 속한 문서들을 하나의 큰 문서로 묶어, 그 안에서 자주 등장하지만, 전체 코퍼스에서는 상대적으로 덜 등장하는 단어에 더 높은 가중치를 부여하는 방법
 - 이를 통해 각 토픽을 가장 잘 대표하는 핵심 키워드를 효율적으로 도출할 수 있음
 - 이러한 절차를 통해 토픽별로 10개의 대표 단어를 산출할 수 있으나, 이 단계에서 도출되는 단어는 각 토픽에서 상대적으로 자주 등장하고 다른 토픽에서 덜 등장하는 단어들의 집합에 불과하므로, 실제로 해당 토픽을 가장 잘 설명 및 대표하는 단어를 선정하기 위해 추가적인 후처리 과정을 수행할 필요가 있음
- 토픽별로 도출된 10개의 대표어 가운데, 해당 토픽을 가장 잘 대표하는 2개의 키워드를 최종적으로 선정하였음
 - 토픽별 후보 키워드에 대해서는 표현 파이프라인을 활용하였음
 - 먼저 Maximal Marginal Relevance (MMR)를 적용하여 토픽과의 관련성은 유지하면서 서로 의미가 과도하게 중복되는 후보를 제거 및 재정렬하였음
 - ※ 예를 들어 '생성형 AI' 토픽의 후보 단어가 {생성형, 챗GPT, 오픈AI, 인공지능, AI, 대규모언어모델}이라면, MMR은 먼저 토픽과 가장 관련성이 높은 단어(예: '생성형')를 선택한다. 이후 두 번째 단어를 고를 때는 토픽과의 관련성은 높으면서도 이미 선택된 단어와는 의미가 너무 겹치지 않는 단어(예: '챗GPT'나

‘대규모언어모델’)에 더 높은 점수를 부여

- 이를 통해 ‘AI, 인공지능’처럼 거의 같은 의미의 단어들만 반복적으로 선택되는 것을 방지하고, 토픽을 다양한 관점에서 설명할 수 있는 대표 단어 조합을 얻을 수 있도록 하였음
- 이후 KeyBERT를 이용해 토픽 대표 벡터를 도출하고, MMR을 통해 도출된 키워드와 BERTopic을 통해 도출된 토픽 대표 벡터 간의 코사인 유사도를 계산한 후 코사인 유사도가 가장 높은 두 개의 키워드를 선정해 키워드 순위를 재도출하였음
 - ※ KeyBERT는 문서(또는 토픽 대표 벡터)와 후보 키워드 간의 코사인 유사도를 계산하여, 문서와 가장 유사한 키워드/키프레이즈를 상위에 배치하는 기법을 말함
- 이를 통해 중복도가 낮으면서 실제 기사 문맥과 의미적으로 가장 잘 부합하는 대표 키워드 집합을 최종적으로 도출하였음
- 최종적으로는 10개의 단어 중 의미적으로 가장 적합한 2개를 선정하여, 단어1, 단어2와 같은 형식의 단일 문자열 형태의 키워드로 결합하였음

3. ICT 키워드 선정

- 앞에서의 방식을 통해 토픽별 대표 키워드를 산출한 후, 이 중 ICT 관련성이 높은 키워드를 별도로 정제하는 작업을 수행하였음
 - 토픽별 대표 키워드가 모두 ICT 키워드라고 보기는 어려운 경우가 발생할 수 있어 신뢰성 있는 ICT 키워드 도출을 위해 검토 작업을 거침
- 이를 위해 Google의 Gemini API를 활용하였음
 - Gemini API는 구글이 제공하는 생성형 인공지능 서비스로, 텍스트 입력에 대해 요약, 분류, 키워드 필터링 등 다양한 자연어 처리 작업을 수행할 수 있음
 - 키워드 선정 시에는 프롬프트 엔지니어링을 활용하였으며, 기본적인 구조는 먼저 ICT 키워드의 정의를 설명한 뒤, 키워드 선정 규칙을 명시하고 이후 입력값으로 자동 산출된 토픽별 대표 키워드 목록을 제공하여 이 중에서 ICT 관련 키워드를 선별하도록 하는 방식임

- 다만 새로운 기술이나 서비스(예: LG의 '안티딥보이스')처럼 Gemini의 사전 학습 어휘에 포함되지 않은 용어가 등장하는 경우, 모델이 해당 개념에 대한 충분한 이해 없이 분류를 수행할 수 있다는 한계가 존재함
- 이를 보완하기 위해 Gemini의 검색 기능을 활용하도록 프롬프트를 개선하였음
- 구체적으로, 모델이 인식하지 못하는 키워드가 입력될 경우 인터넷 검색을 통해 해당 키워드의 정의나 설명을 먼저 조회한 후, 그 결과를 바탕으로 ICT 관련 여부를 판단하도록 지시하는 내용을 프롬프트에 추가하였음

〈표 4-1〉 프롬프트 예시(일부 발췌)

너는 한국어 뉴스 키워드에서 ICT(정보통신기술) 관련 주제를 골라 대표적인 ICT 키워드로 묶어주는 어시스턴트이다.

[ICT 키워드의 정의]

- ICT란 'ICT'는 **정보통신기술(Information and Communication Technology)**의 약자로, 정보를 처리하고 전달하는 데 필요한 모든 기술을 통칭하는 개념으로 디지털 기술, 컴퓨팅, 네트워크, 통신 인프라, 소프트웨어·플랫폼 서비스, 디지털 디바이스 및 관련 하드웨어를 포함한다.
- 예: AI, 생성형 AI, 온디바이스 AI, 클라우드, 5G·6G·통신망, 스마트 디바이스, 로봇, 자율주행·모빌리티, 블록체인, 스마트홈, 스마트시티, 우주, 위성 통신 등.

[ICT가 아닌 키워드 예시 (선택하지 말 것)]

- 순수 날씨·기상(폭염, 대설 등), 일반 건강·의료(치매, 혈압 등), 순수 정치·선거, 뉴스 자정/언론사 IT, 게임, 단순 금융 이벤트(임시 주총, 사장단 인사, 주주총회, 할인·쿠폰 등)처럼 디지털 기술·플랫폼·네트워크와 직접적인 연관이 없는 경우는 ICT 키워드로 보지 않는다.

[작업 설명]

Step 1 (내부 작업, 출력하지 말 것):

- 아래 입력에서 ICT와 직접 관련된 키워드 쌍만 선택한다.
- 키워드 쌍 안에 하나라도 ICT 기술 또는 서비스/플랫폼이 포함되어 있지 않다면 그 쌍은 버린다.

Step 2 (실제 출력):

- Step 1에서 선택된 키워드 쌍을 의미가 비슷한 것끼리 그룹으로 묶는다.
- 각 그룹을 하나의 일반적인 ICT 키워드(짧은 명사구)로 대표한다.

[대표 키워드 선택 규칙]

- 1) 일반 기술 카테고리 사용
 - 대표 키워드는 원칙적으로 '기술·서비스·플랫폼' 그 자체를 가리켜야 한다.
 - 다음과 같은 대표 카테고리들에 잘 맞는 경우, 해당 카테고리명을 대표 키워드로 사용한다.
예) AI, 생성형 AI, 온디바이스 AI, 클라우드, 반도체, AI 반도체, 통신/5G, 스마트폰, 스마트홈, 스마트시티, 로봇, 자율주행/모빌리티, 블록체인, 보안/개인정보, 우주, 위성 통신, 플랫폼 서비스, OTT/스트리밍, 디지털 플랫폼, 스마트 공장, 디스플레이(OLED) 등.
 - 단, 구체적인 신기술·서비스명을 일반 카테고리로 역지로 대체하지 않는다.
- 2) 신규·구체 기술명(세부 키워드) 처리
 - 다음 조건을 만족하는 경우에는, 일반 카테고리 대신 해당 기술명을 그대로 대표 ICT 키워드로 사용할 수 있다.
예) 안티딥보이스, 텔코 LLM, GPUaaS, xAI, 제미니이 CLI 등 새로운 기술·모델·패러다임 이름
 - 입력 키워드 쌍 안에 등장하는 고유한 기술명으로,
 - '클라우드', '보안', 'AI', '플랫폼 서비스' 같은 상위 개념으로만 대체하면 정보가 손실되는 경우

- 이 경우 출력 JSON의 key는 "안티딤보이스", GPUaaS 처럼 원래 기술명을 그대로 사용한다.
- 단, 단순 버전명·제품명·캠페인명 보다는, 재사용 가능한 기술 개념에 가까운 경우에만 유지한다.
(예: "그록3", "CES 2025"처럼 특정 버전/이벤트명은 가급적 "생성형 AI", "AI 모델" 등으로 통합)

3) 그룹핑 원칙

- 의미가 매우 유사한 키워드 쌍들(동일 기술/서비스를 가리키는 표현)은 하나의 대표 키워드로 묶는다.
- 대표 키워드는 가능하면 2~6글자 정도의 짧은 명사구로 유지한다.
- 동일한 신기술명이 여러 입력 쌍에서 반복되면, 그들을 모두 같은 대표 키워드(그 신기술명) 아래에 모은다.

[검색(grounding) 사용 규칙]

- 입력 키워드 쌍에서 생소한 용어/약어/신기술명이 등장하여 의미가 불확실하면, Google Search를 사용해 해당 용어의 정의와 ICT 관련성을 먼저 확인한다.
- 검색 결과 기준으로 ICT 기술·서비스·플랫폼에 직접 해당할 때만 대표 키워드로 유지한다.
- 검색해도 ICT와 직접 관련이 약하거나 단순 기업/제품/이벤트/정책/사건명으로 판명되면 대표 키워드로 쓰지 말고 상위 기술 카테고리로 통합하거나 제외한다.
- 이미 널리 알려진 일반 ICT 카테고리(AI, 반도체, 클라우드, 통신/5G 등)는 검색하지 않는다.
- "처음 보는 고유 기술명/약어"일 때만 검색한다.

자료: 연구진 작성

- 서브 카테고리별로 12개월 동안 도출된 대표 키워드를 한 번에 Gemini에 입력하고, ICT와 무관한 키워드를 제거한 뒤 의미가 유사한 키워드를 하나의 주제로 묶어 재표현하도록 설계하였음
- 최종 출력은 상위 카테고리 아래에 '생성형 AI', '온디바이스 AI' 등 통합된 ICT 키워드가 위치하고, 각 통합 키워드마다 이를 구성하는 세부 뉴스 키워드 목록이 덕서너리 형태로 정리된 구조를 가짐

〈표 4-2〉 최종 출력 예시 - IT 일반 카테고리 ICT 키워드 선정 결과

```

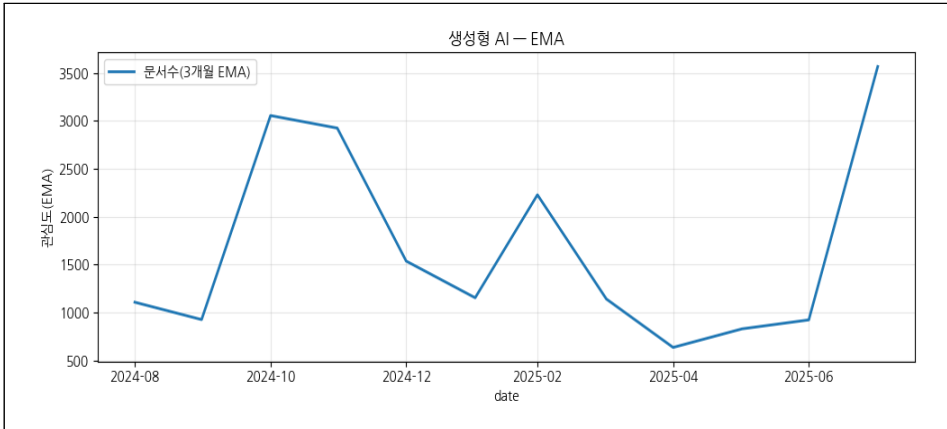
"IT_일반": {
  "생성형 AI": [
    "AI 기술 · AI",
    ...
    "테슬라 주가 · 오픈 AI"
  ],
  ...
  "온디바이스 AI": [
    "애플 AI · 애플 인텔리전스",
    ...
    "갤럭시 AI · 노트북",
  ],
}

```

자료: 연구진 작성

- 이렇게 선정한 ICT 키워드별 월별 관심도 지표(3개월 지수이동평균으로 처리된 문서 수)를 계산하여 시계열 추세를 확인할 수 있음
 - 시계열 추세는 해당 키워드에 대한 관심도 측정의 방법으로 활용 가능할 것으로 보임

[그림 4-4] 시계열 추세 결과물 예시 - 생성형 AI 키워드



자료: 연구진 작성

제3절 결과 및 평가

- 구축한 자동화 모형을 통해 분석 기간 대상으로 도출한 최종 키워드 세트는 다음과 같음

〈표 4-3〉 IT 일반 카테고리 키워드 추출 결과

키워드	평균 문서수	표준편차
생성형 AI	1873.25	2055.66
딥시크	388.00	1191.68
보안/개인정보	225.50	277.59
가상자산/블록체인	212.00	106.50
AI 반도체	176.67	127.37
AI 플랫폼	151.17	226.94
플랫폼 서비스	128.67	197.29
반도체	115.17	74.35
통신/모바일	110.33	162.09
자율주행/모빌리티	108.08	99.69
스마트폰	97.333333	107.48
HBM	96.58	101.60
온디바이스 AI	68.00	101.73
우주/위성	64.50	52.46
로봇	64.25	86.48
양자 컴퓨팅	58.67	51.11
OTT/스트리밍	54.50	71.99
스마트홈	54.42	111.24
딥페이크	42.83	87.68
클라우드	38.08	68.88
스마트 시티	31.75	68.70
안티딥보이스	31.17	59.21
국가 AI	30.92	46.22
메타버스	20.67	51.79
xAI	13.92	33.35
AI 비서	11.83	40.99
스마트디바이스	8.75	30.31
스마트 공장	7.00	24.25
의료 AI	6.25	21.65

키워드	평균 문서수	표준편차
제미나이 CLI	5.92	20.50
디지털 트윈	5.83	20.21
드론	5.58	19.34
AR	5.42	18.76
금융 AI	4.58	15.88

주: 분석 기간 IT 일반 카테고리 기사를 활용, 딥러닝기반 프레임워크로부터 도출된 키워드를 정리하고 각 키워드별로 통계치(평균, 표준편차)를 제시하였음

자료: 연구진 작성

〈표 4-4〉 타 카테고리 키워드 추출 결과

서브 카테고리	키워드	평균 문서수	표준편차
과학 일반	AI	259.25	413.07
	우주/위성 기술	173.50	151.44
	배터리	140.83	154.81
	양자 컴퓨팅	11.25	26.52
	의료 AI	10.16	35.21
	로봇	5.25	18.18
	유전자 편집	4.33	15.01
	CCU 기술	4.25	14.72
모바일	스마트폰	581.75	614.47
	통신/네트워크	186.58	283.83
	AI 서비스	135.75	230.33
	AI 에이전트	119.66	414.53
	보안/개인정보	79.25	144.34
	온디바이스 AI	73.50	114.70
	AI PC	38.41	93.59
	플랫폼 서비스	26.33	51.85
	스마트폰 시장	21.50	40.47
인터넷/SNS	AI 모델	13.25	45.89
	생성형 AI	737.83	885.37
	AI 서비스	368.33	721.35
	플랫폼 서비스	242.50	571.32
	가상자산	62.75	87.44
	AI 반도체	62.08	162.39
	OTT/스트리밍	26.16	90.64
	모빌리티	25.58	61.33
딥페이크	22.91	79.38	

서브 카테고리	키워드	평균 문서수	표준편차
	통신/모바일	22.33	54.13
	보안	20.33	48.40
통신/뉴미디어	AI	942.00	743.84
	통신/모바일	269.75	383.07
	보안	152.33	148.98
	USIM	124.25	233.17
	OTT/스트리밍	107.33	146.28
	IPTV/미디어 플랫폼	75.83	133.20
	스마트폰	24.00	53.37
	양자암호/PQC	13.50	32.35
	반도체	12.66	30.35
	온디바이스 AI	11.75	40.70
	AI 클라우드	11.25	38.97
	위성 통신	8.83	30.59
	안티딥보이스	8.08	28.00
	AI 반도체	8.00	27.71
	컴퓨터	생성형 AI	287.91
클라우드		150.33	287.66
AI 에이전트		54.50	97.72

주: 분석 기간 각각의 서브 카테고리별로 기사를 활용, 딥러닝기반 프레임워크로부터 도출된 키워드를 정리하고 각 키워드별로 통계치(평균, 표준편차)를 제시하였음
 자료: 연구진 작성

- IT 일반 서브 카테고리에서 도출된 연간 키워드 집합은 ‘생성형 AI’, ‘온디바이스 AI’, ‘AI 반도체’, ‘딥페이크’, ‘양자 컴퓨팅’, ‘의료 AI’, ‘우주/위성’, ‘클라우드’, ‘플랫폼 서비스’ 등 다수의 핵심 ICT 이슈를 포함하고 있음
 - 흥미로운 점은, IT 일반 카테고리에서 도출된 키워드 집합이 다른 서브 카테고리에서 도출된 키워드 집합과 상당 부분 중복된다는 것임
 - 이는 ‘IT 일반’이 명칭 그대로 전체 ICT 이슈를 포괄하는 상위 범주에 해당하기 때문이며, 해당 카테고리의 기사 수가 전체 기사 중 약 50%를 차지해 상대적으로 다양한 주제를 포함하고 있기 때문으로 판단됨
 - 이러한 특성으로 인해 IT 일반 서브 카테고리에서 추출된 키워드 집합은 다른 서브 카테고리의 핵심 이슈를 상당 부분 포괄하는 “대표적 키워드 풀”로 해석할 수 있음

- 기존 연구에서 수행했던 보고서 기반 ICT 키워드와 뉴스 및 딥러닝 기반의 ICT 키워드를 비교해보면 몇 가지 특징이 발견됨
- 보고서 기반 키워드는 주로 거시적 수준의 개념을 중심으로 구성되어 있는 반면, 뉴스 키워드에서는 해당 개념에 포함될 수 있는 다양한 키워드들을 함께 발굴할 수 있는 장점이 있음
 - ※ 보고서에서는 ‘생성형 AI’라는 키워드만 제시되는 반면, 뉴스 기반 방식에서는 ‘온디바이스 AI’, ‘국가 AI’, ‘AI 비서’, ‘의료 AI’ 등 생성형 AI에 매칭될 수 있는 후보 키워드가 함께 제시됨
- 두 키워드 모두 ICT 정책/산업 논의에서 주목받는 핵심 키워드 및 이슈를 포함하고 있으며, 거시적인 측면에서의 키워드는 유사하다고 볼 수 있음
- 보고서 기반에서는 과거에 핵심 키워드로 등장한 키워드에 대한 관심이 낮아졌다 해도 당장 확인이 어려운 반면, 뉴스 기반 키워드에서는 이러한 키워드가 제외되어 보다 시의성 있는 결과를 확인할 수 있었음
- 보고서 기반 키워드보다 뉴스 기반 키워드의 성격이 보다 이질적인데, 보고서 기반 키워드처럼 ‘자동화’, ‘네트워크’, ‘플랫폼’ 등 상위 개념도 포함하고 있지만 ‘AI 반도체’, ‘금융 AI’, ‘의료 AI’ 등 특정 산업/응용 분야가 결합된 영역도 함께 제시하고 있음
- 즉, 기사 기반 키워드는 보고서 기반 키워드와 마찬가지로 거시적인 기술 축을 포함하면서도, 그 축 아래에서 실제 언론 보도와 시장 담론에서 언급되는 구체적인 기술 수준의 정보까지 함께 포착하고 있다는 점에서 더욱 세분화되고 동적인 구조를 가진다고 할 수 있음

〈표 4-5〉 보고서 기반 키워드와 뉴스 기반 ICT 키워드 비교

보고서 기반 ICT 키워드	뉴스 기반 ICT 키워드
생성형 AI	생성형 AI
	온디바이스 AI
	국가 AI
	AI 비서
	의료 AI

보고서 기반 ICT 키워드	뉴스 기반 ICT 키워드
	금융 AI
자동화	스마트 공장 스마트홈 스마트 시티
네트워크 6G	통신/모바일
보안 디지털 안전	보안/개인정보
로봇	로봇
우주	우주/위성
플랫폼	플랫폼 서비스 AI 플랫폼
모빌리티	자율주행/모빌리티
디지털 트윈	디지털 트윈
딥페이크	딥페이크
퀀텀 컴퓨팅	양자 컴퓨팅
드론	드론
X	스마트디바이스
X	스마트폰
X	가상자산/블록체인
X	반도체
X	AI 반도체
X	OTT/스트리밍
X	클라우드
X	메타버스
X	AR

주: 앞서 보고서 기반으로 추출한 키워드 리스트와 뉴스 기반 방식으로 추출한 키워드 리스트를 비교하여 매칭한 결과임

자료: 연구진 작성

- 본 연구의 결과는 보고서 검토와 수작업에 크게 의존하던 기존의 키워드 선정 과정을 뉴스 데이터에 기반한 정량 분석 체계로 전환하는 데 기여할 것으로 예상됨
 - 특히 기사 수집에서부터 키워드 도출까지의 전 과정을 자동화함으로써, 추후 분석 시간을 확장하거나 새로운 연/월 데이터를 추가하더라도 동일한 절차를 반복 적용할 수 있는 재현성과 확장성을 동시에 확보하였다는 점에서 방법론적 성과를 가짐
 - 이는 향후 ICT 뉴스경기지수 산출 과정 전반을 보다 자동화하고, 민간 관심 기반의 키워드 체계를 상시적으로 갱신할 수 있는 기반 인프라를 마련했다는 점에서 의미가 큼
- 또한 딥러닝 및 LLM 기반의 분석 방법을 구축하여 키워드 선정 과정에서의 효율성을 높였다는 것도 의미가 있음
 - 기존에는 연구자가 토픽별 키워드 목록을 일일이 검토하며 ICT 관련성을 판별하고, 유사 표현을 수작업으로 통합해야 했던 과정을 상당 부분 자동화하게 되었음
 - 더불어 특정 개인의 직관이나 판단에 의존하기보다, 명시적인 규칙과 프롬프트를 바탕으로 동일한 입력에 대해 반복 가능한 결과를 산출함으로써, 키워드 선정 과정의 일관성과 재현성을 높였다는 점에서도 방법론적 성과를 가짐
- 뉴스 기반의 키워드 도출은, 민간 부문에서 실제로 반복/집중적으로 언급되는 이슈를 데이터 기반으로 확인할 수 있게 되었다는 점에서 의의를 가짐
 - 더불어 단순 빈도 집계가 아니라 의미 기반 군집화와 후속 필터링을 거친 “대표 키워드 목록”으로 정리할 수 있게 되었다는 점에서 가치가 있음
 - 실제 뉴스 보도에서 반복적으로 등장하는 이슈를 바탕으로 지금 민간이 무엇에 주목하고 있는지에 대해 보다 구체적으로 파악할 수 있게 되었음

제5장 온라인 노동지수 구축

제1절 개요

- 본 장에서는 구인구직 정보를 통해 일자리의 흐름을 시의성 있게 확인하기 위해 구축한 온라인 노동지수에 대해 설명하고 주요 결과를 정리
- 디지털 전환의 가속화로 경제사회적 변화가 촉진되면서, 신속한 온라인 정보를 활용한 노동지수 구축의 필요성이 증대되고 있음
 - － 디지털 전환이 빠르게 확산되면서, 기존 조사 통계 방식으로는 현황 파악에 한계가 드러나고 있음
 - － 이를 보완하기 위해, 실시간 온라인 정보 기반의 통계 작성 방안에 대한 연구가 활발히 진행되고 있음
 - － 특히 고용은 사회적 관심이 높고 외부 충격에 민감하게 반응하는 대표적인 경제 지표
 - － 고용의 중요성과 특성을 고려할 때, 온라인 노동지수를 구축하여 시의성과 신뢰성을 갖춘 고용 현황을 제공하는 것은 큰 의미가 있음

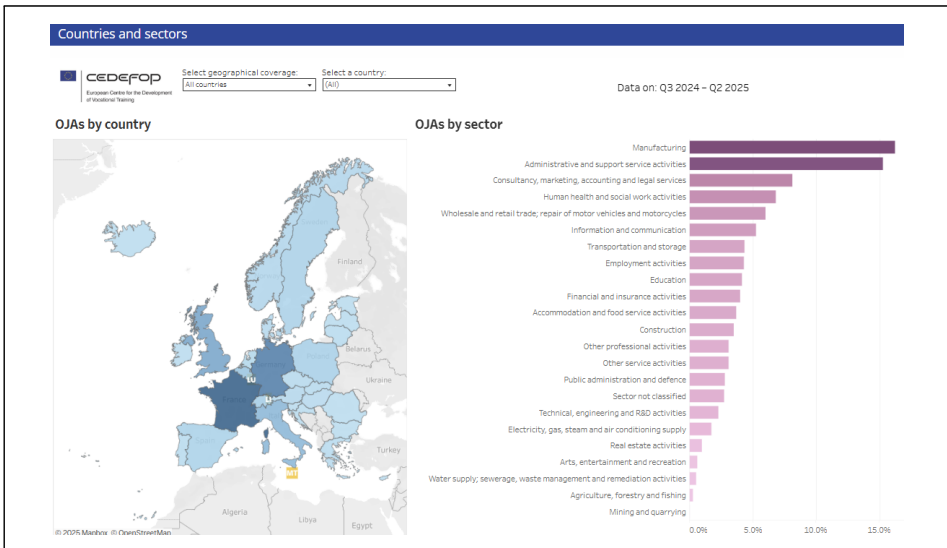
제2절 온라인 정보를 통한 노동지수 구축: 해외사례⁷⁴⁾

- 온라인 노동지수의 등장은 각 이코노미로 대표되는 노동시장 구조 변화와 밀접한 관련이 있음
 - － 각 이코노미의 확산은 온라인 기반의 구직 활동과 프리랜서 증가로 이어짐
 - － 전통적인 통계 조사 방법으로는 각 이코노미 내 노동시장 현황을 정확히 파악하기 어렵다는 지적이 있음

74) Cedefop은 Cedefop(2019), 미국 사례는 홍광표 외(2019), Oxford Internet Institute의 온라인 노동지수에 대한 설명은 Kassi & Lehtonvirta(2018)를 참조

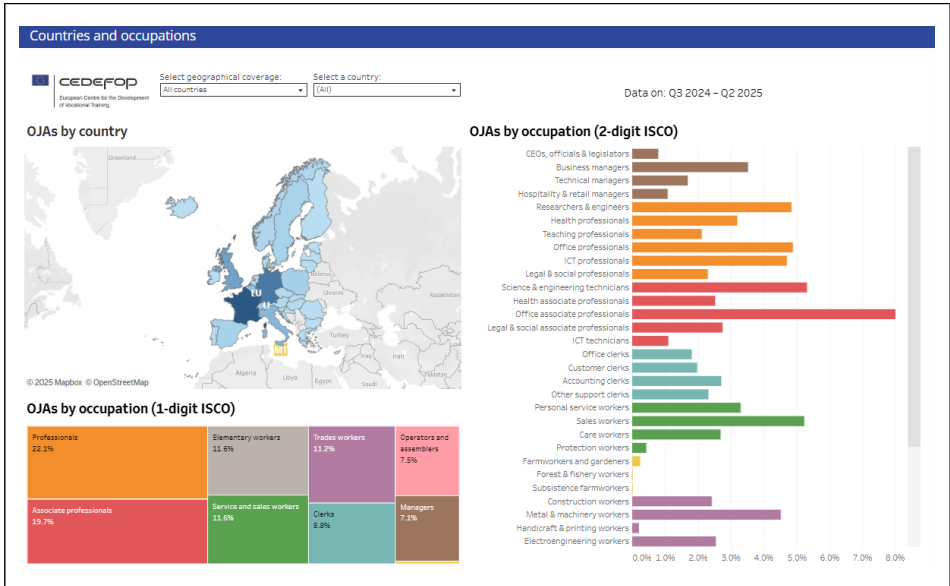
- 예를 들어, 기존 직업을 가진 개인이 온라인 구인구직을 통해 새로운 부업에 참여하는 규모를 파악하기 어려움
- 또한, 통계 조사 결과는 조사 후 일정 시간이 지나서 공표되므로, 신속한 노동시장 정보 확인에 제약이 있음
- 이러한 한계를 보완하기 위해, 실시간으로 제공되는 온라인 구인구직 정보를 활용하여 노동시장 현황을 파악하는 방안이 주목받고 있음
- 선진국에서는 온라인 정보를 활용하여 노동 현황을 분석하는 다양한 연구가 진행되고 있음
 - Cedefop(European Centre for the Development of Vocational Training)은 온라인 구인구직 사이트의 데이터를 바탕으로 EU 지역 노동시장에 대한 신속하고 시의성 높은 정보를 제공하는 프로젝트를 운영
 - Cedefop 웹사이트에서는 직종별 채용공고 비율, 지역별 채용공고 비율, 직종별로 요구되는 기술 순위 등 다양한 데이터를 제공

[그림 5-1] 국가/산업별 온라인 구인구직 광고 비중: Cedefop



자료: Cedefop, <https://www.cedefop.europa.eu/en/tools/skills-online-vacancies/markets/countries-sectors>

[그림 5-2] 국가/직종별 온라인 구인구직 광고 비중: Cedefop

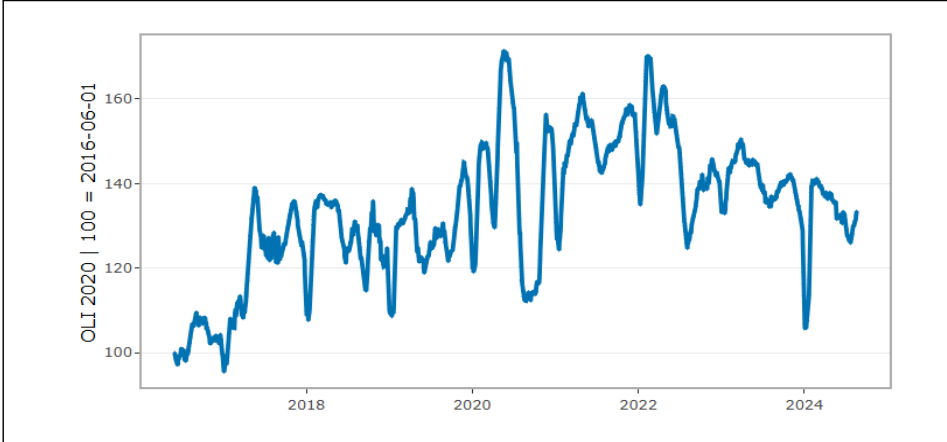


자료: Cedefop, <https://www.cedefop.europa.eu/en/tools/skills-online-vacancies/markets/countries-occupations>

- 미국에서는 피츠버그, 플로리다주, 펜실베이니아주, 뉴저지주 등에서 온라인 실시간 정보를 활용한 노동시장 분석 방안을 검토
 - 실시간 고용 데이터를 통해 지역별 노동시장의 특징을 분석하고, 이를 정책 아이디어로 연결하려는 시도가 이루어지고 있음
- International Labour Organization과 Oxford Internet Institute는 온라인 노동지수 (The Online Labour Index)를 작성하여 2024년 8월까지 공표한 바 있음
 - 이 지표는 인터넷 데이터를 활용하여 온라인 노동시장의 공급과 수요를 분석하고, 당시의 현황을 실시간으로 제공하는 것을 목표로 하였음
 - 영미권의 최대 온라인 구인구직 사이트 5개와 비영어권 언어 플랫폼 6개를 대상으로 하여 모든 프로젝트 및 직무와 관련된 구인구직 건수 수집
 - 이를 바탕으로 매일 노동지수를 작성하며, 직종별·국가별로 세분화된 데이터를 추가로 제공

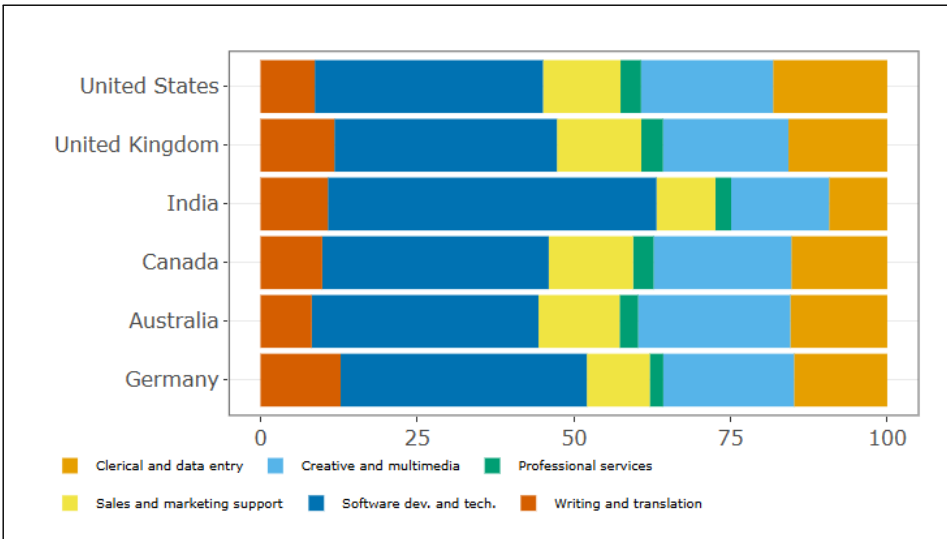
— 플랫폼별 프로젝트와 업무 수를 추적하여 국가 및 직종 간 온라인 프리랜서의 수요와 공급을 통계적으로 분석하고자 시도

[그림 5-3] 온라인 노동 지수(수요): OLI



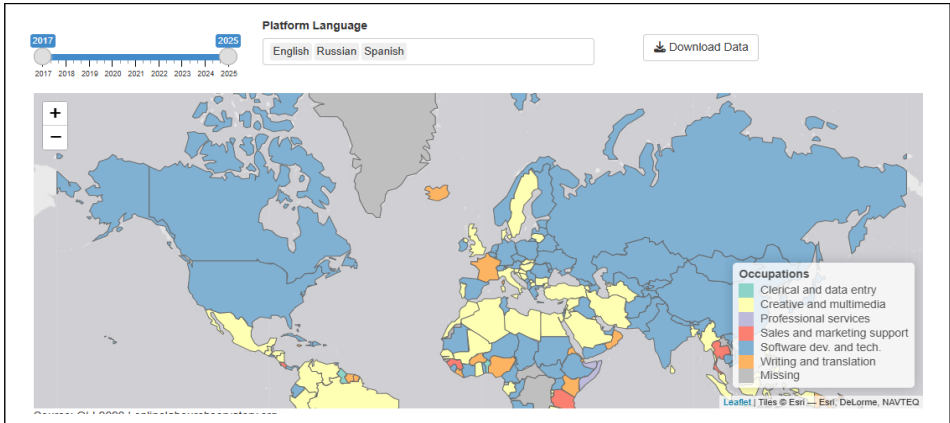
자료: OLI, <http://onlinelabourobservatory.org/oli-demand/>

[그림 5-4] 온라인 노동지수(수요) - 국가/직종별: OLI



자료: OLI, <http://onlinelabourobservatory.org/oli-demand/>

[그림 5-5] 온라인 노동지수(공급) - 국가별: Oxford Internet Institute



자료: OLI, <http://onlinelabourobservatory.org/oli-supply/>

- Lehdonvirta & Ernkvist(2011), Kuek et al.(2015)는 전문가 인터뷰와 온라인 노동 플랫폼 데이터를 활용하여 연구를 진행했으며, Farrel & Gregg(2016)는 JPMorgan Chase 데이터를 바탕으로 노동시장을 추정한 연구를 수행
- 선진국에서 이루어진 온라인 노동지수 구축은 매일의 구인구직 현황을 실시간으로 파악하고, 방대한 데이터를 활용하여 다양한 분석을 가능하게 한다는 점에서 그 가치를 인정받고 있음
 - 전통적인 통계조사로는 확인하기 어려운 프리랜서 고용 형태도 파악할 수 있는 장점이 있음
- 온라인 구인구직 사이트의 활성화가 지속적으로 이루어지는 상황에서, 온라인 기반 노동지수는 향후 고용 시장 현황을 신속하고 정확하게 파악하는 데 중요한 역할을 할 것으로 기대
 - 특히, ICT 분야는 산업 특성상 온라인에서 구인구직 활동이 활발하며, 프리랜서 비중이 높아 온라인 노동지수의 활용 가치가 더욱 큰 영역으로 평가
 - 온라인 노동지수의 성공적인 구축은 ICT 산업의 고용 현황을 빠르게 파악할 수 있게 하고, 이를 통해 효과적인 고용정책 수립에 기여할 수 있을 것으로 기대
 - 본 연구는 이러한 관점에서 앞서 정리한 선진국 사례를 참고하여, 한국형 온라인 노동지수 구축 방안을 모색하고자 함

제 3 절 온라인 노동지수 구축 방법

- 노동지수 구축을 위해 전 산업 및 전 직종을 다루는 온라인 구인구직 사이트의 정보를 수집
- 선정된 구인구직 사이트에서 제공하는 데이터를 크롤링하기 위해 전용 웹서버를 구축
 - 크롤링 대상 정보에는 일별 구인광고 총계, 직종별 및 산업별 구인광고 건수 등이 포함
 - 데이터 수집을 위한 자동화 프로세스를 구축하였으며, 수집된 데이터는 데이터베이스(DB)에 저장하여 체계적으로 관리
 - 이러한 작업을 통해 전 산업 범위를 포괄하는 대표 노동지수를 작성하였고, 이후 산업별 및 직종별로 세부 노동지수를 추가하여 분석 범위를 확대

[그림 5-6] 온라인 노동지수 플랫폼 시스템 구성도



자료: 연구진 작성

- 구축된 플랫폼을 활용하여 웹 크롤링을 통해 데이터를 수집하고, 이를 기반으로 온라인 노동지수를 산출
- 작성한 노동지수는 시각화 작업을 통해 ITSTAT에 공개중

[그림 5-7] 온라인 노동지수 웹페이지



자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

제 4 절 온라인 노동지수 분석

- 온라인 노동지수의 현황 및 특징을 살펴보기 위해, 2020년 4월 12일부터 2025년 12월 14일까지의 크롤링 자료를 활용하여 분석을 수행
 - 각 사이트의 일별 채용공고 건수를 주별 평균으로 환산하고, 최초 수집 시점의 채용공고 건수를 기준값(100)으로 설정하여 지수 형태로 변환⁷⁵⁾
 - 사이트별로 계산한 지수를 평균하여 최종 온라인 노동지수를 산출
 - ICT 산업 및 관련 직종의 지수를 타 산업 및 타 직종과 비교하기 위해, 전 산업 채용공고를 제공하는 채용공고 사이트 데이터를 활용
 - 기존에는 4개 사이트를 바탕으로 분석을 진행하였으나, 2023년 9월 이후에는

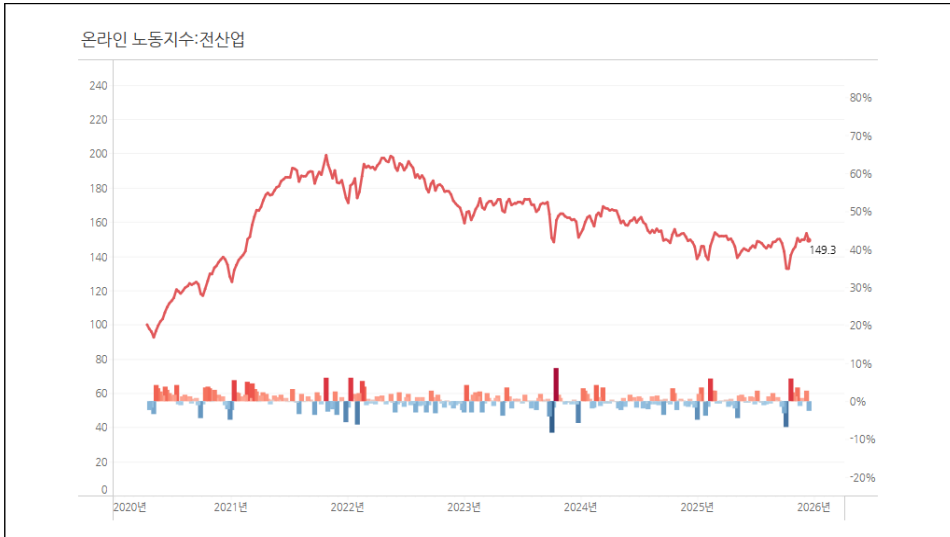
75) 산업별 및 직종별 구분 시 일별 공고 건수의 변동성이 큰 점을 보완하기 위해 주간 결과를 사용하였다.

효율적인 노동지수 산출을 위해 자료 가용성을 검토한 결과 2개 사이트를 활용하는 방향으로 개편⁷⁶⁾

- 전산업, 산업별, 직종별 노동지수를 각각 구축
 - 산업별, 직종별 노동지수 작성을 위해 각 사이트에서 제공하는 산업/직종 분류를 정리하여 매칭테이블을 생성
 - 사이트별로 산업과 직종을 분류하는 단계와 정보가 다르기에, 가장 자세하게 제공되는 분류 수준을 검토하여 연계 가능성을 확인
 - ICT 산업 중심의 정보 제공을 위해 산업과 직종 모두에서 ICT 관련 분류를 구성
 - 산업은 ICT 서비스업, ICT 제조업, 비ICT 서비스업, 비ICT 제조업 4개 분류로 나누었음
 - 직종은 IT, 전기전자생산, 보건의료, 그 외 생산제조, 그 외 서비스 5개 분류로 구분
- 전산업 노동지수의 주요 결과는 다음과 같음
 - 2025년 전산업 온라인 노동지수는 2024년 대비 소폭 감소하는 추세를 보이며, 전체 기간으로도 하향세를 그리고 있음을 확인 가능
 - 2023년 대비 2024년에는 감소폭이 일시적으로 축소되었으나, 2024년 대비 2025년에는 감소 폭이 소폭 증가
 - 2025년 첫째 주에 141을 기록한 지수는 2월 마지막 주에 154로 최고점을 기록한 후 5월 첫째 주에 139까지 하락한 뒤 반등하였으며, 이후 등락을 거쳐 12월 둘째 주에는 149를 기록
 - 채용공고가 활발하게 게시되는 신년 초에 공고가 가장 높은 것을 확인할 수 있으며, 10월 초 급감 이후 공고 수준이 회복되는 양상을 보임
 - 최초 시점인 2020년 4월에 비해서는 여전히 공고수가 높은 수준임을 확인 가능함

76) 2023년 9월 이전 자료의 경우 2개 사이트를 기준으로 지수를 재산출하였다.

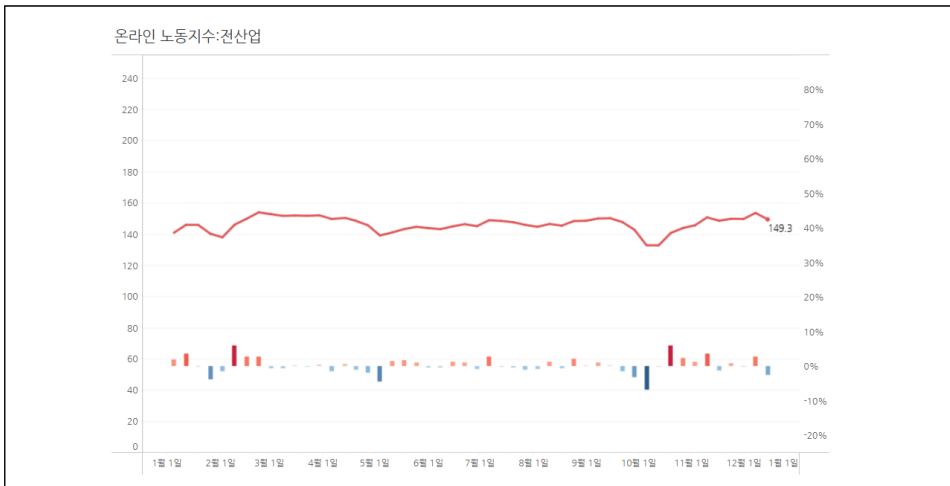
[그림 5-8] 전산업 온라인 노동지수 현황



주: 2020. 4. 12.~2025. 12. 14.까지의 2개 사이트 산업별 채용공고를 이용. X축은 해당 연도 주차를 나타냄

자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

[그림 5-9] 전산업 온라인 노동지수 현황: 2025년

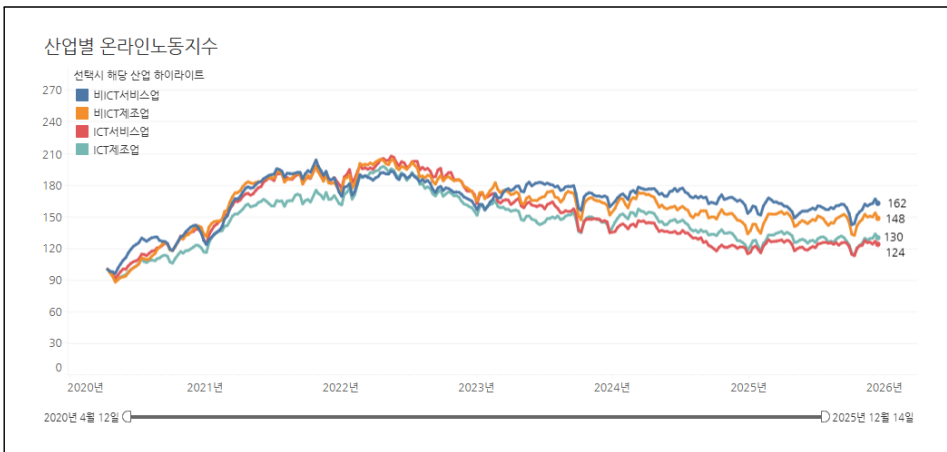


주: 2025년의 2개 사이트 산업별 채용공고를 이용. X축은 주차를 나타냄

자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

- 구인구직 사이트의 산업별 공고 정보를 수집 후 분류하여 ICT 서비스업, ICT 제조업, 비 ICT 서비스업, 비 ICT 제조업 이상 4개의 대분류별로 노동지수를 구축
- 산업별 노동지수의 주요 결과는 다음과 같음
 - 산업 전반적으로 시작점인 2020년 대비 2022년까지 공고수가 빠르게 증가하였으나, 이후에는 서서히 감소세로 전환되고 있음
 - 비 ICT서비스업의 경우 2023년부터 타 산업 대비 채용공고수 증가가 두드러지며, 2025년 기준으로 가장 높은 값을 기록하고 있음
 - 반대로 ICT 서비스업은 2022년까지 가장 높은 값을 기록하였으나, 2025년에는 상대적으로 낮은 수준에 머물고 있음
 - 2025년에는 모든 산업에서 연중 큰 변동 없이 등락을 보이며 전반적으로 안정적인 흐름을 나타냄

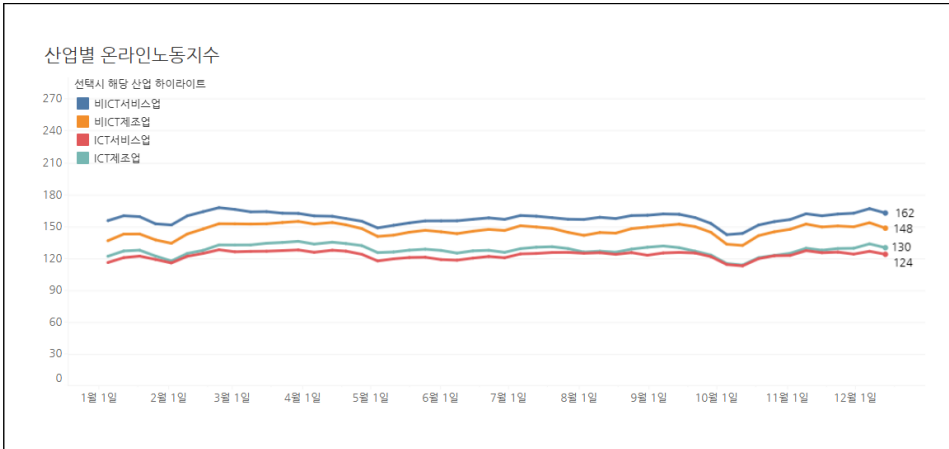
[그림 5-10] 산업별 온라인 노동지수 현황



주: 2020. 4. 12.~2025. 12. 14.까지의 2개 사이트 산업별 채용공고를 이용. X축은 해당 연도 주차를 나타냄

자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

[그림 5-11] 산업별 온라인 노동지수 현황: 2025년

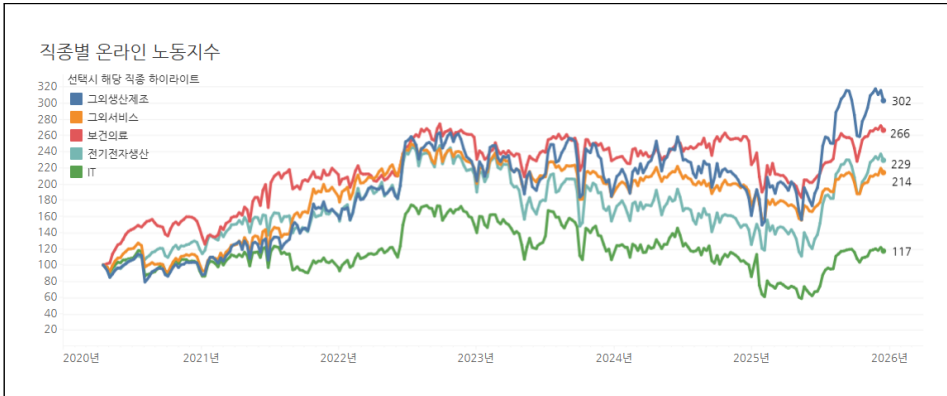


주: 2025년의 2개 사이트 산업별 채용공고를 이용. X축은 주차를 나타냄

자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

- 구인구직 사이트의 직종별 공고수를 수집하여 IT, 전기전자생산, 보건의료, 그 외 생산제조, 그 외 서비스 5개의 대분류로 분류 후 노동지수를 구축
- 직종별 노동지수의 주요 결과는 다음과 같음
 - 직종별 구인구직은 2022년 후반까지 꾸준한 상승세를 보이고 있으며, 2023년 들어 하락 추세를 보이다 2025년 중반 이후 전 직종에서 상승세를 보이고 있음
 - 그 외 생산제조가 2025년 중반 이후부터 가장 높은 수준의 공고수를 기록하고 있으며, 보건의료와 전기전자생산이 그 뒤를 잇고 있음
 - IT 직종은 2020년 대비로 공고수가 크게 증가하지 않은 것을 확인
 - 2025년 기간을 보면 보건의료는 중반부터 상승세를 보이며, 연말에는 연초보다 높은 수준의 공고수를 기록하고 있음을 확인 가능
 - 그 외 생산제조는 2025년 중반 이후 상승세가 본격화되며, 하반기에는 모든 직종 중 가장 큰 폭의 증가를 보이며 연말까지 최고 수준을 유지

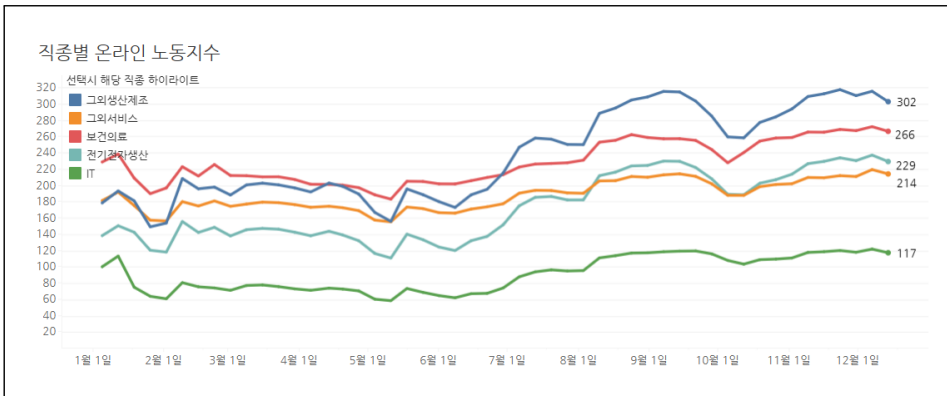
[그림 5-12] 직종별 온라인 노동지수 현황



주: 2020. 4. 12.~2025. 12. 14.까지의 2개 사이트 직종별 채용공고를 이용. X축은 해당 연도 주차를 나타냄

자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

[그림 5-13] 직종별 온라인 노동지수 현황: 2025년



주: 2025년의 2개 사이트 직종별 채용공고를 이용. X축은 주차를 나타냄

자료: ITSTAT, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html>

제5절 IT 직종 공고 분석

- 본 절에서는 IT 개발자 관련 공고를 분석한 결과를 정리하였음
- IT 직종은 디지털 전환과 생성형 AI의 활성화에 직접적으로 연결되어 있는 직종이며, 이러한 변화로 인한 영향 또한 크게 받고 있음
 - IT 직종의 공고 변화를 살펴보는 것은 이러한 변화를 확인하고 IT 관련 직무의 수요가 어떻게 재편되고 있는지 확인할 수 있다는 점에서 의의가 클 것으로 판단
- 2025년 2월 5일~11월 17일 기간에 수집 가능했던 10,781개의 공고 정보를 활용
 - 공고에는 직무, 요구 스킬, 요구 커리어 정보가 포함되어 있음
 - 이러한 정보를 사용하여 2025년 IT 직종 공고의 흐름과 특징을 살펴보고자 함
- 직무를 구분하기 위해 직무 공고 정보, ITSQF 직무체계, 글로벌 IT 트렌드 등을 반영하여 분류를 구성
 - 서버/백엔드 개발자, 프론트엔드 개발자, 풀스택 개발자, 앱 개발자, AI/ML 엔지니어, LLM/생성형 AI 엔지니어, 데이터 전문가, 클라우드 엔지니어, DevOps 엔지니어, 시스템 /네트워크 엔지니어, 임베디드 펌웨어 엔지니어, 로봇 자율주행 엔지니어, 의료IT/디지털헬스케어 엔지니어 이상 13개 범주로 구분
 - 각각의 직무 분류를 나타내는 키워드를 작성하고, 해당 키워드가 직무 공고 설명에 포함되어 있을 경우 직무 분류에 포함되는 공고로 정의하였음
 - ※ 한 공고가 여러 직무 분류에 포함될 수 있음

〈표 5-1〉 직무 분류 및 키워드 목록

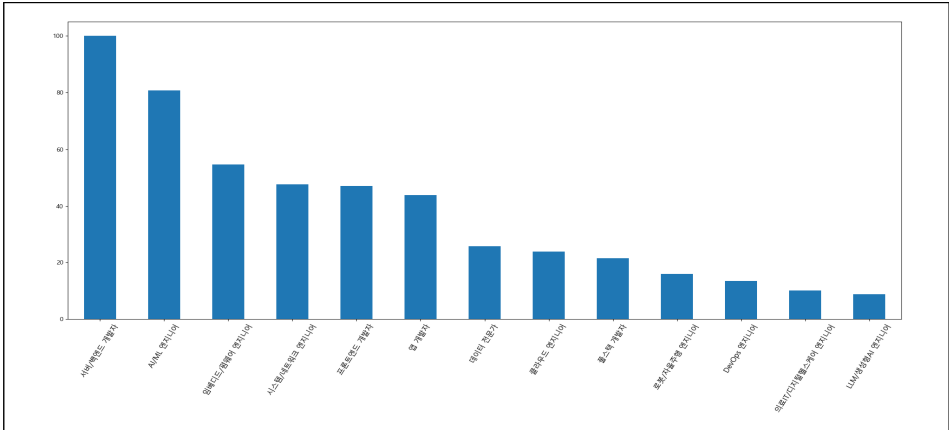
직무명	키워드 목록
서버/백엔드 개발자	백엔드, backend, back-end, 서버, server, 백엔드
프론트엔드 개발자	프론트엔드, frontend, front-end
풀스택 개발자	풀스택, fullstack, full-stack
앱 개발자	앱, 어플리케이션, application, 애플리케이션, app, 모바일, mobile, android, ios
AI/ML 엔지니어	ai, 인공지능, 머신러닝, 딥러닝, 알고리즘, machinelearning, deeplearning, ml

직무명	키워드 목록
LLM/생성형AI 엔지니어	llm, 생성형, generative, gpt, 언어모델
데이터 전문가	데이터, data, 빅데이터, bigdata
클라우드 엔지니어	클라우드, cloud, 인프라, infrastructure, aws, azure, gcp
DevOps 엔지니어	devops, 자동화, automation, cicd, ci/cd, kubernetes, docker
시스템/네트워크 엔지니어	시스템, system, 네트워크, network
임베디드/펌웨어 엔지니어	펌웨어, firmware, 하드웨어, hardware, 임베디드, embedded, hw, h/w
로봇/자율주행 엔지니어	로봇, robot, 자율주행, autonomous, 드론, drone
의료IT/디지털헬스케어 엔지니어	의료기기, 의료영상, 디지털헬스케어, medical, healthcare, 헬스케어

자료: 연구진 작성

- 직무 분류별 공고수 정보를 살펴보면 다음과 같음
 - 서버/백엔드 개발자 관련 공고가 가장 많았으며 그 다음으로 AI/ML 엔지니어가 위치하였음
 - 전통적인 서버/백엔드 개발자 다음으로 AI/ML 엔지니어의 공고수가 높아진 것은, AI 관련 직종에 대한 수요가 꾸준히 상승하고 있음을 확인할 수 있는 결과
 - 특정 분야의 맞춤형 직무라고 볼 수 있는 로봇/자율주행과 의료IT/디지털헬스케어 직무도 별도의 공고로 등장하고 있다는 점도 최근 IT 채용 시장의 흐름을 잘 보여주고 있음
 - 월별 추세를 보면 시간에 따른 추세는 유사하며, 7월에 공고가 많았음을 알 수 있음
 - 공고가 많았던 7월에 AI/ML 엔지니어에 대한 수요가 가장 많았다는 것도 최근 AI 관련 직종에 대한 수요를 보여주는 모습으로 해석

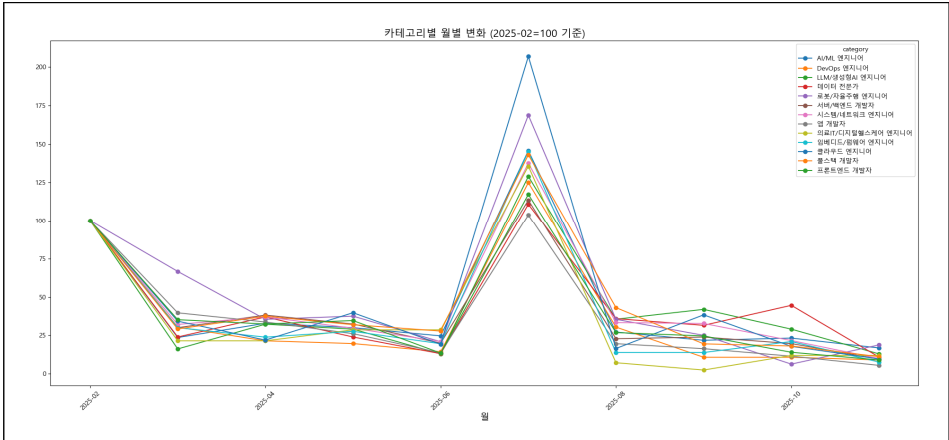
[그림 5-14] 직무별 공고수 합계



주: 가장 많은 공고수를 기록한 서버/백엔드 개발자를 100으로 두고, 각 직종 공고수의 상대값을 시각화하였음

자료: 연구진 작성

[그림 5-15] 직무별/월별 공고 변화



주: 시작 지점인 2월의 공고수 합계를 100으로 했을 때 각 월의 공고수를 의미

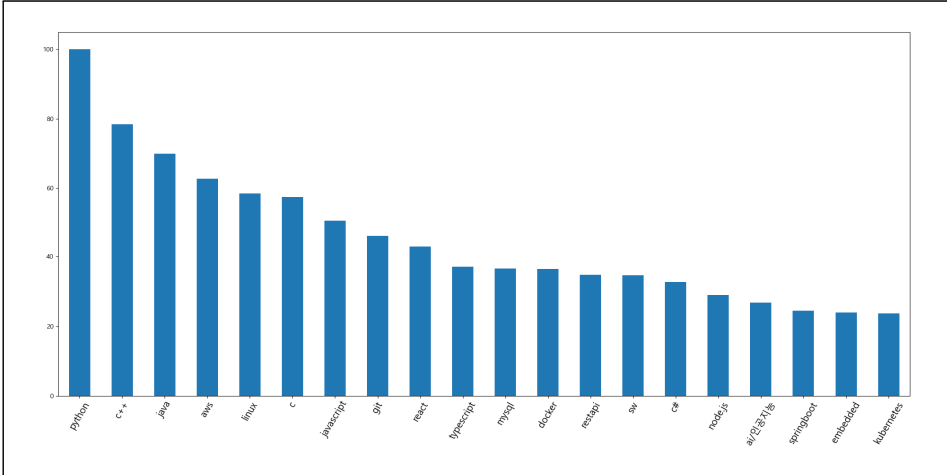
자료: 연구진 작성

○ 기술별 공고수 정보를 살펴보면 다음과 같음

- 요구사항으로 가장 많이 언급된 기술은 Python으로, 실무에서 가장 많이 활용되는 프로그래밍 언어임을 확인할 수 있음

- 다음으로 언급량이 많은 것은 C++, Java, AWS 순이었음
- 흥미로운 것은, 요구 기술에 “AI/인공지능”이 별도의 키워드로 등장하고 있으며 빈도수도 상위 20위 안에 들어간다는 점

[그림 5-16] 기술별 공고수 합계(상위 20개)



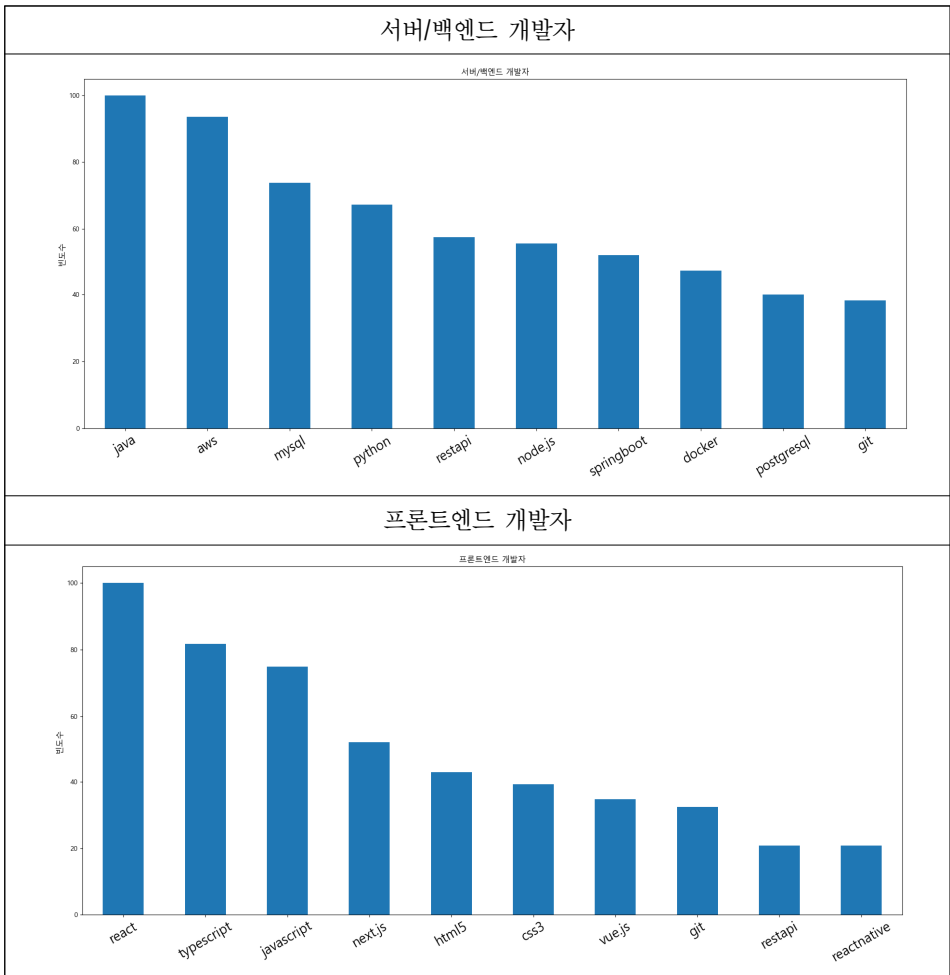
주: 공고수 합계 기준으로 상위 20개 기술을 정리한 결과이며, 가장 많은 공고수를 기록한 Python을 100으로 두고 각 기술 공고수의 상대값을 시각화하였음
 자료: 연구진 작성

- 직종별 / 기술별로 공고수 정보를 살펴보면 다음과 같음
 - 서버/백엔드에서는 Java 기술에 대한 수요가 많았으며, AWS와 MySQL도 높은 수준을 기록
 - 프론트엔드와 풀스택에서는 React 관련 공고가 많았으며, TypeScript와 JavaScript도 높은 비중을 기록
 - 앱 개발에서는 iOS가 가장 높은 점유율을 기록하였으며 Flutter, Git에 대한 수요도 높았음
 - AI/ML 엔지니어와 LLM/생성형AI 엔지니어, 데이터 전문가는 Python 수요가 높았고, AI/인공지능이라는 범용적인 기술 수요를 요청하는 경우도 많았음
 - 클라우드 엔지니어에서는 AWS와 Linux, DevOps 엔지니어에서는 Kubernetes와

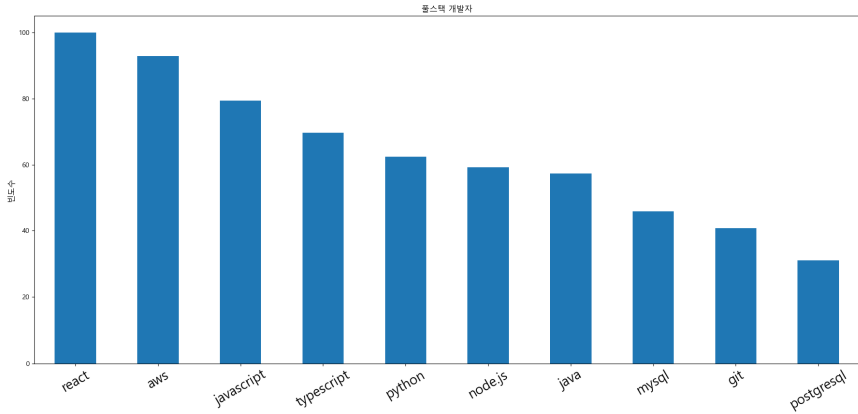
AWS 관련 수요가 높았음

- 시스템/네트워크 엔지니어에서는 Linux와 C++, 임베디드/펌웨어 엔지니어에서는 C++에 대한 기술 수요가 많았음
- 로봇/자율주행 엔지니어와 의료IT/디지털헬스케어 관련 직종에서도 C++과 C에 대한 수요가 높은 것을 확인
- 대부분의 직종에서 Python에 대한 기술 수요가 높다는 점도 알 수 있음

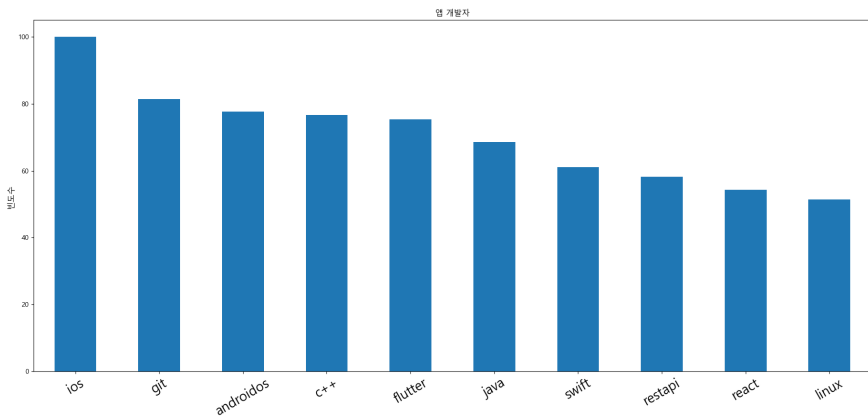
[그림 5-17] 직무별/기술별 공고수 합계(상위 10개)



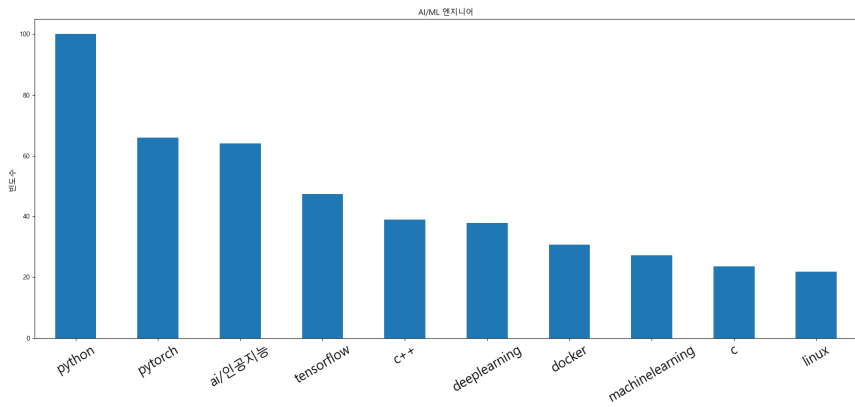
풀스택 개발자



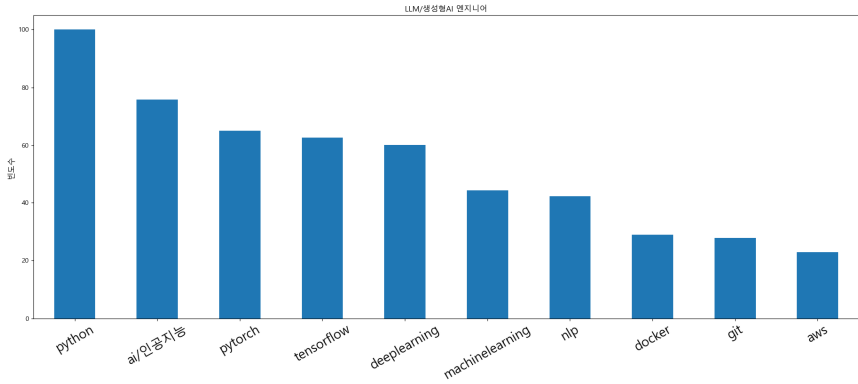
앱 개발자



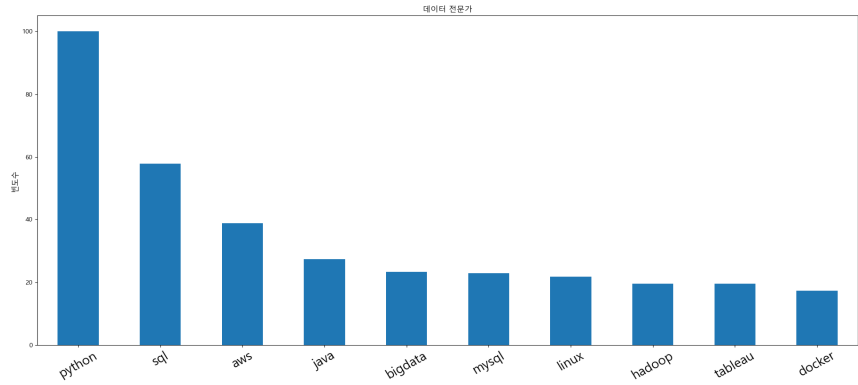
AI/ML 엔지니어



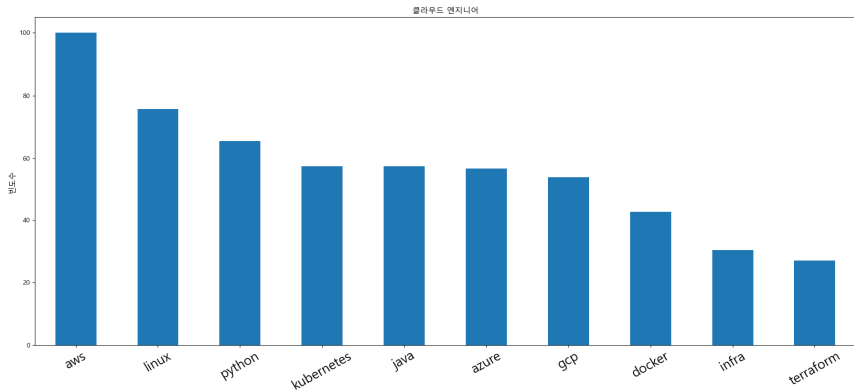
LLM/생성형 AI 엔지니어



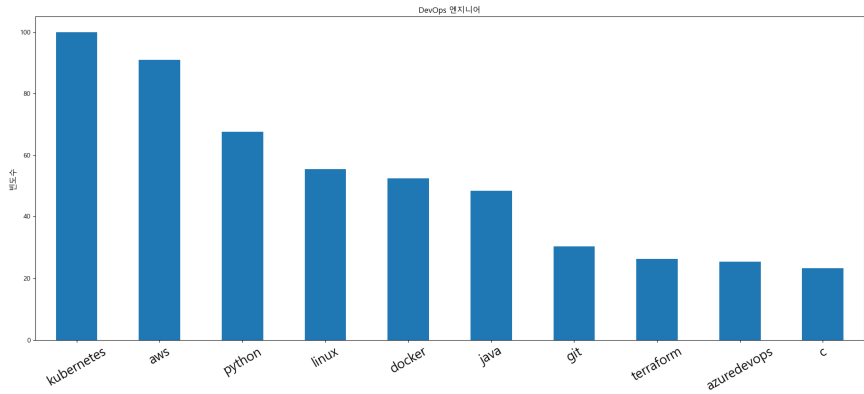
데이터 전문가



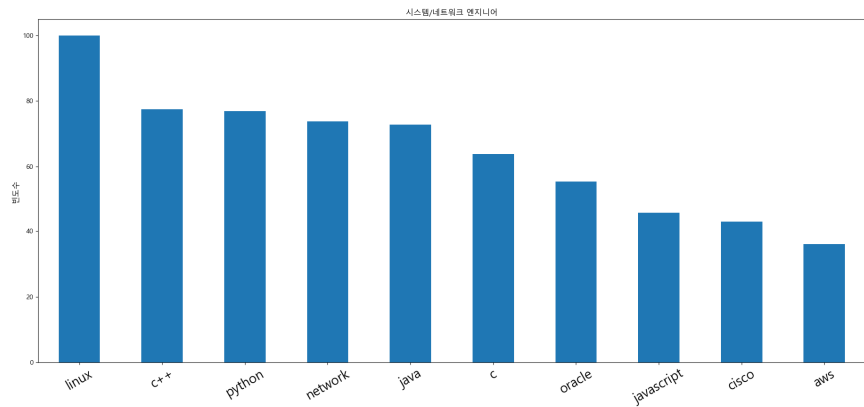
클라우드 엔지니어



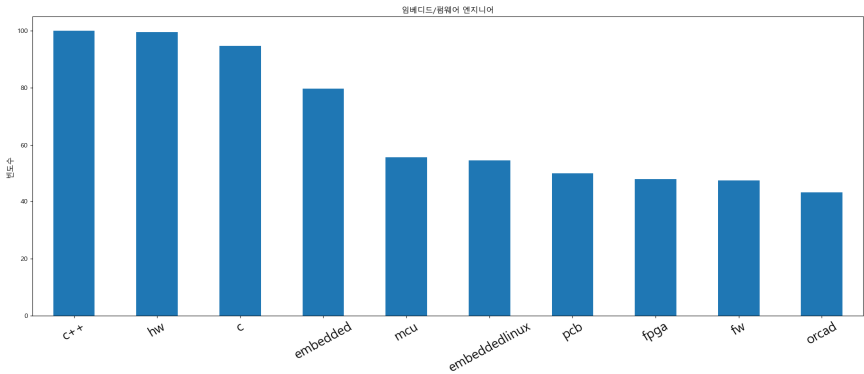
DevOps 엔지니어



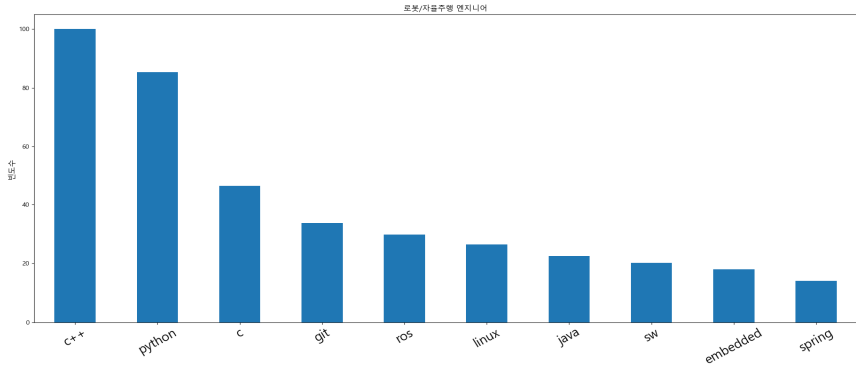
시스템/네트워크 엔지니어



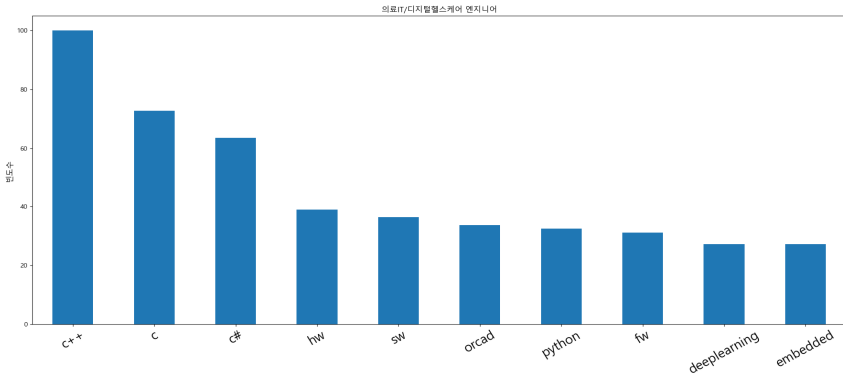
임베디드/펌웨어 엔지니어



로봇/자율주행 엔지니어



의료IT/디지털헬스케어 엔지니어



주: 각 직무별로 공고수가 많았던 상위 10개 기술의 등장 빈도를 합산 후, 가장 빈도가 높았던 기술을 100으로 두어 상대적 빈도를 시각화한 결과

자료: 연구진 작성

참 고 문 헌

[국내 문헌]

강기운 외 9인(2022), “디지털 헬스케어 진흥 및 보건의료데이터 활용 촉진에 관한 법률안”, (의안번호 177751). 대한민국 국회.

강유진, 김주원(2022.11.30.), 「헬스케어 디지털 트윈」, 《KISTEP브리프》, 46호, 한국과학기술기획평가원.

과학기술정보통신부 보도자료(2025.03.12.), “양자기술, 연구개발을 넘어 산업화로”.
_____ (2024.04.25.), “양자과학기술 대도약, 디지털을 넘어 “퀀텀의 시대”로”.

_____ (2024.2.16.), “제2차 인공지능 데이터 융합 네트워크 개최”.
과학기술정보통신부(2023.06.), “대한민국 양자과학기술 전략”.

_____, 『ICT실태조사』.

과학기술정보통신부·정보통신기획평가원, 『ICT 수출입동향』.

공개SW포털(2025.05.27.), “[기획브리핑] 국내 주요 기업의 오픈소스 LLM 공개 현황 분석”,
https://www.oss.kr/oss_guide/show/9246eca5-f639-484c-be09-797d76fc9582

국민일보(2022.02.06.), “코로나, 디지털헬스케어에 폭발력 더하다… IT기업·통신사 ‘눈독””,
<https://n.news.naver.com/mnews/article/005/0001503247>

김지은 외(2020.12.), 「디지털 헬스 산업 분석 및 전망 연구」, 한국보건산업진흥원.

김현중, 임종호, 이해영, 이상호(2019), 『온라인 뉴스 기사를 활용한 경제심리보조지수 개발』,
《국민계정리뷰》, 2019년 제2호, 한국은행.

뉴스1(2023.08.19.), “한미일 파트너십 새시대…안보·공급망, 인태·우주까지 협력”,
<https://n.news.naver.com/mnews/article/421/0006997988>

_____ (2022.06.27.), “‘양자 시대 연다’…과기정통부 ‘2022 양자정보주간’ 진행”,
<https://n.news.naver.com/mnews/article/421/0006180376?rc=N&ntype=RANKING>

_____ (2022.01.12.), “과기정통부, ICT원천연구개발에 904억 투자…양자컴·차세대 반도체 개발”, <https://n.news.naver.com/mnews/article/421/0005840490>

- _____ (2020.05.05.), “이번주 ‘한국판 뉴딜’ 밑그림 공개…디지털·의료·일자리 키워드”,
<https://www.news1.kr/economy/policy/3925148>
- 네이버 CLOVA(2025.04.23.), “AI 생태계에 씨앗을 뿌리다: 상업용 오픈소스 AI, HyperCLOVA X SEED”, <https://clova.ai/tech-blog/ai-생태계에-씨앗을-뿌리다-상업용-오픈소스-ai-hyperclova-x-seed>
- 동아일보(2022.01.21.), “현대車, 양자컴퓨터로 차세대 배터리 개발 나선다”,
<https://n.news.naver.com/mnews/article/020/0003407064>
- _____ (2022.01.11.), “LG전자 “美 IBM과 양자컴퓨팅 기술 개발 협력””,
<https://n.news.naver.com/mnews/article/020/0003404977>
- 디지털경제뉴스(2023.12.15.), “업스테이지, 사전학습 LLM ‘솔라’ 10.7B 모델 공개”,
<https://www.denews.co.kr/news/articleView.html?idxno=28062>
- 디지털비즈온(2024.07.19.), ““디지털 헬스케어”… 이미 시작된 미래”,
<https://www.digitalbizon.com/news/articleView.html?idxno=2336495>
- 머니투데이(2025.06.12.), “젠슨 황 한 마디에 또다시 들썩…양자컴퓨팅 ETF 주목도 ↑”,
<https://n.news.naver.com/mnews/article/008/0005206753>
- _____ (2025.01.30.), “LG유플러스, PQC 기술 활용한 통합 계정관리 솔루션 ‘알파키’ 출시”,
<https://n.news.naver.com/mnews/article/008/0005146778>
- _____ (2025.01.15.), “마이크로소프트 ‘올해는 양자컴퓨팅 준비하는 해’ 선언”,
<https://n.news.naver.com/mnews/article/008/000514167>
- _____ (2025.01.07.), ““모빌리티·디지털헬스, 그리고 ‘…CES 2025 휩쓴 ‘이 기술’”,
<https://n.news.naver.com/mnews/article/008/0005137693>
- _____ (2024.12.27.), “KOSEF 미국양자컴퓨팅 ETF, 7거래일만에 순자산 500억 돌파”,
<https://n.news.naver.com/mnews/article/008/0005133795>
- 메디게이트뉴스(2025.01.07.), “국내 디지털 헬스 CES혁신상 수상 제품·서비스 23개”,
<https://medigatenews.com/news/1815064460>
- 문장원·윤형진·선미란(2019.12.11.), 「해외 디지털 헬스케어 규제개선 동향」, 《이슈리포트》, 2019-37호, 정보통신산업진흥원.

박지원, 손녕선(2025), “멀티모달 인공지능 기술 트렌드 분석”, 《ICT 통계분석》, 2025-02호, 정보통신정책연구원.

법제처 국가법령정보센터, 「양자과학기술 및 양자산업 육성에 관한 법률」, 법률 제19990호, 제정 2023.10.31., 시행 2024.11.1.

보건복지부 보도자료(2023.06.01.), “국민건강 증진을 위해, 디지털 헬스케어 산업을 육성하겠습니다”.

_____ (2022.06.27.), “데이터 기반으로 바이오·디지털헬스 글로벌 도약 추진”. 산업통상자원부 보도자료(2023.01.30.), “디지털헬스 산업 생태계 조성을 위한 4대 정책방향 제시”.

서법석, 이영환, 조형배(2022), 『기계학습을 이용한 뉴스심리지수(NSD)의 작성과 활용』, 《국민계정리뷰》, 2022년 제1호, 한국은행.

서울경제(2025.04.28.), “AI 빗장 푼 삼성전자 반도체…메타 ‘라마4’ 전 부문 도입”, <https://www.sedaily.com/NewsView/2GRP2ULW1B>

_____ (2022.08.25.). “삼성SDS, AWS·MS 우수 기업과 ‘양자암호’ 기술 표준 만든다”, <https://n.news.naver.com/mnews/article/011/0004091499?rc=N&ntype=RANKING>

손녕선(2025), “개방형 AI 모델 동향 분석”, 《ICT 통계분석》, 2025-04호, 정보통신정책연구원.
손녕선, 이은영, 진정민, 김재민, 박재홍(2023), 『ICT 주요이슈 분석체계 구축』, 정책연구 23-28.

세계일보(2025.03.21.), “‘개발 20년 걸릴 것’ 젠슨 황, 양자컴 기업에 사과…‘상장사 있는 줄 몰랐다’”, <https://n.news.naver.com/mnews/article/022/0004020981>

_____ (2022.01.11.), “LG, IBM과 양자컴퓨팅 개발 협력”, <https://n.news.naver.com/mnews/article/022/0003656164>

신현호, 손녕선(2025), “디지털 헬스케어에 대한 온라인 인식 변화”, 《ICT 통계분석》, 2025-01호, 정보통신정책연구원.

연미영(2023.06.30.), 「디지털 헬스케어 정책 현황: 소비자 참여 기반의 디지털 헬스케어 활성화를 위한 검토」, 《소비자정책동향》, 130호, 한국소비자원.

연합뉴스(2023.08.10.), “美 “투자가 위험 키워”…AI·반도체·양자컴퓨팅 對中투자 제한(종합)”, <https://n.news.naver.com/mnews/article/001/0014122863>

이데일리(2022.04.25.), “‘바이오헬스 한류시대’ 연다…제약바이오혁신위 신설”,
<https://n.news.naver.com/mnews/article/018/0005199828?rc=N&ntype=RANKING>

인공지능신문(2019.01.03.), “인공지능(AI) 언어모델 BERT(버트)는 무엇인가”,
[https:// www.aitimes.kr/news/articleView.html?idxno=13117](https://www.aitimes.kr/news/articleView.html?idxno=13117)

전자신문(2025.07.20.), “삼성SDS, IBM AI ‘그라니트’ 도입 검토…LLM 백엔드 다변화”,
<https://www.etnews.com/202507180000073>

전종준, 안승환, 이문희, 황희진(2020), 『경제용어 감성 사전 구축방안 연구』, 《국민계정리뷰》,
 2020년 제3호, 한국은행.

조선비즈(2024.01.30.), “LG전자, 중기부와 ‘디지털 헬스케어’ 스타트업 발굴 나선다”,
<https://n.news.naver.com/mnews/article/366/0000966526>

_____ (2019.10.31.), “구글 ‘10년 뒤엔 양자컴퓨터…슈퍼컴퓨터로 1만년 할 일을 200초만에”,
<https://n.news.naver.com/mnews/article/366/0000446509>

조수지, 김홍규, 양철원(2021), 『Building the Korean Sentiment Lexicon for Finance (KOSELF)』,
 《한국증권학회지》, 50(2), 135-170.

진정민, 손녕선(2025), “양자컴퓨팅 키워드의 언론 인식 변화”, 《ICT 통계분석》, 2025-03호,
 정보통신정책연구원.

파이낸셜뉴스(2025.01.12.), “美 MS·아이온큐와 투자협약… 부산 ‘제2 시애틀’로 만든다”,
<https://n.news.naver.com/mnews/article/014/0005294383>

프레시안(2022.01.24.), “이재명·윤석열·안철수·심상정, 포스트 코로나 체제의 비전은?”,
<https://n.news.naver.com/mnews/article/002/0002228489>

한겨레(2024.11.28.), “스마트 헬스케어, 건강관리 해결책 ‘주목’”,
<https://n.news.naver.com/mnews/article/028/0002718409>

_____ (2019.05.17.), “그 알고리즘이 너무 매력적이었기에 그들은 구글을 세웠다”,
[https:// www.hani.co.kr/arti/science/technology/894339.html](https://www.hani.co.kr/arti/science/technology/894339.html)

한국산업기술진흥원(2024.04.), 「2024년 산업기술 환경예측(디지털헬스케어)」.

한국인공지능·소프트웨어산업협회(2025), 『IT 분야 / 융합산업분야 SQF 직무기술서』,
<https://www.sw.or.kr/site/sw/ex/board/View.do?cbIdx=308&bcIdx=53300>

홍광표·설귀환·박상오·문혜정(2019), 『실시간 노동시장 정보를 활용한 숙련수요 분석』, 기본 연구 2019-15, 한국직업능력개발원.

KDI경제정보센터(2021), 「디지털 헬스케어 산업 활성화 정책…발자취와 새로운 도약, 《해외동향-디지털헬스케어편》, 2021-02호.

KISTI(2010), “2010년 ICT 분야의 R&D 보고서”,
<https://scienceon.kisti.re.kr/srch/selectPORSrchTrend.do?cn=GT201001232>

MoneyS(2021.11.16.), “IBM, ‘양자컴퓨팅 시대’ 선도한다… 127큐비트 ‘이글’ 공개”,
<https://n.news.naver.com/mnews/article/417/0000754665>

SBS Biz(2025.03.30.), “AI·보안·편의성 더했다…삼성전자, 비스포크AI 가전 신제품 공개”,
<https://n.news.naver.com/mnews/article/374/0000432256>

ZDNet Korea(2025.08.07.), “개방형 AI 모델도 급이 있다… ‘오픈소스’와 ‘오픈웨이트’ 차이?”,
<https://zdnet.co.kr/view/?no=20250807091040>

_____ (2024.09.19.), “금융권 노린 네이버클라우드, ‘하이퍼클로바X’ 소형 LLM 구축”,
<https://zdnet.co.kr/view/?no=20240919095825>

_____ (2021.02.02.), “MS, 양자컴퓨팅 개발환경 누구에게나 제공”,
<https://n.news.naver.com/mnews/article/092/0002212585>

[해외문헌]

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4), 1593-1636.

Baltrušaitis, T., Ahuja, C., & Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423-443. <https://doi.org/10.1109/TPAMI.2018.2798607>

Cedefop. (2019). Annual report 2019. Luxembourg: Publications Office. <https://www.cedefop.europa.eu/en/publications/4175>.

Christian, N. A., Turuwhenua, J., & Norouzifard, M. (2025). AI and generative models in 360-degree video creation: Building the future of virtual realities. *Applied Sciences*, 15(17), 9292.

- Emanuel Knill. (2010). Quantum Computing. *Nature*, Vol.463, pp.441-443, <https://www.nature.com/articles/463441a>
- Farrel, D. & Gregg, F. (2016). Paychecks, Paydays, and the Online Platform Economy: Big Data on Income Volatility. *Proceedings of the Annual Conference on Taxation*. 109, pp. 1~40.
- Fenter, F. (2025). The Top 10 emerging technologies of 2025. *Forrester*.
Financial Times. (2025.07.20.). Space is the new factory floor. <https://www.ft.com/content/ec5918cd-78f6-413a-a903-5bd57b3c3b53>
- Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint*, arXiv:2203.05794. <https://arxiv.org/abs/2203.05794>
- Kässi, O., & Lehdonvirta, V. (2018). Online labour index: Measuring the online gig economy for policy and research. *Technological forecasting and social change*, 137, 241-248.
- Kuek, S. C., Paradi-Guilford, C., Fayomi, T., Imaizumi, S., & Ipeirotis, P. (2015). *The Global Opportunity in Online Outsourcing*, Washington, D.C.: World Bank Group.
- Lehdonvirta, V. & Ernkvist, M. (2011). *Knowledge Map of the Virtual Economy: Converting the Virtual Economy into Development Potential*. World Bank.
- Li, H. (2022). Language models: Past, present, and future. *Communications of the ACM*, 65(7), 56-63.
- Preskill, J. (2012). Quantum computing and the entanglement frontier. *arXiv preprint*, arXiv:1203.5813.
- Scott, S. L. and Varian, H. R. (2014). Predicting the present with bayesian structural time series. *International Journal of Mathematical Modelling and Numerical Optimisation*, 5(1-2), 4-23.
- Touvron, H., et al. (2023). Llama: Open and efficient foundation language models. *arXiv preprint*, arXiv:2302.13971.
- U.S. Congress. (2018). *National Quantum Initiative Act (Public Law No: 115-368)*.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017).

Attention is all you need. NeurIPS, 30.

Villars, R., Ng, S., Hamel, J., Madsen, P., Marshall, C., Schneider, L., Lava, S., Graham, S., Loomis, A., & Cooke, J. (2025). IDC FutureScape: Worldwide IT industry 2025 predictions. IDC.

WHO. (2021). Global strategy on digital health 2020-2025.

[웹사이트]

네이버 뉴스 플랫폼, <https://news.naver.com/>

네이버 Datalab, “검색어 트렌드”, <https://datalab.naver.com/keyword/trendSearch.naver>

네이버 클라우드(2025.04.23.), “AI 생태계에 씨앗을 뿌리다: 상업용 오픈소스 AI, HyperCLOVA X SEED”, <https://tinyurl.com/4zbf5dp5> (2025.07.01. 접속).

썸트렌드, <https://some.co.kr/>.

카카오(2021.12.15.), “카카오브레인, 초거대 AI 두 번째 모델 ‘minDALL-E’ 공개”, <https://www.kakaocorp.com/page/detail/9638> (2025.07.01. 접속).

Andreessen Horowitz(al6z). (2024.03.21.). 16 Changes to the Way Enterprises Are Building and Buying Generative AI. <https://a16z.com/generative-ai-enterprise-2024/> (2025.11.25. 접속).

Baig, A., Bawcom, A., Wiesinger, A., Stüer, C., Bauer, F., Brown, J. S., et al. (2025.01.06.). New Year’s resolutions for tech in 2025. McKinsey & Company. <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/new-years-resolutions-for-tech-in-2025> (2025.11.25. 접속).

Bechtel, M., Briggs, B., & Raskovich, K. (2024.12.11.). Tech Trends 2025. Deloitte Insights. <https://www.deloitte.com/us/en/insights/topics/technology-management/tech-trends.html> (2025.11.25. 접속).

Belitsoft. (2025.04.01.). BloombergGPT is live: A custom large language model for finance. <https://belitsoft.com/bloomberggpt> (2025.11.25. 접속).

Cedefop. Skills Online Vacancies-Countries and Sectors. <https://www.cedefop.europa.eu/en>

/tools/skills-online-vacancies/markets/countries-sectors (2025.12.18. 접속).

ChainPlay. (2025). NFT market declines drastically for the fifth quarter in a row. <https://chainplay.gg/blog/nft-market-decline-q2-2025/> (2025.11.25. 접속).

CNBC. (2025.01.21.). Goldman Sachs launches AI assistant, <https://www.cnbc.com/2025/01/21/goldman-sachs-launches-ai-assistant.html> (2025.11.25. 접속).

Data Science Dojo. (2025.08.11.). The complete history of OpenAI models: From GPT-1 to GPT-5. <https://datasciencedojo.com/blog/the-complete-history-of-openai-models/> (2025.11.25. 접속).

DeepSeek AI. (n.d.). DeepSeek-R1. <https://github.com/deepseek-ai/DeepSeek-R1> (2025.10.27. 접속).

Digitalhealth. (2023.06.09.). Cloud technology in healthcare benefits highlighted in new report. <https://www.digitalhealth.net/2023/06/cloud-technology-in-healthcare-benefits-highlighted-in-new-report> (2025.04.01. 접속).

Epoch AI. (2024.11.04.). How behind are open models?. <https://epoch.ai/blog/open-models-report> (2025.11.25. 접속).

European Commission. (2025.05.22.). Quantum Technologies Flagship (2018~2028). <https://digital-strategy.ec.europa.eu/en/policies/quantum-technologies-flagship> (2025.09.10. 접속).

FDA. (2020). What is Digital Health?. <https://www.fda.gov/medical-devices/digital-health-center-excellence/what-digital-health> (2025.04.01. 접속).

Gartner. (2025.07.). Hype Cycle for Compute, 2025. <https://www.gartner.com/interactive/hc/6671434?ref=solrAll&refval=493038320> (2025.09.16. 접속).

Gartner. (2025.04.). Emerging Tech Horizon for Generative AI.

Google Trends, <https://trends.google.co.kr/trends/>.

Google. (2023.12.06.). Introducing Gemini: our largest and most capable AI model. <https://blog.google/technology/ai/google-gemini-ai/> (2025.07.01. 접속).

Google AI. (n.d.). Gemma 모델 개요. <https://ai.google.dev/gemma/docs?hl=ko> (2025.10.27.

접속).
 GoOver.AI. (2024.06.01.). Llama 3 vs GPT-4: An in-depth comparative analysis. <https://seo-goover.ai/report/202405/go-public-report-en-8e483565-7cd2-4d41-9127-0b7d01247c96-0-0.html> (2025.11.25. 접속).
 Grootendorst, M. (n.d.). The Algorithm – BERTopic. <https://maartengr.github.io/BERTopic/algorithm/algorithm.html> (2025.10.27. 접속).
 Harness. (2025.03.12.), Platform engineering: Beyond the trough of disillusionment. <https://www.harness.io/blog/platform-engineering-beyond-the-trough-of-disillusionment> (2025.10.27. 접속).
 HKDCA. (2024.05.16.). 15 graphs that explain the state of AI in 2024. <https://www.hkdca.com/15-graphs-that-explain-the-state-of-ai-in-2024/> (2025.11.25. 접속).
 IBM.(n.d.). What is multimodal AI?. <https://www.ibm.com/think/topics/multimodal-ai> (2025.05.19. 접속).
 IBM.(n.d.). What is quantum computing?. <https://www.ibm.com/think/topics/quantum-computing?>(접속일자 : 2025.09.10.)
 ITSTAT, “온라인노동지수”, <https://www.itstat.go.kr/itstat/kor/labor/Labor.html> (2025.12.27. 접속).
 ITSTAT, “ICT 뉴스경기지수”, <https://www.itstat.go.kr/itstat/kor/economic/economicList.html> (2025.12.27. 접속).
 K21Academy. (2025.01.20.). Top 10 AI job roles in 2025. <https://k21academy.com/ai-ml/top-10-ai-job-roles-in-2025-k21academy/> (2025.11.25. 접속).
 KISA, “암호이용활성화 홈페이지”, <https://seed.kisa.or.kr/kisa/index.do>
 Martyn Eeles. (2023.01.01.). The Role of AI in Digital Health: How to Turn This Technology Into A Competitive Advantage. <https://www.linkedin.com/pulse/role-ai-digital-health-how-turn-technology-advantage-martyn-eeles> (2025.04.01. 접속).
 MassChallenge. (2022.05.25.). Rising Trends in Digital Health: 5 Technologies That Will Define the Future of Healthcare. <https://masschallenge.org/articles/digital-health-trends/> (2025.04.01. 접속).
 Meta. (2025.04.05.). The Llama 4 herd: The beginning of a new era of natively multimodal

- AI innovation. <https://ai.meta.com/blog/llama-4-multimodal-intelligence/> (2025.07.01. 접속).
- ____ (2024.07.23.). Introducing Llama 3.1: Our most capable models to date. <https://ai.meta.com/blog/meta-llama-3-1> (2025.07.01. 접속).
- ____ (2024.06.18.). Sharing new research, models, and datasets from Meta FAIR. <https://ai.meta.com/blog/meta-fair-research-new-releases/> (2025.07.01. 접속).
- MIT Technology Review. (2025). MIT's 10 breakthrough technologies for 2025. MIT. (2025.11.25. 접속).
- MPG ONE. (2025.08.06.). GPT OSS: OpenAI's shocking return to open-source. <https://mpgone.com/gpt-oss-guide/> (2025.11.25. 접속).
- OLI. Online Labour Observatory-Demand. <http://onlinelabourobservatory.org/oli-demand/> (2025.12.18. 접속).
- ____. Online Labour Observatory-Supply. <http://onlinelabourobservatory.org/oli-supply/> (2025.12.18. 접속).
- OpenAI. (2024.05.13.). Hello GPT-4o. <https://openai.com/index/hello-gpt-4o/> (2025.07.01. 접속).
- ____. (2023.03.14.). <https://openai.com/index/gpt-4-research/> (2025.07.01. 접속).
- Precedence Research. (2025). Digital Health Market Size and Forecast 2025 to 2034. <https://www.precedenceresearch.com/digital-health-market> (2025.04.01. 접속).
- Rockwell Automation. (2025.06.03.). Ninety-five percent of manufacturers are investing in AI to navigate uncertainty and accelerate smart manufacturing. <https://www.rockwellautomation.com/en-us/company/news/press-releases/Ninety-Five-Percent-of-Manufacturers-Are-Investing-in-AI-to-Navigate-Uncertainty-and-Accelerate-Smart-Manufacturing.html> (2025.11.25. 접속).
- Stability AI. (2024). Introducing Stable Diffusion 3.5. <https://stability.ai/news/introducing-stable-diffusion-3-5> (2025.11.25. 접속).
- Sustainability Directory. (2025.04.02.). Digital safety. <https://lifestyle.sustainability-directory.com/term/digital-safety/> (2025.11.25. 접속).
- The Guardian. (2025.08.05.). OpenAI takes on Meta by launching free and customisable

AI models. <https://www.theguardian.com/technology/2025/aug/05/openai-meta-launching-free-customisable-ai-models> (2025.11.25. 접속).

TheSequence. (2023.04.09.). The LLaMA effect: How an accidental leak sparked a series of impressive open source alternatives to ChatGPT. <https://thesequence.substack.com/p/the-llama-effect-how-an-accidental> (2025.11.25. 접속).

Together.ai. (2023.04.17.). RedPajama. <https://www.together.ai/blog/redpajama> (2025.11.25. 접속).

Vellum. (2024.05.08.). Llama 3 70B vs GPT-4: Comparison analysis. <https://www.vellum.ai/blog/llama-3-70b-vs-gpt-4-comparison-analysis> (2025.11.25. 접속).

Vivasoft Ltd. (2025.03.05.). Top 15 types of software engineering jobs in 2025. <https://vivasoftltd.com/types-of-software-engineering-jobs/> (2025.11.25. 접속).

Zilliz. (2025). BAAI / bge-m3. <https://zilliz.com/ai-models/bge-m3> (2025.11.25. 접속).

● 저 자 소 개 ●

손 념 선

- 서강대학교 경제학 박사
- 현 정보통신정책연구원 연구위원

박 지 원

- 연세대학교 통계데이터사이언스학 석사
- 현 정보통신정책연구원 연구원

진 정 민

- 한양대학교 경제학 석사
- 현 정보통신정책연구원 전문연구원

신 현 호

- 한밭대학교 신소재공학과 학사
- 전 정보통신정책연구원 위촉연구원
- 현 한국특허전략개발원 연구원

박 재 흥

- University of Texas at Austin 경영학 박사
- 현 경희대학교 경영학과 교수

정책연구 25-29-04

ICT 주요이슈 분석체계 구축

(Establishment of Analysis System for ICT Issue)

2025년 12월 일 인쇄

2025년 12월 일 발행

발행인 과학기술정보통신부 장관

발행처 과학기술정보통신부

세종 갈매로 477 정부세종청사 4동

Homepage: www.msit.go.kr

인 쇄 경 성 문 화 사



과학기술정보통신부
Ministry of Science and ICT