

# 빅카인즈를 활용한 GenAI(생성형 인공지능) 기술 동향 분석: ChatGPT 등장과 스타트업 영향 평가

이현주 (동국대학교 기술창업학과 박사과정)\*

성창수 (동국대학교 기술창업학과 교수)\*\*

전병훈 (동국대학교 기술창업학과 교수)\*\*\*

## 국 문 요 약

기술 창업 및 스타트업 분야에서는 인공지능(AI)의 발전이 사업 모델 혁신의 핵심 주제로 부상하였다. 이를 통해 벤처기업들은 경쟁력 확보를 위해 AI를 중심으로 다양한 노력을 기울이고 있다. 본 연구는 GenAI 기술의 발전과 스타트업 생태계 간의 관계를 국내 뉴스 기사를 분석하여, 기술 창업 분야의 동향을 파악하는 것을 목적으로 하였다. 본 연구는 빅카인즈(BIG Kinds)를 활용하여 1990년부터 2023년 8월 10일까지의 국내 뉴스 기사에서 ChatGPT의 등장 전후를 중심으로 GenAI 관련 뉴스 기사, 주요 이슈 및 트렌드의 변화를 조사하였으며, 네트워크 분석 및 키워드 시각화를 통해 관련성을 시각화하였다.

연구결과, 2017년부터 2023년까지 GenAI에 대한 언급이 기사 내에서 점차 증가하였다. 특히, OpenAI의 GPT-3.5를 기반으로 한 ChatGPT 서비스가 주요 이슈로 부각 되었는데, 이 서비스는 OpenAI의 DALL-E, Google의 MusicLM, VoyagerX의 Vrew 등과 같은 언어 모델 기반 GenAI 기술의 대중화를 시사하였다. 이로써 생성형 인공지능은 다양한 분야에서의 유용성을 입증하며, ChatGPT 출시 이후 국내 기업들의 한국어 언어 모델 개발 활동이 활발히 이루어지고 있는 것으로 확인되었다. 리튼 테크놀로지스와 같은 스타트업들도 GenAI를 활용하여 기술 창업 분야에서의 영역을 확장하고 있다.

본 연구에서는 GenAI 기술과 스타트업 창업 활동 간의 연관성을 확인하였으며, 이는 혁신적인 비즈니스 전략의 구축 지원을 시사하며 GenAI 기술의 발전과 스타트업 생태계의 성장을 지속해서 형성할 것으로 전망된다. 더 나아가 국제적 동향 및 다양한 분석 방법의 활용, 실제 현장에서의 GenAI 응용 가능성을 모색하는 연구가 요구 된다. 이러한 노력은 GenAI 기술의 발전과 스타트업 생태계의 성장 발전에 이바지할 것으로 기대된다.

핵심주제어: 생성형인공지능, 빅카인즈, 스타트업, ChatGPT, 기술 동향

## I. 서론

최근 10년 사이 폭발적인 속도로 고도화된 AI 기술은 다양한 산업 분야와의 융합을 통해 디지털 전환(digital transformation)의 혁신을 주도하고 있다. 이제 AI는 전문가뿐 아니라 대중에게도 국가경쟁력의 핵심으로 인식되고 있다(한규동·전병훈, 2022; 윤성만·성창수, 2023). 특히 'Generative Artificial Intelligence(생성형 인공지능, 생성 AI, 생성형 AI, 이후 GenAI)' 분야는 인간과 유사한 창의적 결과물을 생성하는 AI 모델로 주목받고 있으며, 콘텐츠 제작, 디자인, 커뮤니케이션 등의 산업분야에서 새로운 가능성을 열고 있다(Hataya et al., 2022). GenAI 초기 모델인 Google의 DeepDream과 Microsoft의 Tay는 대중의 관심보다는 외면받았다. 그 이유는 Google의

DeepDream은 2015년에 출시된 이미지 생성 모델로 인공 신경망을 사용하여 환각적이고 꿈같은 이미지를 생성할 수 있었으나 DeepDream은 이미지의 의미를 파악하는데 적합하지 않았고 종종 노이즈와 착시를 생성하여 환각적이고 불안한 이미지를 생성하여 대중에게는 관심을 받지 못했다(Suzuki et al., 2017). Microsoft의 Tay는 2016년에 출시된 챗봇으로 Twitter에서 인간과 대화를 나눌 수 있었지만, Tay는 인종차별적이고 혐오스러운 트윗으로 인해 Microsoft는 서비스를 중단하였다(Wolf et al., 2017).

OpenAI의 GPT-2는 2019년에 출시된 언어 모델로 방대한 텍스트 데이터 세트로 학습되었으며, 인간과 같은 텍스트를 생성할 수 있었으나 가짜 뉴스와 선전을 만드는데 악용될 가능성을 우려하여 소스 코드를 제한적으로 공개했었다(Radford et al., 2019). 그러나 2021년 6월에 발표된 대화형 생성 인공지능

\* 주저자, 동국대학교 기술창업학과 박사과정, hj.lee@maxted.kr

\*\* 교신저자, 동국대학교 기술창업학과 교수, redsun44@dongguk.edu

\*\*\* 교신저자, 동국대학교 기술창업학과 교수, bhjeon@dongguk.edu

투고일: 2023-07-14 · 1차 수정일: 2023-08-11 · 2차 수정일: 2023-08-24 · 게재확정일: 2023-08-28

챗봇인 ChatGPT의 등장은 이러한 우려를 넘어 GenAI 기술의 실제적인 가능성을 대중에게 보여주었다. 이러한 변화의 중심에는 OpenAI와 같은 혁신적인 스타트업들이 있다.

본 연구의 주요 목적은 GenAI 기술의 동향을 분석하고, ChatGPT의 등장이 스타트업 및 기술 창업 분야에 미친 영향을 조사하는 것이다. 이를 위해 빅데이터 분석을 활용할 예정이며, 현대 사회에서 빅데이터 분석은 의사결정, 시장 예측, 고객 행동 분석 등 다양한 분야에서 중요한 도구로 자리 잡고 있다. 빅카인즈(BIG Kinds)를 활용하여 GenAI 관련 뉴스 기사를 분석할 계획이다. GenAI 기술의 시기별 트렌드와 주요 이슈를 파악하며, 특히 ChatGPT의 등장 전과 후의 GenAI 기술 분야의 변화를 중점적으로 조사할 것이다. 이를 통해 스타트업 및 기술 창업 분야에서의 GenAI 기술 활용 가능성과 그 영향을 탐구하려 한다. 이러한 연구는 GenAI 기술의 현재 상황과 미래 전망을 이해하는 데 중요한 역할을 할 것이며, 기술 창업과 스타트업 분야에서 유니콘 기업으로의 성장 가능성 탐색에 큰 의미가 있다(김주희 외, 2022).

GenAI와 스타트업에 대한 심도 있는 연구가 시작단계에 있다. 본 연구를 통해 국내 스타트업에게 유용한 인사이트와 지침을 제공하며 국내에서 시작 단계인 GenAI와 스타트업 분야의 연구와 발전에 기여하고자 한다.

## II. 이론적 배경

### 2.1 생성형인공지능(GenAI)

2022년 Gartner의 주요 전략적 기술 트렌드(Forti et al., 2022)에서는 12가지 중요한 트렌드 중 하나로 GenAI를 소개하고 있다. 이에 따르면, “많은 AI는 특정 결론을 위해 훈련받지만, 진정한 혁신을 가져오는 AI는 스스로 변화와 발전을 이룬다. GenAI는 샘플 데이터를 통해 디지털 아티팩트를 학습하며, 이를 기반으로 훈련 데이터와 유사하되 반복되지 않는 새롭고 현실적인 아티팩트를 창조한다. 이러한 능력으로, 생성형 AI는 기업들에게 혁신의 새로운 동력을 제공할 것이다.” 더불어, Forti et al.(2022)에 따르면, “영국의 금융 행동 감독 기관(Financial Conduct Authority)은 생성형 AI를 활용해 실제 500만 건의 결제 데이터를 바탕으로 합성 결제 데이터를 만들었다. 이 데이터는 개인 정보를 노출시키지 않으면서 새로운 결제 방식을 개발하는 데 활용될 예정이다. 2025년까지 생성형 AI가 처리하는 데이터의 비율은 현재 1% 미만에서 10%로 크게 증가할 것으로 예상된다. 시작하려면, 입증된 생성형 AI 사용 사례를 선택하여 콘텐츠 제작과 연구 개발을 가속화하고, 새로운 제품 개발을 빠르게 진행하며, 아티팩트를 개인화하는 방향으로 나아가야 한다.”라고 GenAI의 중요성을 강조하고 있다.

GenAI는 Transformer(2022), GAN, VAE와 같은 다양한 기술을 포함하고 있다. 이 중에서도 Transformer는 딥러닝의 한 분

야로, 주로 문장 내의 단어들 간의 관계를 파악하여 맥락과 의미를 학습하는 신경망으로 알려져 있다. 이 기술은 이미지 인식, 분류, 그리고 큰 데이터 세트에서 텍스트나 이미지를 생성하는 등의 다양한 작업에 활용되며, 구글의 번역 서비스에서도 이를 활용하고 있다. Transformer 기반의 AI 모델로는 BERT, GPT-3, ChatGPT, LaMDA, Wu-Dao 등이 있으며, 이들은 어텐션 메커니즘을 통해 입력 데이터의 중요한 부분을 감지한다(OpenAI, 2022). 한편, GAN은 Ian Goodfellow와 그의 팀에 의해 개발되었다(Goodfellow et al., 2014). GAN은 두 개의 신경망, 즉 생성 네트워크와 판별 네트워크로 구성되어 있다. 생성 네트워크는 기존 데이터와 유사한 새로운 데이터나 콘텐츠를 만들어내는 반면, 판별 네트워크는 원본과 생성된 데이터를 비교하여 차이를 감지한다(Isola et al., 2018). 이 두 네트워크는 서로 경쟁하며 학습을 진행, 더욱 정교한 결과물을 생성한다(Goodfellow et al., 2014). 마지막으로, VAE는 기존의 오토인코더를 발전시킨 모델로, 입력 데이터를 압축한 후 다시 원래의 형태로 복원하는 과정을 통해 학습한다. 이 과정을 통해 VAE는 원본 데이터와 매우 유사한 새로운 데이터를 생성할 수 있다(정승원 외, 2019).

Goodfellow와 그의 팀이 제시한 Generative Adversarial Nets 논문은 GenAI 분야의 큰 발전을 이끌었다. 변정호·권용주(2023)의 연구에서 “생성적 인공지능이란 사전에 학습된 텍스트, 오디오, 이미지 데이터를 기반으로 새로운 콘텐츠를 생성하거나 기존 데이터를 변형하는 인공지능 기술”이라고 정의하였다. 최준영(2023)은 문화일보를 통해 “생성적 인공지능은 데이터 샘플의 맥락을 분석하고 사용자의 입력 간의 상호 비교가 가능한 인공지능”으로 설명하였다.

GenAI는 현재 다양한 분야에서 활용되고 있다. 특히, 콘텐츠 제작, 이미지 변환, 비디오 관련 애플리케이션 등에서 그 효과를 볼 수 있다. 텍스트 생성에 있어서는, ChatGPT와 같은 대화형 GenAI 도구를 사용하여 사용자와의 상호작용을 통해 원하는 스타일의 텍스트를 생성한다. 이러한 ChatGPT는 OpenAI의 LLM GPT-3 모델을 기반으로 하며, 다양한 온라인 데이터를 학습한 LLM GPT-3.5 모델도 있다(Aydn & Karaarslan, 2022; Bommarito & Katz, 2022; Wu et al., 2023). Microsoft(2023)는 코파일럿 기능을 통해 워드, 엑셀, 파워포인트, 아웃룩 등의 애플리케이션에서 텍스트 콘텐츠를 자동으로 생성하고 수정, 이로써 사용자의 작업 효율성을 향상시킨다. Google의 Bard, Notion AI, DeepMind의 Sparrow, 네이버의 HyperCLOVA 등은 대화형 GenAI 애플리케이션의 주요 예시로 알려져 있다(성현희, 2022; 박수현, 2023). 더불어, DALL-E 같은 도구는 텍스트 설명을 기반으로 현실감 있는 이미지를 생성하는 능력을 보여준다(Cho et al., 2022; Reddy et al., 2021).

이미지 변환 분야에서는 Semantic-Image-to-Photo Translation 기술과 Face Frontal View Generation 기술이 주요 관심을 받고 있다. Semantic-Image-to-Photo Translation은 스케치나 시맨틱 이미지를 현실적인 이미지로 바꾸는 기술이다. 반면, Face Frontal View Generation은 여러 각도의 얼굴 이미지를 기반으

로 정면 얼굴 이미지를 생성하여 얼굴 인식 시스템에 적용할 수 있다. 더 나아가, GenAI를 활용하면 이미지를 만화 스타일로 바꾸거나, 특정 연령대의 얼굴 이미지를 생성하는 것도 가능하다(Heo et al., 2021).

또한, GenAI는 프로그래밍과 코드 작성 분야에서도 큰 잠재력을 보이고 있다. ChatGPT와 같은 도구를 사용하면, 다양한 프로그래밍 언어에서 자연어를 코드로 바꾸는 것은 물론, 코드 검토, 버그 수정, 리팩토링 등의 작업도 지원한다. 이러한 기능은 프로그래머들의 작업 효율성을 크게 향상시킬 수 있다(Zheng et al., 2023; Cao et al., 2023).

## 2.2 GenAI 관련 국내외 스타트업 현황

<표 1>과 <표 2>는 각각 해외와 국내에서 GenAI를 활용한 서비스를 제공하는 사이트의 목록을 보여준다. 먼저, <표 1> 해외 GenAI Service 목록을 보면 텍스트를 이미지로, 또는 이미지를 텍스트로 변환 및 생성을 위한 DALL-E2, stable-diffusion-reimagine, prompthero 등이 있고 스크립트 입력으로 영상을 생성해주는 서비스인 pictory가 있으며, 논문이나 PDF 파일을 분석하여 요약해주는 filechat, doclime, ChatPDF, mapdeduce, askyourpdf 서비스가 있으며, 텍스트를 채팅하듯 입력하면 PPT가 생성되는 tome, beautiful도 있다. 또한 위키피디아(2023)에 따르면 AIVA(Artificial Intelligence Virtual Artist)는 프랑스의 음악학회인 SACEM이 인정한 전자 작곡 서비스로 영화음악부터 탱고에 이르기까지 다양한 스타일의 음악을 최대 3분 정도의 짧은 작곡을 생성할 수 있다. 유니콘타임즈의 김보라(2023.03.03)에 의하면 독일 AI 기반 언어 번역 기업인 DeepL은 올해 1월 1억 달러의 기업가치로 실리콘밸리 벤처 캐피탈인 아이브이피(IVP)와 베세머 벤처 파트너스(Bessemer Venture Partners), 아토미코(Atomico), 윌드이노베이션랩(WiL) 등으로부터 투자유치에 성공하며 유니콘에 도달한 새내기 스타트업이다. 이 외에도 텍스트 채팅, 음성 채팅 및 대면 채팅, 원하는 캐릭터를 선정할 수 있는 채팅으로는 Chat-GPT, chat.d-id, character가 서비스되고 있다.

<표 1> 해외 GenAI Service 목록(스타트업 포함)

분류	명칭	사이트	대표서비스
음악	AIVA	https://creators.aiva.ai	AI 음악 생성 및 음악 관련 솔루션, 원하는 스타일과 분위기를 선택, 음악생성서비스
이미지	DALL-E2	https://openai.com/product/dall-e-2	AI 이미지 생성 모델, 텍스트를 이미지로 바꿔주는 Openai서비스
	stable-diffusion-reimagine	https://clipdrop.co/stable-diffusion-reimagine	AI 이미지 생성 모델, 이미지를 비슷한 이미지로 만들어주는 서비스
	prompthero	https://prompthero.com	AI 기반의 텍스트 생성 솔루션 이미지를 키워드로 검색, 이미지 프롬프트 출력 서비스
영상	pictory	https://pictory.ai	AI 기반의 이미지 검색 및 관리 솔루션, 스크립트 입력 영상 생성

			서비스
PDF	filechat	https://www.filechat.io	파일 공유 및 협업을 위한 채팅 앱, PDF, docx, html, txt 내용 요약 서비스
	mapdeduce	https://mapdeduce.com	PDF 요약
	askyourpdf	https://askyourpdf.com	PDF 요약
	doclime	https://doclime.com	PDF 요약
	ChatPDF	https://www.chatpdf.com	AI 기반의 텍스트 분석 및 자연어 처리 솔루션, 텍스트입력 PPT 생성 서비스
마인드맵	whimsical	https://whimsical.com	다이어그램 작성 및 시각화를 위한 툴, 인공지능 마인드맵 서비스
채팅	chat.d-id	https://chat.d-id.com	음성 채팅, AI와 얼굴을 보며 대화
	Chat-GPT	https://ai.com	openai에서 만든 대화형 AI
	poe	https://poe.com	4가지 Chat-AI비교
	character	https://beta.character.ai	원하는 캐릭터와 채팅
번역	DeepL	https://www.deepl.com/translator	AI 번역
3D, 여행	journeymade	http://www.journeymade.io/	AI 기반의3D 모델링 및 시각화 솔루션, 간단한 정보 입력,여행 일정 생성 서비스
API	FutureTools	https://www.futuretools.io/	AI 기반의 텍스트 분석 및 자연어 처리 솔루션, AI 도구모음사이트
	GPT API	https://platform.openai.com/overview	GPT API.
PPT	tome	https://beta.tome.app	AI 기반의 텍스트 분석 및 자연어 처리 솔루션, 텍스트 입력 PPT 생성 서비스
	beautiful	https://www.beautiful.ai	텍스트를 입력하면 PPT 생성 서비스

<표 2>는 국내 GenAI Service 목록으로 AI 휴먼을 서비스하는 디오비스튜디오, 딥브레인AI, 이너버즈, 빔스튜디오가 있다. 그중 2022년 KOREA AI STARTUP 100에 선정된 디오비스튜디오는 2020년에 창업한 스타트업으로, GenAI 기반의 캐릭터 생성 및 AI 버추얼 휴먼 제작을 주요 사업으로 하고 있다(코리아AI스타트업, 2022). 엑스엘에이트에이아이는 2021년에 창업하여 인공지능 기반의 영상 번역 솔루션을 제공하며, 드랩은 2022년에 창업하여 AI를 활용한 광고용 이미지 생성 CDN 솔루션을 제공한다. 또한, 박준호(2023.06.27.)에 의하면 AI 튜터, AI 기반 문제 풀이 솔루션, 글쓰기 훈련 서비스 등을 제공하는 GenAI를 활용한 작문 보조 솔루션 제품을 개발한 워튼테크놀로지스는 과학기술정보통신부(이하 '과기정통부')의 2023년 글로벌 ICT 미래 유니콘 육성기업으로 선정되었다. 또한, 아티피셜소사이어티는 영어 지문과 문제를 자동 생성 서비스와 AI 기반 수학 교육 콘텐츠를 서비스하고 있다(고석용, 2023).

<표 2> 국내 GenAI Service 스타트업 목록

분류	명칭	사이트	대표 서비스
AI 휴먼	디오비스트UDIO	http://dobstudio.co.kr	AI 기반 캐릭터 생성, AI 버추얼휴먼 제작
	딥브레인AI	https://www.deepbrain.ai	AI 기반 이미지 생성, AI휴먼 서비스
	이너버즈	http://innervz.io	버추얼인플루언서제작
	빔스튜디오	https://studio-vim.com	디지털 휴먼 실시간AI 딥페이크솔루션
3D	리빌더에이아이	https://rebuilderai.com	2D 이미지를 3D모델 변환해주는AI 솔루션
	플루언트	https://fluentt.ai	3D 아바타 모션 캡처 생성
	나니아랩스	https://narnia.ai	제조업체를 위한 AI 학습용 3D 합성데이터를 자동으로 생성
	빗날덴탈스튜디오	http://bitnalcorp.com	3D 객체 구현 알고리즘을 헬스케어 산업에 적용한 치아 및 바이오 모델링 솔루션
	리콘랩스	https://reconlabs.io	AI기반 3D 모델링 솔루션을 제공
영상	웨인힐스브라이언트AI	http://www.waynehills.co	Text-To-Video, Speech-to-Video 텍스트로 영상 콘텐츠로 제작
	Vrew (VoyagerX)	https://vrew.voyagerx.com/ko	영상 편집을 위한 AI 기반 비디오 편집 플랫폼
음악	Musia	https://musia.ai/ko	자동 멜로디 생성 서비스
	슈퍼톤	https://supertone.ai	AI 기반 음악 제작 솔루션을 제공
번역	엑스엘에이 트에이아이	https://www.xl8.ai	AI 기반 영상 번역 솔루션
	보이스루	https://voithru.com	AI 동영상 번역 및 자막 생성 플랫폼
교육	뤼튼테크놀로지스	https://wrtn.tech	AI 튜터, AI 기반 문제 풀이 솔루션, 글쓰기 훈련 서비스
	아티피셜스 사이어티	http://artificial.sc	영어 지문과 문제를 자동 생성 서비스, AI 기반 수학 교육 콘텐츠
텍스트 생성	Mixo	https://app.mixo.io/ai-website-builder	웹사이트, 마케팅 자료, 소셜 미디어 콘텐츠를 포함한 다양한 목적으로 사용할 수 있는 인공 지능 기반 텍스트 생성 도구
글쓰기 첨삭	투블럭에이아이	https://twoblockai.com	글쓰기 첨삭AI 키워티
이미지, 웹툰	클레온	https://kleeon.com	AI 기반 이미지 생성, AI 기반 웹툰 제작
이미지	드랩	http://draph.ai	AI기반 이미지 생성CDN 솔루션
이미지, 배경	커넥트브릭	https://connectbrick.com	AI 배경 콘텐츠 생성 솔루션
데이터 합성	씨엔에이아이	http://www.cnai.ai	AI 합성데이터를 생성하는 플랫폼 솔루션
채팅	native	https://www.native.me	모바일에서 Chat-GPT를 이용할 수 있는 서비스 한글로 입력하면 자동으로 영어로 물어보고, 답변을 한글로 보여줌.
	AskUP	http://pf.kakao.com/_BhxkWxj	카카오톡에서 ChatGPT를 사용. b y 업스케일

모바일에서 Chat-GPT를 이용할 수 있는 서비스로 한글을 입력하면 자동으로 영어로 물어보고, 답변을 한글로 보여주는 native도 있으며, 업스케일의 AskUP은 카카오톡에서 ChatGPT를 사용할 수 있다. 이외에도 리콘랩스는 GenAI 기반 3D 모델링 솔루션을 제공하고 빗날덴탈스튜디오는 3D 객체 구현 알고리즘을 헬스케어 산업에 적용하여 치아 및 바이오 모델링 솔루션을 제공한다. 리빌더에이아이는 2D 이미지를 3D 모

델로 변화해주고 플루언트는 3D 아바타 모션 캡처 생성해주고, 나니아랩스는 제조업체를 위한 AI 학습용 3D 합성데이터를 자동으로 생성해주는 솔루션을 제공한다.

### 2.3 빅카인즈(뉴스 빅데이터 분석)

뉴스 빅데이터 분석 플랫폼인 빅카인즈는 한국언론진흥재단이 1990년부터 현재까지 54개 언론매체에서 약 7천만 건 이상 뉴스 콘텐츠의 빅데이터 화하여 검색한 뉴스의 메타데이터(언론사, 기고자, 제목 등)와 개체명(인물, 기관, 장소 등) 분석 데이터를 엑셀 파일로 제공하는 서비스이다. 1990년 1월 1일부터 2021년 9월 30일까지 노동 인권교육 관련 기사를 대상으로 시계열, 관계도, 연관어 분석을 실시하여 17대(이명박 정부), 18대(박근혜 정부), 19대(문재인 정부) 순으로 노동 인권교육 관련 보도가 급증한다는 결과를 도출하였다(이서영, 2021). 서울시의 문제 및 이슈를 도출하기 위하여 약 5년간 뉴스 기사를 수집하고 키워드분석 및 토픽 모델링을 적용하여 서울시의 문제 20개 및 주요 이슈들을 제시하였다(권민지, 2019). 키워드가 스마트 팩토리인 뉴스를 분석하여 텍스트 네트워크 분석을 하여 산업 활성화에 도움이 될 수 있는 이슈들을 도출하였다(권정흠·이호, 2019). 2019년과 2020년의 언론사 뉴스 기사들을 중 키워드를 무역으로 하여 비교 분석하였다. 이는 코로나19 사태가 무역 분야에 어떠한 영향을 주었는지 시기별 무역 관련 주요 이슈들을 도출하였다(윤희영, 2021). 소비자 연구 동향을 파악하고자 워드 클라우드 분석 방법을 적용하였다(이은영 외, 2019). 코로나19 이전과 이후에 항공사 및 승무원 관련 뉴스 기사를 비교하고 분석한 결과 항공 관련 이벤트가 발생하였을때 폭발적으로 기사 건수가 증가하였고 구조조정 관련 뉴스로 트렌드의 변화가 일어나고 있음을 확인하였다(김정오·권충훈, 2020).

2019-12-01~2020-11-30까지의 홈트레이닝에 관한 기사 수집과 분석을 통하여 홈트레이닝 트렌드를 알아보려 실시한 결과 인공지능의 기술과 등장이 홈트레이닝을 주도하는 것과 운동 종목으로는 필라테스의 선호도가 높고, 이와 관련된 운동용품의 수요가 증가함에 따라 상표출원도 영향이 있음을 알 수 있었다. 이를 통하여 대략적인 언론사들의 키워드 부여 형태를 살펴볼 수 있다는 점과 기존의 연구가 측정하지 못했던 최근 트렌드를 신속히 알아볼 수 있었다는 점에서 의미를 부여했다(지동철·김상호, 2021). 약 10개월간 코로나와 위기를 키워드로 하여 언론사의 관련 기사들을 추출하여 키워드 간의 상관계수를 파악하고, 연관어 분석을 통해 경제, 사회, 국제, 문화 각 대표 카테고리별로 COVID-19로 인해 위기 및 해결해야 할 과제들을 제시하였다(김나현·오하영, 2021).

최대 규모의 뉴스 기사 데이터베이스에 빅데이터 분석기술을 접목하여 만들어졌으며, 비정형 텍스트를 정형화된 데이터로 변경하여 분석할 수 있도록 만들었으며 이를 통해 사회현상을 연구하고 미래를 예측할 수 있는 정보로 뉴스 분석 콘

텐츠 서비스를 제공하고 있다. 빅카인즈는 텍스트 마이닝 분석 단계인 데이터수집, 전처리, 특징 추출, 모델링 및 분석, 결과해석, 시각화 및 보고서 생성하여 제공한다(김나현·오하영, 2021). 빅카인즈는 다년간의 다양한 매체의 뉴스 기사를 지역별, 사건별, 매체별 등으로 분석하고, 키워드분석 및 토픽 모델링을 적용하여 텍스트 네트워크 분석을 통해 트렌드 및 제시하는 이슈에 대해 시대상과의 관계 분석을 수행한다(김정오·권충훈, 2020).

뉴스 기사를 분석하면서 신뢰할 수 있는 객관적인 매체를 통해 트렌드나 이슈와 사건을 반영하여 시대 변화를 파악할 수 있음에 큰 의미가 있다.

### III. 연구 방법

#### 3.1 분석대상

GenAI과 스타트업 관련 뉴스 기사 데이터를 수집하기 위해 한국언론진흥재단의 빅카인즈(BIGKinds) 서비스를 이용하였다. 먼저 고려해야 할 사항은 추출하려는 데이터가 ‘생성형 인공지능’과 스타트업이라는 단어를 동시에 포함한 뉴스 기사이므로 Generative Artificial Intelligence를 표현하는 ‘생성형 AI’, ‘생성 AI’ 등과 startups를 표현하는 ‘스타트업’을 <표 3>의 검색 조건식을 참조하여("생성 AI", "생성형 AI") AND "스타트업"과 같이 조건식을 만들었다.

<표 3> 검색 조건식

번호	검색 조건식	의미
1	생성, AI	생성이나 AI 중에 1개 이상 포함된 뉴스 검색
2	"생성" OR "AI"	
3	생성 AI	생성과 AI 두 단어 모두 포함된 뉴스 검색
4	생성 + AI	
5	"생성" AND "AI"	정확히 포함된 뉴스
6	"생성 AI"	
7	생성 - AI	생성이 포함된 뉴스에서 AI가 포함되지 않은 뉴스만 검색
8	"생성" NOT "AI"	

<표 4>는 수집 기간, 수집 대상, 검색 조건식을 정리한 수집 데이터 보기표이다. 1990-01-01~2023-08-10 기간의 뉴스 기사를 수집하였으나 분석 대상으로 알맞은 의미 있는 기사 데이터는 2019년부터 도출되어 수집 기간을 2019-01-01~2023-08-10으로 설정하였다.

<표 4> 수집 데이터 보기표

구분	내용		
수집 기간	2019-01-01 ~ 2023-08-10		
단일 조건 검색어	"생성 AI"	"생성형 AI"	"스타트업"
검색 조건식	("생성 AI", "생성형 AI") AND "스타트업"		
뉴스 기사 수집 대상 (54)	전국일간지 (11)	국민일보, 경향신문, 동아일보, 조선일보, 한국일보, 서울신문, 세계일보, 문화일보, 중앙일보, 한겨레, 내일신문	
	경제일간지 (8)	매일경제, 서울경제, 한국경제, 머니투데이, 파이낸셜뉴스, 아시아경제, 헤럴드경제, 아주경제	
	지역일간지 (28)	경인일보, 부산일보, 매일신문, 중도일보, 영남일보, 중부일보, 대전일보, 충청일보, 경기일보, 중부매일, 무등일보, 경남신문, 전북도민일보, 경상일보, 충북일보, 광주일보, 국제신문, 충청투데이, 경남도민일보, 울산매일, 전남일보, 강원일보, 전북일보, 광주매일신문, 제민일보, 강원도민일보, 대구일보, 한라일보	
	방송사(5)	YTN, KBS, MBC, OBS, SBS	
	전문지(2)	전자신문, 디지털타임스	

<표 5>는 각 검색 조건식에 따라 수집된 뉴스 기사 데이터 수량과 ChatGPT 출시 전후의 데이터 건수, 인용문, 사실, 단순 동정 등의 제외 건수, 분석 대상 건수를 정리한 것이다. 검색 조건식 ("생성 AI", "생성형 AI") AND "스타트업"으로 2019-01-01~2023-08-10 기간에 수집된 뉴스 기사는 총 1,785건, 제외 대상 41건, 분석 대상 1,744건이다. ChatGPT 출시 전 2019년부터 2022년 10월까지 19건, 출시 후 2022년 11월부터 2023년 8월10일까지 총 1,766건, 제외 대상 41건, 분석 대상 1,725건이다. 검색 조건식에 의한 분석 대상 데이터의 상세 내용은 <표 5>와 같다.

<표 5> 뉴스 기사 데이터 건수

검색 조건식	기사 (건)수	ChatGPT 출시 전 (19.1~22.10)	ChatGPT 출시 후 (22.11~23.8.10)	합 계
"생성 AI"	총	91	2,522	2,613
	제외	0	61	61
	분석대상	91	2,461	2,552
"생성형 AI"	총	25	5,676	5,701
	제외	1	183	184
	분석대상	24	5,493	5,517
"스타트업"	총	147,593	71,384	218,977
	제외	3,747	934	4,681
	분석대상	143,846	36,173	180,019
("생성 AI", "생성형 AI") AND "스타트업"	총	19	1,766	1,785
	제외	0	41	41
	분석대상	19	1,725	1,744

#### 3.2 분석도구 및 분석방법

빅카인즈의 관계도 분석 시 개체명 추출 방식은 검색 결과 중 정확도 상위 100건의 뉴스 본문을 형태소 분석하여 명사 상당 어구를 추출하고 추출된 명사 상당 어구에 Structured SVM(개체명 분석 알고리즘 Support Vector Machine)을 적용하며 개체명의 관련 기사 건수를 고려해 가중치를 부여하였다

(이서영, 2021). 이와 같은 단계를 거쳐 개체명은 검색어 및 동일한 뉴스에서 추출된 다른 개체명과 선으로 연결되며 빈도에 따라 선의 굵기가 다르게 표시된다. 키워드 트렌드 분석은 피어슨 상관계수를 사용하여 검색어들의 상관관계의 정도를 나타낸다. 연관어 분석 시 사용되는 토픽랭크 알고리즘은 사용자가 입력한 키워드(A, B)와 검색 결과에서 제외하고 싶은 단어가 (C)인 경우, 쿼리 문법에 맞게 쿼리문((A or B) NOT C)으로 변형하여 검색한다. 검색 결과를 공시적 분석과 키워드 클러스터링을 하여 동적으로 상호 작용하는 시맨틱 네트워크를 생성하는 알고리즘이다. 쿼리문에 의해 제시된 키워드의 배열은 의미적 유사도 수식에 따라 결정되어 같은 기본 질의라 하더라도 사용자가 원하는 결과를 얻어내기 위해 입력한 다양한 조건(날짜, 카테고리, 매체)에 따라 시맨틱 네트워크를 구성한다(김창식 외, 2019; 박중순·김창식, 2019; 권민지, 2019). 토픽랭크 작동원리는 첫째, 주어진 용어와 관련 있는 N개의 Top 문서를 가져올 쿼리문을 검색엔진에 전송한다. 두 번째, 관련 용어 수집을 위하여 쿼리 키워드를 중심으로 하는 주변 단어들의 셋인 snippet 들을 구한 후 snippet의 용어들을 모두 추출하고, 쿼리 키워드를 중심으로 거리를 계산한 TF(weighted term frequency), 특정 단어 또는 구(phrase)가 포함된 문서의 수 DF(document frequency)를 구한다. 그리고 두 개의 이상의 요소가 같은 문맥이나 구조 내에서 함께 나타나는 co-occurrence 용어 매트릭스 생성한다. M개의 Top 용어를 선택하기 위해 복합 명사에 가중치를 주게 된다. 그 결과 M개의 Top 용어를 이용하여, 용어 벡터를 구성하게 되고 M개의 Top 용어 벡터에 대한 클러스터링을 수행한다(이서영, 2021).

빅카인즈에서 검색한 뉴스 기사의 본문에서 단순 숫자나 이메일 주소, 밤, 낮 등 시간을 뜻하는 단어를 제외한 정제된 데이터를 키워드로 제공해 주기 때문에 텍스트 전처리 과정이 단축되었다. 이러한 메타정보, 분석 결과 등을 엑셀 파일로 내려받아 Python, SPSS 등의 분석 프로그램을 통해 추가 분석할 수 있다. 본 연구에서는 이차적으로 재정제를 위하여 Python의 KoNLPy를 이용하여 빈도수가 현격히 높지만 본 연구에서는 의미 없는 '고도화', '수년간', '다년간', '마지막' 등의 단어들을 제거하였고 'ChatGPT', '챗GPT', '챗지피티' 등과 같이 'ChatGPT'를 의미하는 단어는 하나로 만드는 전처리 과정을 진행하였다.

본 연구에서는 재정제된 데이터를 사용하여 키워드 네트워크 분석을 통해 네트워크 시각화를 진행하였다. 네트워크 시각화를 위해서 Python의 networkx 라이브러리와 Gephi 같은 시각화 도구를 활용하였다. 이를 통해 중심성이 높은 키워드, 키워드 클러스터 등의 정보를 파악할 수 있다. 또한, 토픽랭크 알고리즘을 기반으로 단어의 중요도를 가중치로 계산하여 연관어의 빈도수를 워드 클라우드로 시각화하였다. 이를 통해 GenAI와 스타트업의 연관관계를 직관적으로 파악할 수 있었다.

## IV. 연구 결과

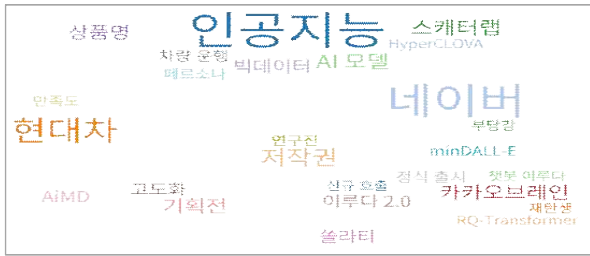
### 4.1 ChatGPT 출시 전 뉴스 기사 분석

ChatGPT 출시 전 2019년~2022년 10월 기간에 GenAI와 스타트업에 대한 기사 빈도를 분석한 결과는 <표 6>과 같다. 단일 검색어들("생성 AI", "생성형 AI")과 검색 조건식("생성 AI", "생성형 AI") AND "스타트업")에 의한 뉴스 기사 빈도를 보면 공통으로 '인공지능', '이루다', 'GAN', '네이버 하이퍼클로' 등이 고빈도로 나타나고 있다. "생성 AI"로 검색 시 키워드 중 스캐터랩과 인공지능, 이루다, AI 모델의 빈도가 높게 나오는데 이는 스캐터랩에서 개발한 이루다는 AI 챗봇으로, 사용자와의 대화 데이터를 기반으로 학습하며, 감정, 성격, 관계 등을 분석하여 대화를 진행하는 형식으로 한때 큰 인기를 끌었으나, 개인정보 유출과 차별, 혐오 발언 관련된 논란으로 서비스를 잠정 중단했었다. 그 후 약 1년 9개월 만에 재출시되는 과정에서 다시금 개인정보 유출, 보안 등의 이슈가 제기되었기 때문이다. "생성형 AI"로 검색하면 LG, AI 연구원, AI 엑사원, 파슨스, GAN, 메타, 메타버스, 메타의 CAIRaoke, 딥페이크 등의 빈도가 높게 나오는 것은 초거대 AI '엑사원'을 기반으로 LG와 파슨스가 GenAI 기술과 방법론을 공동 연구를 발표했기 때문이다. 또한, 메타는 프로젝트 CAIRaoke를 미래의 AI 어시스턴트로 발표하며 양방향 대화형 AI의 구축을 선보였다. 조건식("생성 AI", "생성형 AI") AND "스타트업"로 검색을 진행하면 스캐터랩이나, 플리토와 같은 스타트업이 키워드로 표출되지만 이런 경우 스타트업의 고객사로 네이버나 카카오 등이 소개되어 오히려 네이버관련 키워드의 빈도수가 높은 것을 알 수 있다. 또한 애플의 시리가 고객과 대화하는 내용을 시리뿐만 아니라 애플의 협력사 직원까지 대화를 엿듣는 일이 발생하는 문제가 있었다. 이와 같이 2019년부터 2022년 10월 사이의 뉴스 기사를 보면 GenAI와 관련된 스타트업에 대한 언론의 관심 매우 약하고 오히려 우려의 소리가 크다는 것을 파악할 수 있다. 반면, 초거대기업이나 메타버스, AI 반도체와 LLM에 대한 관심이 높고 TOP10에는 속하지 못한 키워드에는 'fake', '딥페이크', '가짜 뉴스', '얼굴 합성' 등 GAN에 대한 우려와 인식이 좋지 않음을 알 수 있다.

<표 6> ChatGPT 출시 전 GenAI 관련 뉴스 기사 빈도 분석

no	"생성 AI"		"생성형 AI"		("생성 AI", "생성형 AI") AND "스타트업"	
	키워드	빈도	키워드	빈도	키워드	빈도
1	인공지능	195	LG-파슨스	855	네이버클라우드	103
2	이루다 2.0	71	메타	211	빅데이터	18
3	현대차	52	GAN	183	스캐터랩	18
4	스캐터랩	50	AI 엑사원	150	플리토	14
5	AI 모델	47	CAIRacke	34	GPT-	9
6	카카오브레인	44	딥페이크	26	스케	7
7	빅데이터	35	인공지능	24	이스트소프트	6
8	HyperCLOVA	35	연구팀	18	시리	6
9	솔라티	25	AI 반도체	16	한들	5
10	AIMD	25	KAIST	5	페리지	4

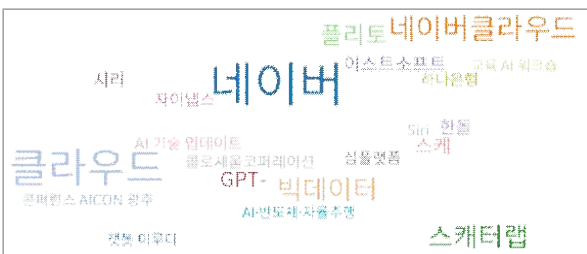
<그림 1> 부터 <그림 3>는 ChatGPT 출시 전 검색어별 관련 키워드를 빈도수 100회 이상으로 설정하여 워드클라우드로 시각화를 구현하였다. 이를 통해 ChatGPT 출시 전에는 GenAI에 대한 관심과 사업화가 스타트업들보다 주로 전문가와 빅테크 기업에서 더 강조됨을 알 수 있다.



<그림 1> ChatGPT 출시 전 "생성 AI"



<그림 2> ChatGPT 출시 전 "생성형 AI"



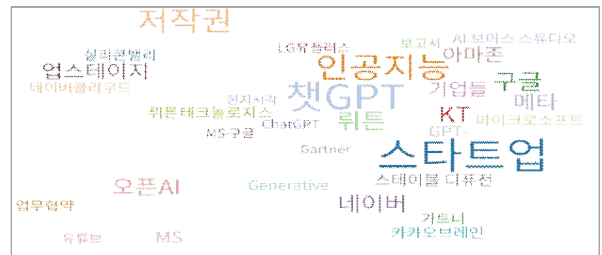
<그림 3> ChatGPT 출시 전 ("생성 AI", "생성형 AI") AND "스타트업"

## 4.2 ChatGPT 출시 후 뉴스 기사 분석

ChatGPT 출시 후 2022년 11월~2023년 8월 기간에 GenAI와 스타트업에 대한 기사 빈도를 분석한 결과는 <표 7>과 같다. 단일 검색어들과 검색 조건식(GenAI와 스타트업)에 의한 뉴스 기사 빈도를 보면 높은 빈도의 기사는 'ChatGPT', '스타트업', '뤼튼테크놀로지', '엔비디아', '가트너' 등이다. <표 5>의 ChatGPT 출시 전후 뉴스 기사 건수를 비교해보면 "생성 AI"는 27배, "생성형 AI"는 229배, ("생성 AI", "생성형 AI") AND "스타트업"은 91배 증가했음을 알 수 있다. ChatGPT 출시 전에 비하면 전문가뿐만 아니라 일반 대중들도 GenAI와 이를 서비스하는 스타트업체에 관심이 높은 것을 알 수 있다.

<표 7> ChatGPT 출시 후 GenAI 관련 뉴스 기사 빈도 분석

no	"생성 AI"		"생성형 AI"		("생성 AI", "생성형 AI") AND "스타트업"	
	키워드	빈도	키워드	빈도	키워드	빈도
1	가트너	428	챗GPT	196	챗GPT	142
2	스타트업	191	AWS	124	칼로	106
3	MS	209	인공지능	106	AWS	87
4	인공지능	132	구글	54	KT	68
5	저작권	124	SK	38	구글	61
6	구글	82	엔비디아	33	언어모델	45
7	오픈AI	70	ETF	25	챗봇	44
8	메타	62	AI 시대	24	실리콘밸리	43
9	뤼튼테크놀로지	61	보고서	23	업스테이지	43
10	업스테이지	60	머신러닝	21	마이크로소프트	28



<그림 4> ChatGPT 출시 후 "생성 AI"

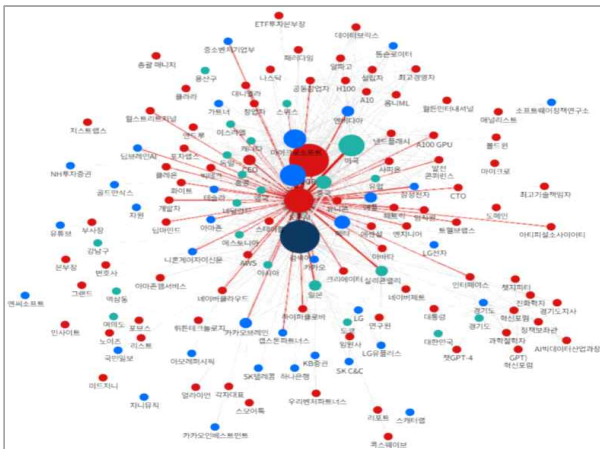


<그림 5> ChatGPT 출시 후 "생성형 AI"



<그림 6> ChatGPT 출시 후 ("생성 AI", "생성형 AI") AND "스타트업"

<그림 4>부터 <그림 6>까지는 ChatGPT 출시 후의 검색어별 관련 키워드를 빈도수 500회 이상으로 설정하여 wordcloud 모델로 시각화한 내용이다. 이를 통해 ChatGPT 출시 전과 비교했을 때, 스타트업 분야에서 GenAI에 대한 관심이 많이 증가한 것을 확인할 수 있다. 또한, <그림 7>은 관계도 분석을 네트워크 형태로 시각화한 결과다. 이를 통해 유니콘 기업인 OpenAI가 직접 투자한 국내 스타트업인 스피키지랩스와 디스크립트를 확인할 수 있었으며, GenAI를 활용하여 혁신적인 상품화에 성공한 스캐터랩, 콕스웨이브, 스모어톡과 같은 스타트업들도 확인할 수 있다.



<그림 7> ("생성 AI", "생성형 AI") AND "스타트업" 관계도 분석

## V. 결론

### 5.1 연구 결과

본 연구는 GenAI(Generative Artificial Intelligence) 기술 분야의 이슈와 트렌드 동향을 국내 뉴스 기사를 통해 분석하고, 이를 바탕으로 GenAI 기술을 활용한 기술 창업 분야의 연구 방향을 탐색하는 것을 목표로 설정하였다. 이를 위해 빅카인즈(2023) 뉴스 기사 빅데이터 분석 플랫폼을 활용하여 1990년 1월부터 2023년 8월 10일까지의 기간 동안 “생성 AI”, “생성형 AI”, “스타트업” 각각 검색한 뉴스 기사와 (“생성 AI” ,

“생성형 AI”) AND “스타트업”을 여러 조건을 넣어 뉴스 기사를 검색하였다. 그러나, 유의미한 뉴스 기사를 수집하기 위해서 시작 시기를 바꾸어야 했다. 그리고 GenAI 기술 분야에서 큰 이변이 발생했던 시기를 기점으로 나누어 기사 내용을 중심으로 동향을 파악하였다. 먼저 시작 시기는 GenAI와 스타트업이 뉴스 기사에 동시에 나오기 시작한 시기이므로 연구의 시작점으로 판단하였다. GenAI 분야에서 획기적인 일이 2022년 11월을 기점으로 발생한 것을 뉴스 기사의 데이터양으로 파악이 되어 ChatGPT 출시 전과 출시 후로 나누어 비교 분석하였다. 출시 전에는 GenAI에 대한 우려와 부정적인 관점이 우세하였으며, 출시 후에는 직업 대체 우려와 함께 GenAI를 실생활에 적용하는 방법, 비서 역할, 스타트업 관련 기사 등 긍정적인 측면이 강조되는 것을 확인하였다.

관계도 분석 결과에서 가장 많이 표출된 키워드는 '창업자', '인사이드', '인스타그램', '코파일럿', '정보통신정책연구원', '경기도', '교육부', '카카오브레인', '소프트웨어 정책연구원' 등이다. 전체 기간 건수에 따른 연관어 분석 결과에서는 '챗GPT', '스타트업', '구글', '혁신포럼'이 가중치와 기사 건수의 높은 비중을 차지하는 주요 키워드로 나타났다. 이는 관계도 분석 결과와 유사한 결과를 보여주었다. ChatGPT 서비스 이후 키워드와 연관어의 빈도를 집중적으로 분석한 결과 2022년 11월부터 12월까지 AI 글쓰기 퓌튼을 개발한 스타트업 퓌튼테크놀로지스, 비주얼 AI 기반 비브스튜디오스, 반도체 스타트업인 리벨리온(오수연, 2023), 2023년 1월부터 3월까지 인공지능 경진대회, 경기도, 스타트업, 개척자, 2023년 4월부터는 경진대회, 창의적 아이디어, 저작권, HelloAI 대표 등이 언론매체에 많이 언급되고, 독자들에게 많은 주목을 받고 있음을 알 수 있다. GenAI 기술은 스타트업의 비즈니스 모델 혁신과 창업 활동을 촉진하는 주요한 동력으로 작용하며 GenAI기술 분야의 이슈와 트렌드 동향이 기술창업 생태계와 밀접한 관련이 있음을 확인하였다. 스타트업 창업자 및 기업 리더들에게 GenAI 기술을 통한 혁신적인 비즈니스 전략을 구축하는데 큰 도움을 제공할 것으로 기대된다.

### 5.2 연구의 한계와 제언

본 연구는 몇 가지 한계를 지니고 있다. 첫째, 데이터 수집 범위가 제한적이다. 본 연구에서 사용된 데이터는 빅카인즈를 통해 얻은 국내 54개 언론매체의 뉴스 기사에 국한되어 있다. 이 때문에 국내 다른 전문 신문사 자료는 포함되지 않았으며, 따라서 데이터의 완전성과 다양성에 한계가 있을 수 있고, 중요한 정보가 누락될 가능성이 있다.

둘째, 본 연구의 분석 방법론은 주로 네트워크 관계도 분석, 키워드 트렌드 분석, 연관어 분석 등에 초점을 맞추었다. 그러나 GenAI 기술과 관련된 스타트업 간의 관계나 기술 분야의 상관관계 분석, 기술적 측면의 깊은 탐구를 위한 다양한 방법론은 적용되지 않았다. 이러한 한계를 극복하기 위해 향



후 연구에서는 보다 다양한 분석 방법을 활용하여 깊이 있는 연구를 수행할 필요가 있다. 특히, 해외 GenAI 기술 트렌드와 스타트업 간의 관계에 대한 포괄적인 분석이 요구된다. 이를 통해 국제적인 시각에서 GenAI 기술의 발전 동향과 기술 창업 분야의 전략적 방향을 더욱 명확하게 파악하는 것이 중요하다.

향후 연구 방향에 대해 다음과 같이 제안할 수 있다. 첫째, 국내외 다양한 데이터를 활용해 GenAI 기술 분야의 최신 이슈와 트렌드를 분석하여, 국제적인 시각에서 GenAI 기술의 발전 동향을 파악하는 연구가 필요하다. 다양한 분석 방법론을 적용해 더욱 상세한 인사이트를 도출하고, 기술 창업 분야에서의 전략적 접근법을 깊게 탐구해야 한다. 둘째, 기술 창업 생태계와의 긴밀한 협력을 통해 GenAI 기술의 실질적인 응용 및 현장 적용 가능성을 탐색하는 것이 중요하다. 이를 위해 기술 창업자와의 협력 강화, 기술 지원 정책 개발, 그리고 다양한 이해관계자와의 협력을 통한 새로운 비즈니스 모델 및 사회적 적용 전략을 연구해야 한다. 이러한 연구는 GenAI 기술 분야에 대한 이해를 확대하고, 기술 창업 생태계의 성장과 사회적 적용을 촉진하는 데 크게 기여할 것으로 기대된다.

## REFERENCE

고석용(2023.05.11.). "AI로 영어 지문·문제 생성" 아티피셜소사이 어티, 서비스 공개. 머니투데이, Retrieved from <https://news.mt.co.kr/mtview.php?no=2023013112281246955>.

권민지(2019). 토픽 모델링 기반 뉴스기사 분석을 통한 서울시 이슈 도출. *한국방송 미디어공학회: 학술대회논문집*, 11-13, 한국방송·미디어공학회.

권정흠·이호(2019). 스마트팩토리의 트렌드 및 인식 연구: 뉴스 네트워크 분석을 중심으로. *한국지식정보기술학회논문지*, 14(6), 605-614.

김나현·오하영(2021). 대규모 빅데이터 분석 기반 COVID-19 Pandemic 분석결과. *한국정보통신학회논문지*, 25(4), 495-500.

김보라(2023.03.03). "오픈AI에 밀리지 않아" 獨 AI 번역 '다크호스' '달열' 유니콘타임즈, Retrieved from <http://www.unicomtimes.kr/news/articleView.html?idxno=530>

김정오·권충훈(2020). 코로나19 이전(2019) 과 이후(2020)에 항공사와 직원 관련 언론사 뉴스 기사 비교분석. *한국컴퓨터정보학회논문지*, 25(7), 167-173.

김주희·정애린·김선우(2022). 국내 유니콘 기업군의 실태분석과 특정에 관한 연구. *벤처창업연구*, 17(1), 63-77.

김창식·김남규·곽기영(2019). 머신러닝 및 딥러닝 연구동향 분석: 토픽모델링을 중심으로. *디지털산업정보학회논문지*, 15(2), 19-28.

박수현(2023.05.19.). "'챗GPT'보다 한국어 6500배 잘해"... 네이버 '하이퍼클로바X' 들고 대화형 AI 전쟁 뛰어든다. 조선비즈, Retrieved from <https://biz.chosun.com/it-science/ict/2023/02/27/6Y46R5SPXFFYBCJGXG7M6ZU62E/>

박중순·김창식.(2019). 빅데이터 연구동향 분석: 토픽 모델링을 중심으로. *디지털산업정보학회논문지*, 15(1), 1-7.

박준호(2023.06.27.). 과기정통부, ICT 유니콘 육성사업에 린트 등

15개사 선정 전자신문, Retrieved from <https://www.etnews.com/20230627000058>

변정호·권용주(2023). 생성적 인공지능의 교육적 활용 방안 탐색. *한국교원대학교 뇌·AI기반교육연구소*, 13(1), 1-17.

빅인즈(2023). *생성형AI*, Retrieved 2023.05.19. from <https://www.bigkinds.or.kr/v2/news/index.do/2023.02.23>.

성현희(2022.07.19). 네이버 '글로벌 스튜디오', CBT에 스타트업 100곳 참여...AI 대중화 가속. 전자신문, Retrieved 2023.05.19. from <https://www.etnews.com/20220719000093>.

오수연(2023.04.04). '한국의 엔터테인먼트 리벨리온, AI반도체 '아름 앞세워 도약' 아시아경제, Retrieved 2023.05.19. from <https://www.asiae.co.kr/article/2023040409194345764>.

위키피디아(2023). *AIVA*. Retrieved 2023.06.10, <https://en.wikipedia.org/wiki/AIVA>.

윤성만·성창수(2023). 공동창업의 단계별 성공요인에 관한 연구: 기회형 창업기업 사례를 중심으로. *벤처창업연구*, 18(1), 141-158.

윤희영(2021). 뉴스기사 빅데이터를 활용한 무역이슈 변화분석: 2011~2020. *통상정보연구*, 23(1), 113-136.

이서영(2021). 뉴스 기사 빅데이터 분석을 통한 노동인권교육 담론 분석. *사회과학연구*, 28(4), 33-45.

이은영·주경희·이두희(2019). 워드 클라우드 기법을 이용한 최근 소비자학 연구 트렌드 분석. *상품학연구*, 37(1), 1-7.

정승원·손민재·황인준(2019). 지능형 교육 시스템의 학습자 분류를 위한 Variational Auto-Encoder 기반 준지도학습 기법. *한국멀티미디어학회지*, 22(11), 1251-1258.

지동철·김상호(2021). 신문 빅데이터와 키워드 분석을 이용한 핫트레이닝 트렌드 분석. *한국융합학회논문지*, 12(6), 233-239.

최준영(2023.03.03.). 중기부 '생성형 AI 스타트업과 정책지원 논의' 문화일보, Retrieved 2023.05.09. from <https://www.nunhwa.com/news/view.html?no=2023030301071507025001>.

코리아AI스타트업(2022). *2022 KOREA AI STARTUP 100 선정* Retrieved 2023.05.19. from <https://aistartup100.co.kr/page/s1/s3.php>

한규동·전병훈(2022). 스타트업 실태 및 육성방안에 관한 연구: 성남시 스타트업을 중심으로. *벤처창업연구*, 17(5), 67-80.

Aydın, Ö., & Karaarslan, E.(2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. *Available at SSRN* 4308687.

BIG KINDS(2023). *GenerativeAI*. Retrieved 2023.05.19. from <https://www.bigkinds.or.kr/v2/news/index.do/2023.02.23>.

Bommarito II, M., & Katz, D. M.(2022). GPT takes the bar exam. arXiv preprint arXiv:2212.14402.

Byeon, J. H., & Kwon, Y. J.(2023). An Investigation of Generative AI in Educational Application: Focusing on the Usage of ChatGPT for Learning Biology. *Brain, Digital, & Learning*, 13(1), 1-17.

Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L.(2023). *A comprehensive survey of ai-generated content (aigc): A history of generative ai from gan to chatgpt*. arXiv preprint arXiv:2303.04226.

Chi, D. C., & Kim, S. H.(2021). Home training trend analysis using newspaper big data and keyword analysis. *Journal of the Korea Convergence Society*, 12(6), 233-239.

Cho, J., Zala, A., & Bansal, M.(2022). *DALL-EVAL: Probing the Reasoning Skills and Social Biases of Text-to-Image Generative Models*. arXiv preprint arXiv: 2202.04053.

Choi, J. Y.(2023, March 3). *Ministry of SMEs discusses policy support with generative AI startups*. *Munhwa*

- Ilbo*. Retrieved May 9, 2023, from <https://www.munhwa.com/news/view.html?no=2023030301071507025001>.
- Forti, S., Breitenbücher, U., & Soldani, J.(2022). Trending topics in software engineering. *ACM SIGSOFT Software Engineering Notes*, 47(3), 20-21.
- Go, S. Y.(2023, January 13). "AI generates English fingerprints and questions," *Artificial Society, Service unveiled. MONEY TODAY*(2023). Retrieved 2023.05.11 from <https://news.nt.co.kr/ntview.php?no=2023013112281246955>.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y.(2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
- Han, K. D., & Jeon, B. H.(2022). A Study on the Status of Startups and Their Nurturing Plans: Focusing on Startups in Seongnam City. *Asia-Pacific Journal of Business Venturing and Entrepreneurship*, 17(5), 67-80.
- Hataya, R., Bao, H., & Arai, H.(2022). Will Large-scale Generative Models Corrupt Future Datasets?. arXiv preprint arXiv:2211.08095.
- Heo, Y. J., Kim, B. G., & Roy, P. P.(2021). Frontal face generation algorithm from multi-view images based on generative adversarial network. *Journal of Multimedia Information System*, 8(2), 85-92.
- Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A.(2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125-1134.
- Jung, S. W., Son, M. J., & Hwang, E. J.(2019). Variational Auto-Encoder Based Semi-supervised Learning Scheme for Learner Classification in Intelligent Tutoring System. *Journal of Korea Multimedia Society*, 22(11), 1251-1258.
- Kim, B. R.(2023, March 3). "Not being outdone by OpenAI" German AI translation 'Dark Horse'-'DeepL'. *Unicorn Times*, <http://www.unicorntimes.kr/news/articleView.html?idxno=530>
- Kim, C. S., Kim, N. G., & Kwahk, K. Y.(2019). Research trends analysis of machine learning and deep learning: Focused on the topic modeling. *Journal of Korea Society of Digital Industry and Information Management*, 15(2), 19-28.
- Kim, J. H., Jung, A. R., & Kim, S. W.(2022). A Study on Current Trends and Characteristics of Korean Unicorn Group. *Asia-Pacific Journal of Business Venturing and Entrepreneurship*, 17(1), 63-77.
- Kim, J. O., & Kwon, C. H.(2020). Comparative Analysis of News Articles related to Airlines and Staff the Previous Corona19(2019) and After Corona19(2020). *Journal of The Korea Society of Computer and Information*, 25(7), 167-173.
- Kim, N. H., & Oh, H. Y.(2021). Analysis of COVID-19 Pandemic based on Massive Big Data Analysis. *Journal of the Korea Institute of Information and Communication Engineering*, 25(4), 495-500.
- KoreaAISTartup(2022). *2022 Korea AI Startup 100 Finalists*. Retrieved 2023.05.19. from <https://aistartuptop100.co.kr/page/sl/s3.php>
- Kwon, J. H., & Lee, H.(2019). A study on trends and perceptions in smart factory: news network analysis. *Journal of Knowledge Information Technology and Systems*, 14(6), 605-614.
- Kwon, M. J.(2019). Identifying Seoul city issues based on topic modeling of news article. In Proceedings of the Korean Society of Broadcast Engineers Conference. 11-13, *The Korean Institute of Broadcast and Media Engineers*.
- Lee, E. Y., Chu, K. H., & Lee, D. H.(2019). A study on recent trend analysis in consumer research applying Word Cloud. *Journal of product Research*, 37(1), 1-7.
- Lee, S. Y.(2021). A Discourse Analysis on Human Rights Education on Labor through Big Data Analysis of News Articles. *Journal of Social Studies Education Research*, 28(4), 33-45.
- Microsoft(2023). *Microsoft combines mega AI into all its workplace productivity tools, Copilot*. Retrieved 2023.05.19. from <https://news.microsoft.com/ko-kr/2023/03/17/introducing-microsoft-365-copilot/>.
- Oh, S. Y.(2023, April 4). 'Korea's NVIDIA' Rebellion, leaps forward with AI semiconductor 'Atom'. *Asia Economy*. Retrieved May 19, 2023, from <https://www.asiae.co.kr/article/2023040409194345764>.
- OpenAI(2022). *ChatGPT*. Retrieved 2022.12.10. from <https://openai.com>.
- Park, J. H.(2023, June 27). *Ministry of Science and ICT selects 15 companies including Luton for ICT unicorn fostering project*. ET News, <https://www.etnews.com/20230627000058>
- Park, J. S., & Kim, C. S.(2019). Research trends analysis of big data: focused on the topic modeling. *Journal of Korea Society of Digital Industry and Information Management*, 15(1), 1-7.
- Park, S. H.(2023, May 19). "“ChatGPT” is 6500 times better in Korean”... Naver jumps into the conversational AI war with 'HyperClovaX'. *Chosun Biz*, <https://biz.chosun.com/it-science/ict/2023/02/27/6Y46R5SPXFFYBCJGXG7M6ZU62E/>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I.(2019). *Language models are unsupervised multitask learners*. OpenAI blog, 1(8), 9.
- Reddy, M. D. M., Basha, M. S. M., Hari, M. M. C., & Penchalaiah, M. N.(2021). Dall-e: Creating images from text. *UGC Care Group I Journal*, 8(14), 71-75.
- Sung, H. H.(2022, July 19). *Naver 'Clova Studio', 100 startups participate in CBT... Accelerating 'AI popularization'*. ET News, Retrieved May 19, 2023, from <https://www.etnews.com/20220719000093>.
- Suzuki, K., Roseboom, W., Schwartzman, D. J., & Seth, A. K.(2017). *A Deep-Dream Virtual Reality Platform for Studying Altered Perceptual Phenomenology*. Scientific reports, 7(1), 15982. <https://doi.org/10.1038/s41598-017-16316-2>
- Transformer(2022). *Transformer*. <https://blogs.nvidia.co.kr/2022/04/01/what-is-a-transformer-model/2022.05.18>
- Wikipedia(2023). *AIVA*. Retrieved June 10, 2023, from <https://en.wikipedia.org/wiki/AIVA>.
- Wolf, M. J., Miller, K., & Grodzinsky, F. S.(2017). Why we

should have seen that coming: comments on Microsoft's " Tay" experiment," and wider implications. *Acm Sigcas Computers and Society*, 47(3), 54-64.

- Wu, J., Gan, W., Chen, Z., Wan, S., & Lin, H.(2023). *Ai-generated content(aigc): A survey*. arXiv preprint arXiv:2304.06632.
- Yoon, H. Y.(2021). An analysis of trends in international trade issues using big data-based news articles: From 2011 to 2020. *International Commerce and Information Review*, 23(1), 113-136.
- Yun, S. M., & Sung, C. S.(2023). A Study on the Success Factors of Co-Founding Start-up by Step: Focusing on the Case of Opportunity-type Start-up. *Asia-Pacific Journal of Business Venturing and Entrepreneurship*, 18(1), 141-158.
- Zheng, Q., Xia, X., Zou, X., Dong, Y., Wang, S., Xue, Y., ... & Tang, J.(2023). *Codegex: A pre-trained model for code generation with multilingual evaluations on humaneval-x*. arXiv preprint arXiv:2303.17568.

# GenAI(Generative Artificial Intelligence) Technology Trend Analysis Using Bigkinds: ChatGPT Emergence and Startup Impact Assessment

Lee, Hyun Ju\*  
Sung, Chang Soo\*\*  
Jeon, Byung Hoon\*\*\*

## Abstract

In the field of technology entrepreneurship and startups, the development of Artificial Intelligence(AI) has emerged as a key topic for business model innovation. As a result, venture firms are making various efforts centered on AI to secure competitiveness(Kim & Geum, 2023). The purpose of this study is to analyze the relationship between the development of GenAI technology and the startup ecosystem by analyzing domestic news articles to identify trends in the technology startup field. Using BIG Kinds, this study examined the changes in GenAI-related news articles, major issues, and trends in Korean news articles from 1990 to August 10, 2023, focusing on the emergence of ChatGPT before and after, and visualized the relevance through network analysis and keyword visualization.

The results of the study showed that the mention of GenAI gradually increased in the articles from 2017 to 2023. In particular, OpenAI's ChatGPT service based on GPT-3.5 was highlighted as a major issue, indicating the popularization of language model-based GenAI technologies such as OpenAI's DALL-E, Google's MusicLM, and VoyagerX's Vrew. This proves the usefulness of GenAI in various fields, and since the launch of ChatGPT, Korean companies have been actively developing Korean language models. Startups such as Ritten Technologies are also utilizing GenAI to expand their scope in the technology startup field.

This study confirms the connection between GenAI technology and startup entrepreneurship activities, which suggests that it can support the construction of innovative business strategies, and is expected to continue to shape the development of GenAI technology and the growth of the startup ecosystem. Further research is needed to explore international trends, the utilization of various analysis methods, and the possibility of applying GenAI in the real world. These efforts are expected to contribute to the development of GenAI technology and the growth of the startup ecosystem.

*KeyWords: GenAI(Generative Artificial Intelligence), BIG Kinds, Startup, ChatGPT, Tech Trends*

---

\* First Author, Ph.D. Candidate, Dongguk University, Department of Technology Startup, hj.lee@maxted.kr

\*\* Corresponding Author, Professor, Department of Technology Entrepreneurship, Dongguk University, redsun44@dongguk.edu

\*\*\* Corresponding Author, Professor, Department of Technology Entrepreneurship, Dongguk University, bhjeon@dongguk.edu