

Deloitte.



생성형 AI-2편

새로운 인공지능(AI) 개척지

생성형 AI의 비즈니스 모델 연구

딜로이트 AI 연구소(Deloitte AI Institute)

2023년 09월
Deloitte Insights

Download on the
App Store

GET IT ON
Google Play



'딜로이트 인사이트' 앱에서
경영·산업 트렌드를 만나보세요!

목차

개요	생성형 AI, 신기원을 열다	03
01	생성형 AI의 '마술' 해독하기	05
02	소비자와 기업의 생성형 AI 활용사례	07
03	생성형 AI의 상업성과 시장경쟁 구도	13
04	생성형 AI 도입 및 상업화 전략	20

개요

생성형 AI, 신기원을 열다

챗지피티(ChatGPT), 달리(DALL·E), 렌사(Lensa) 등 소비자용 애플리케이션이 연이어 출시된 2022년은 인공지능(AI) 분야의 분수령이었다. 생성형 AI(Generative AI)를 일반 대중이 사용할 수 있게 되면서 AI 세계의 패러다임이 바뀌었다. 현 시대 AI는 패턴 포착과 규칙 기반 기술로 데이터를 분석하고 예측 능력을 발휘하지만, 딥러닝 모델인 '트랜스포머'(transformer, 변환기) 아키텍처가 발전하면서 생성형 AI라는 새로운 세계가 펼쳐지고 있다. **생성형 AI는 사전에 훈련을 받은 데이터에 기반해 새로운 데이터를 만들어 냈으로써 사람의 창작 과정을 흉내 낸다. AI가 단순한 실험기술에서 (잠재적으로) 사람의 협업자(co-passenger)로 승격한 것이다.** 시장조사기관 가트너(Gartner)는 이르면 2025년까지 세상에 존재하는 모든 데이터의 10% 이상을 AI가 생성할 것이라고¹, 사람-기계 '공생의 시대'(Age of With™)를 예고했다.

생성형 AI의 초기라 할 수 있는 지금 출시된 제품은 대부분 소비자용으로, 이것만으로도 신기원을 열었다고 할 수 있다. 하지만 생성형 AI는 조만간 사람에 버금가는 상황인식 능력과 의사결정 능력까지 갖춰, 업무 흐름에 자연스럽게 녹아들며 기업 운영 프로세스에도 혁신을 가져다줄 것이다. 현재 그러한 미래 모습을 엿볼 수 있는 솔루션이 이미 등장했다. 구글(Google) 콘택트센터 AI(Contact Center AI, CCAI)는 자연어 기반 고객 서비스 소통을 지원하는 솔루션이며², 엔비디아(NVIDIA)의 산업 특성화 솔루션인 바이오네모(BioNeMo)는

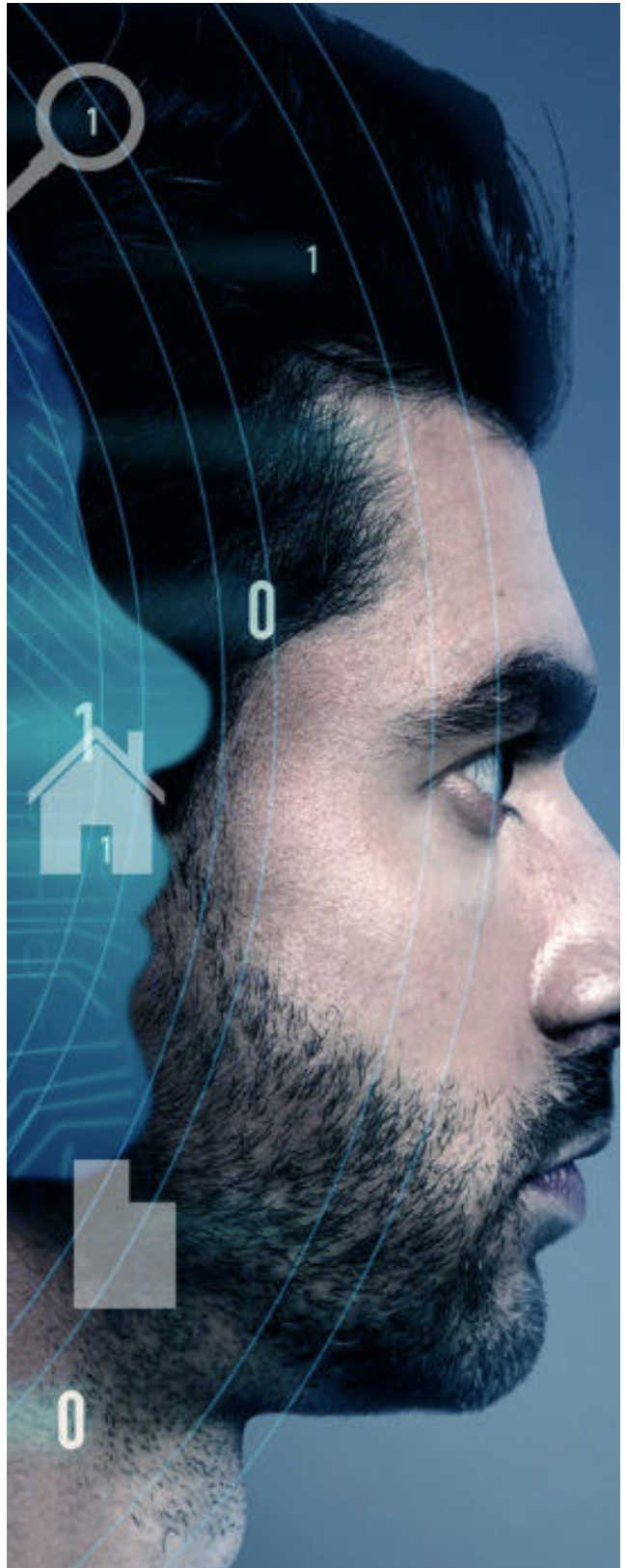
신약 개발 분야의 생성형 AI를 위한 클라우드 서비스다.³ 이러한 잠재력이 점차 구체화되면서, 벤처캐피털(VC) 및 인수합병(M&A) 시장의 전통적 플레이어들뿐 아니라 생태계 파트너십의 새로운 플레이어들까지 생성형 AI가 가져다줄 수익에 눈독을 들이고 있다. 전 세계 VC들이 생성형 AI에 투자한 액수가 2022년 한 해에만 미화 20억 달러를 넘었고⁴, 첨단기술 선도 기업들도 대규모 투자에 나섰다. 마이크로소프트(Microsoft)는 챗지피티 개발사 오픈AI(Open AI)에 100억 달러를 투자했고⁵, 구글은 전 오픈AI 직원들이 공동설립한 스타트업 앤스로픽(Anthropic)에 3억 달러를 투자했다.⁶

소비자와 기업들이 생성형 AI 도입 시 얻게 될 광범위한 잠재력과 가치에 눈을 뜨면서, 실험적 활용사례와 소비자 활용 사례가 갈수록 더욱 강력한 탄력을 얻고 있으며, 곧 기업용 활용 사례도 부쩍 늘어날 것이다. **현재 언론 보도는 소비자용 활용 사례에 초점을 맞추고 있지만, 생성형 AI가 제시할 기회는 이보다 훨씬 광범위하고 일부는 이미 실현되고 있다.** 그럼에도 불구하고 개인과 기업이 어떻게 생성형 AI를 활용해서 효율성을 강화하고, 제품을 개선하고, 새로운 경험을 창출하고, 운영 혁신을 꾀할 수 있는지 그 구체적인 방법을 묻는다면, 실질적인 답을 줄 수 있는 이들은 많지 않다. 또한 생성형 AI의 상업화도 이제 겨우 시작 단계이며, 이를 활용한 지속 가능한 사업모델도 지금으로서는 많이 찾아보기 힘들다.

또한 생성형 AI는 현재 요람기를 거치고 있으며, 초기 확산 단계에 있는 모든 기술이 그렇듯 여러가지 위험을 수반한다. 가장 시급하게 해결해야 할 위험은 개인 정보보호와 데이터 보안, AI의 편향성 관리, AI 창작물의 투명성과 추적가능성, 지식재산권, 평등접근권, 특히 생성형 AI로 일자리 상실 위험이 커진 이들에 대한 고려 등의 문제다. 이에 따라 생성형 AI의 도입과 활용에 관여하는 모든 이들은 상업화, 규제, 윤리, 공동창작, 심지어 철학적 이슈까지 아우르며 균형을 잡아야 한다. 이와 동시에, 이해관계자 사고집단의 범위를 기술 전문가와 도입 찬성론자들로만 제한하는 대신, 이러한 모든 분야의 전문가로 확대하고 반대론자들의 의견까지 수용해야 한다.

결국 생성형 AI는 사람과 기술 간 관계를 인터넷, 스마트폰, 클라우드가 했던 것보다 더욱 심오한 단계로 끌어올릴 것이다.

각계 전문가들의 의견을 종합하면, 생성형 AI 시장 규모는 2032년에 미화 2,000억 달러에 달할 것으로 전망된다.⁷ 이 전망에 따르면, 개인과 기업의 AI 관련 지출 중 생성형 AI가 차지하는 비율이 현재 5%에서 2032년이면 20%까지 늘어난다는 계산이 나온다.⁸ 즉 생성형 AI 시장 규모가 향후 10년간 격년으로 두 배 성장한다는 뜻이다. 더 나아가 생성형 AI가 경제 전반에 미치는 영향은 더욱 강력하고 광범위할 것이다. 우리는 이러한 잠재력을 파악하고 알리기 위한 객관적인 지침서이자 전령으로서 본고를 제시한다. 급변하는 시장을 구조적으로 파악하려는 시도에 해당하는 본고는 1) 생성형 AI의 기본 요소들을 간략히 설명하고 2) 기업용/소비자용 활용사례를 짚어보고 3)시장 참여자들이 지속가능한 사업모델을 구축하는 방식을 탐구한 후 4)생성형 AI의 미래에 대한 고찰과 과감한 예측으로 내용을 마무리한다.



01 생성형 AI의 '마술' 해독하기

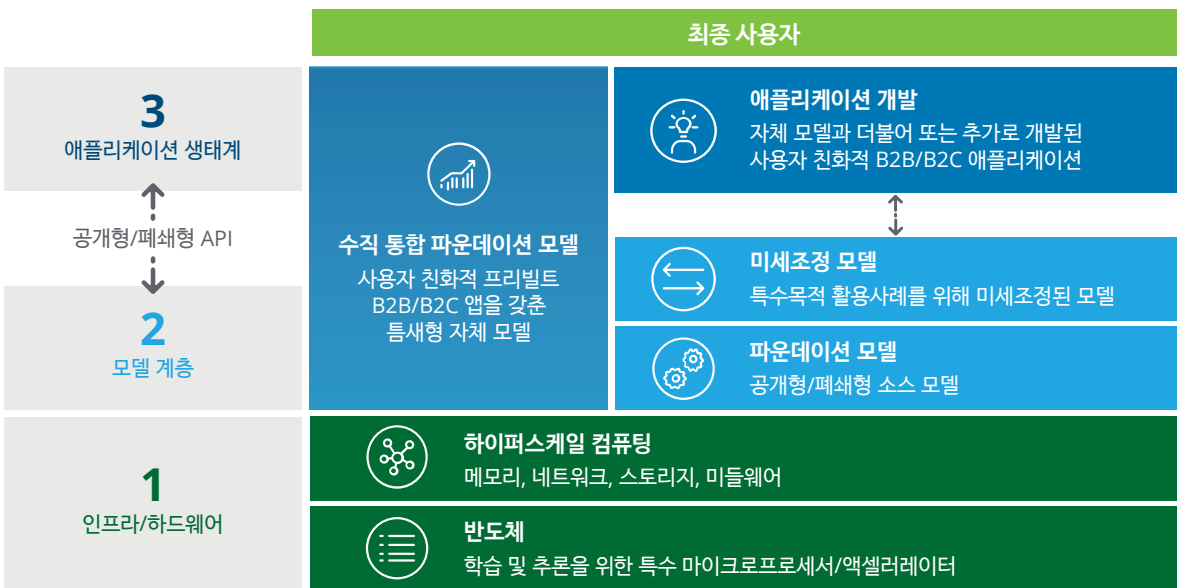
생성형 AI의 잠재력에 대한 기대가 갈수록 높아지고 있다. 이러한 잠재력은 상호 연결된 하드웨어, 소프트웨어, 데이터 제공 생태계 전반에서 끊임없는 발전과 혁신이 이뤄져야 실현될 수 있다.

생성형 AI의 기술 스택(tech stack)은 인프라-모델-애플리케이션 3개 계층으로, 사실 인프라-플랫폼-애플리케이션 계층으로 이뤄진 이전 AI의 기술 스택과 크게 다르지 않다.

이 가운데 인프라는 대개 가장 안정적이고 상업적 기반이 이미 구축된 계층으로 간주된다. 엔비디아의 그래픽처리장치(GPU)로 대표되는 특수 목적용 반도체(마이크로프로세서)와 AI 워크로드 처리에 최적화된 구급의 자체 개발 AI 전문 칩 텐서프로세싱유닛(Tensor Processing Units, TPU) 등 이미 기술과 서비스 수준을 갖춰 놓은 업체들이 메모리, 네트워크, 스토리지를 제공한다. 한편 애플리케이션 계층은 생성형 AI의 플랫폼 계층이라 할 수 있는 파운데이션 모델(foundation model)*을 활용하고 확장하기도 하면서 빠르게 진화하고 있다.

* 파운데이션 모델은 레이블이 없는 방대한 데이터셋에서 훈련한 AI 신경망으로, 미국 스탠포드대 파운데이션모델연구센터(Center for Research on Foundation Models)는 트랜스포머 모델과 대규모언어모델(LLM) 그리고 계속 구축되는 신경망 모델을 이러한 파운데이션 모델 범주로 분류했다.

그림 1. 생성형 AI의 기술 스택



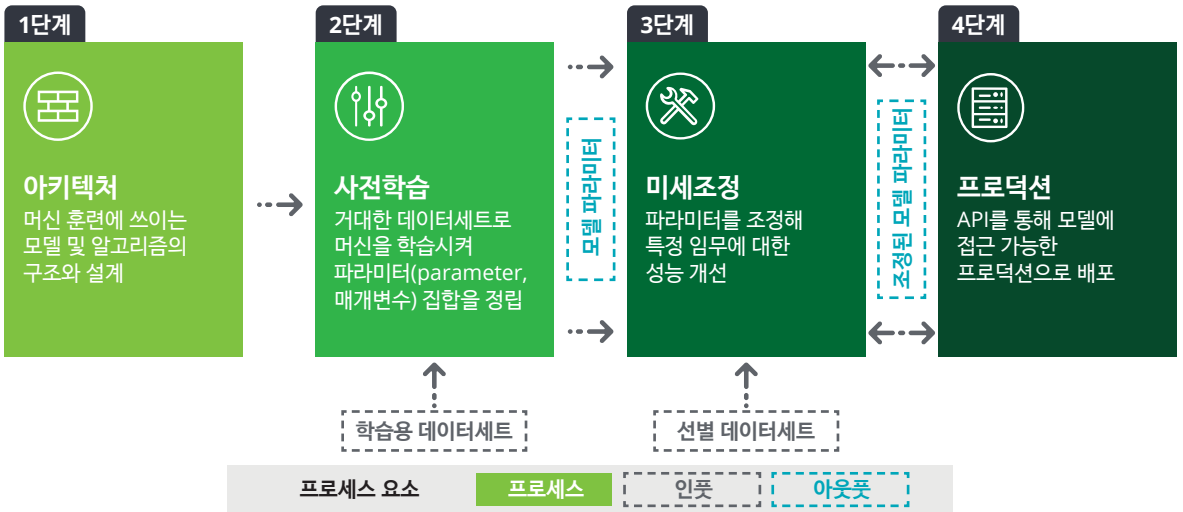
출처: 딜로이트

생성형 AI의 기술 스택과 이전 AI 기술 스택의 차별점은 바로 파운데이션 모델에 있다. 생성형 AI 파운데이션 모델의 핵심은 머신러닝(ML) 모델이다. 광범위한 데이터셋으로 사전 훈련을 받은 ML 모델이 다양한 문제에 대한 해결사로 진화한다.⁹ 예를 들어, 마이크로소프트의 'Win32'가 제공하는 API로 개발자들은 베이스 레벨의 하드웨어와 OS 기능을 활용할 수 있고, 엔비디아의 'CUDA'를 이용하면 게임엔진과 같은 그래픽 집약적 애플리케이션을 만들 때 손쉽게 GPU 리소스를 확보할 수 있다. 이처럼 모델 계층은 한 단계 높은 수준의 애플리케이션을 개발하려는 개발자들에게 최적화된 하드웨어를 제공함으로써, 생성형 AI의 도입을 가속화하고 대중화할 수 있도록 지원하는 역할을 한다.

파운데이션 모델은 대개 공개형/폐쇄형 API를 통해 개발자들에게 제공된다. 이후 개발자들은 학습 데이터를 추가하는 방식으로 미세조정을 거쳐 ML 모델의 상황인식 능력, 상관성, 성능을 강화해 특수목적 활용사례에 배치될 수 있도록 한다. 또 이와 동시에 제공 비용을 최적화하려는 시도를 지속한다.

파운데이션 모델은 대체로 아래 그림에서 도해한 것과 같은 4단계를 거쳐 개발된다.

그림 2. 파운데이션 모델 개발의 4단계



출처: 딜로이트

이러한 프레임워크는 AI 아키텍처 전반에 응용할 수 있지만, GPT-3, 스테이블 디퓨전(Stable Diffusion), 메가트론-튜링(Megatron-Turing) 등 최첨단 파운데이션 모델들은 트랜스포머라는 신경망 아키텍처에 기반한다. 트랜스포머는 2017년 구글 브레인(Google Brain) 연구팀의 논문¹⁰에서 처음 소개된 아키텍처로, 이전 아키텍처와 비교해 ML 성능을 획기적으로 끌어올렸다. 그 결과 맥락 파악, 관계 추적, 결과 예측 능력이 크게 강화됐다. 오늘날 가장 성숙한 생성형 AI 파운데이션 모델은 텍스트 영역으로, 방대한 텍스트 학습 데이터에 기반해 생성형 AI 파운데이션 모델의 한 유형인 대규모언어모델(LLM) 개발이 가속화되고 있다. LLM은 순차적으로 다음에 나올 단어나 단락 내 빠진 단어를 예측해 텍스트를 만들도록 훈련받는다.

게다가 생성형 AI는 코드, 이미지, 비디오, 오디오, 3D 모델 등 다양한 종류의 인공물(artifact)을 만들어낸다. 이로 인해 **카피라이팅부터 연구활동, 소프트웨어 엔지니어링까지 다양한 영역의 생산성이 거대한 파급적 변화를 겪을 수 있다.** 예를 들어, 광고 업계에서 생성형 AI를 도입하면 기계가 사람 대신 참신한 광고 문구와 제품 설명서를 쓰고 필요한 이미지까지 단 몇 초 만에 만들 수 있다. 또 의료 부문에서는 합성 엑스레이 이미지를 만들어 의료인의 진단 훈련에 사용할 수 있다.

또한 생성형 AI는 기업의 운영 방식과 고객 소통 방식에 변화를 가져올 것이며, 현재 우리가 알고 있는 '직원'의 개념까지 재정립할 수 있다. 일부 소비자 및 기업 영역에서 대전환은 이미 시작됐다.

02

소비자와 기업의 생성형 AI 활용사례

지난 2022년 오픈AI가 대규모 이미지 생성 모델 달리2(DALL·E 2)를 출시했을 때 즉각 세상의 관심이 집중됐다.¹¹ 달리2는 '정글 속 사자를 그려줘' 또는 '농구 하는 사자 두 마리를 피카소풍으로 그려줘'와 같은 단순한 프롬프트를 입력하면 똑딱 그림을 그려준다.

이후 챗지피티 출시와 뮤직LM(MusicLM) 프리뷰 등 생성형 AI가 끊임없이 뉴스에 등장했다. 그리고 당연한 수순으로 오픈AI의 챗지피티를 탑재한 Bing(Bing) 검색엔진 등 광범위하게 활용할 수 있는 소비자 활용사례가 확산됐다.¹² 정보검색부터 건강관리까지 소비자용 생성형 AI 활용사례의 '캠브리아기 대폭발' 양상이 진행 중이다.

이러한 폭발적 확산 속에서 일종의 맥락을 잡기 위해 개인의 일상에 변화를 가져다준 소비자 활용사례를 유용성에 기반해 다음과 같이 네 가지로 분류했다.

- ☑ **효율성:** 계획 수립, 연구, 제품 개발 등 업무의 최적화
- ☑ **교육:** 개인 맞춤형 교습 및 학습 콘텐츠 제공
- ☑ **창작:** 사람의 창작 과정을 모방해 콘텐츠를 창작하거나 내용을 보강
- ☑ **엔터테인먼트:** 게임과 가상 페르소나, 기타 엔터테인먼트 제작

이상은 소비자용 생성형 AI 시장을 1차적으로 분석한 결과로, 앞으로 기술과 시장이 발전하면서 상호 겹치는 카테고리가 생길 수 있다. 게다가 다중 모델(multi model) 애플리케이션이나 휴대용 기기만으로도 사용 가능한 애플리케이션 등이 등장해 미래 세대 AI가 성숙기에 접어들면, 새로운 카테고리로 분류돼야 하는 활용사례가 속속 등장할 것이다.



그림 3. 현재 대표적인 생성형 AI 소비자 활용사례 표본 추출

	효율성	교육	창작	엔터테인먼트
유용도 복잡성	건강 및 웰니스 계획 수립	가상 동료와 대화 소통	동영상 파일 생성 및 편집	게임 창작
	신제품 개발	개인 맞춤형 재무계획 수립	인테리어 디자인 모형 제작	팝 문화 유명인사와 채팅
	출처 표기 연구 활동	외국어 교육	의상 코디네이션 및 패션 조언	3D 환경 렌더링
유용도 복잡성	콘텐츠 큐레이션	리서치 보고서 정리	디자인 파일 수정 및 편집	음원 리믹싱 또는 샘플링
	일반적 질의 응답	개인 글쓰기 교습 및 정보 제공	예술품 창작 및 이미지 편집	새로운 단편소설 창작
	샘플 공급업체			
	· 신디시스(Synthesis.ai) · 컨센서스(Consensus)	· 그램머리(Grammarly) · 링고스타(Lingostar.ai)	· 루미나르(Luminar AI) · 렌사(Lensa)	· 재스퍼(Jasper) · 시나리오(Scenario)

출처: 딜로이트

생성형 AI의 발전 속도는 예측이 불가능할 정도지만, 2023년 초에 이미 등장한 다음과 같은 소비자 활용사례는 앞으로도 지속될 것으로 예상된다.

출시 속도

소셜미디어 등을 통해 생성형 AI에 대한 소비자의 인식과 이해가 높아지고 있는 만큼, 관련 업체들은 이러한 열풍에 힘입어 고객 확보에 드는 비용을 줄이고, 제품 결함을 해결하고, 적극적으로 기여도가 높은 사용자 기반을 통해 효율적으로 규모를 확대할 수 있다.

업무 수단

쓰기 도구와 같이 업무환경에서 가치를 창출하는 제품은 지속 가능한 사업 모델에 손쉽게 통합할 수 있다. 반면 소셜미디어 필터와 같이 그 유용성이 다소 과장된 제품은 도입이 쉽지 않을 수 있다.

끊김 없는(seamless) 통합

플랫폼에 통합되는 생성형 AI 솔루션은 기존 업무흐름에 이미 녹아 들어 있으며, 앞으로도 더욱 지속적으로 도입될 것이다. AI 기반 영문법 교정 서비스인 그램머리(Grammarly)는 PC에 앱으로 설치할 수 있고, 마이크로소프트 Bing(Bing)에는 오픈 AI의 챗지피티가 탑재돼 있다.

소비자 활용사례는 기업용 활용사례의 단초를 제공하기도 한다. 하지만 소비자와 달리 기업들은 첨단 사양, 입증된 투자수익률(ROI), 주문제작, 조직화된 콘텐츠, 보안, 기술 지원 등이 추가로 필요하다. 각종 기능의 생성형 AI가 쏟아져 나오고 있는 지금, 가장 일반적으로 활용돼 내부적 또는 B2B 성과를 도출할 수 있는 기업용 활용사례는 각종 산업 및 부서를 망라하고 범용 목적으로 도입할 수 있는 '수평적'(horizontal) 활용사례가 될 것이다. 하지만 이전 기술이 그러했듯, 특정 산업에 특화된 '수직적'(vertical) 기업용 활용사례도 지속적인 가치를 창출할 기회가 늘어나고 있다.

수평적 활용사례를 도입할 수 있는 후보 대상은 이미 구축이 완료된 자동화 센터들이다. 지식 기반, 지원 채팅 기록 등 학습 데이터가 매우 풍부하고, 비용 최적화와 생산성 개선 여지가 많기 때문이다. 예를 들어, 광고 문구 작성, 블로그 글쓰기, 소셜미디어 문구 작성 등 광고 목적의 글쓰기와 같은 창의적 마케팅 업무는 사람이 수행하자면 몇 시간에서 최대 며칠이 걸릴 수도 있다. 하지만 생성형 AI는 쓸 만한 초안을 단 몇 분 만에 써내, 사람이 할 일은 이를 다듬는 일뿐이다.

이러한 효율성 향상으로 해당 직무에 요구되는 자질이 달라질 수 있으며, AI에게 적절한 질문을 던지는 능력 등 프롬프트 엔지니어링(prompt engineering)이 능력 차별화의 기준이 될 수 있다. 궁극적으로 수평적 활용사례는 더욱 특화된 애플리케이션을 위한 상업용 기반을 만들 것이다. 기업들은 이러한 수평적 활용사례를 조기에 도입해 생성형 AI 활용 능력과 지식 기반을 쌓아, 수직적 애플리케이션으로 가치를 창출할 활용사례를 만들어가야 한다.

현재 일부 기업들은 이미 생성형 AI의 수평적 활용사례에 투자해 가시적 성과를 얻고 있다.

연구팀이 제3자의 정보를 요약하거나, 제품 매니저들이 요구사항 정의서를 작성하거나, 소셜미디어 마케팅 담당자들이 광고문구를 다듬거나, 고객센터 팀이 고객사례를 요약하거나 해결책을 제시하는 등 생성형 AI의 수평적 활용사례는 이미 적극 도입되고 있다. 하지만 실질적인 ROI를 얻기 위해서는 서비스 가능한 자체 데이터뿐 아니라 안전한 모델 파티셔닝(partitioning, 분할), 뛰어난 제품 리더, ML옵스(MLOps) 툴을 지원할 수 있는 ML 엔지니어, 새로운 상용 및 운영 모델이 필요하다. 기업들은 생성형 AI 도입에 있어 조기 수용자이건 재빠른 추종자이건 후발 주자이건 이러한 환경을 구축하기 위한 투자 여건을 지금 평가해야만 한다.

그림 4. 기업용 생성형 AI 수직적/수평적 활용사례 표본 추출

	소비자/소매	생명과학/의료	은행/금융서비스	첨단기술	미디어/통신	산업/제조	정부/공공부문
새로운 활용사례	맞춤형 대화형 소매 경험 ●●	디지털 테라피 용도의 AR/VR 콘텐츠 제작 ●●●●●	금융사기 시뮬레이션/패턴 탐지 ●●●	맞춤형 AR/VR 경험 생성 ●●●●●	게임 창작 ●●●●●●●	석유 탐사 목적의 지질 평가 ●●●●●	365일/24시간 사용 가능한 학술 가상 비서 ●●●●●
	맞춤형 제품 설계 & 제안 ●●●●	예측 및 가상 환자 분류 ●●●	세금 및 규제 준수 감사/시나리오 테스트 ●●●●	자동화 제품/하드웨어 설계 ●●●●	예고편/요약본 제작 ●●●●●	생성형 시뮬레이션/안전 테스트 ●●●	인프라 매핑 및 계획 ●●●●
수직적	제품 세부사항 & 사진 생성 ●●●	교육용 3D 해부 영상 ●●●	소매은행 거래 지원 ●●●	맞춤형 자동화 UI/UX 설계 ●●●●	대본/악보 창작 및 자막 생성 ●●●●	유정/파이프라인 등 3D 환경 렌더링 ●●●	재해 복구 시뮬레이션 ●●●
	패션 의상 코디네이션 ●●●●	건강/웰니스 계획 수립 ●●●●	맞춤형 가상 재무 자문 ●●●●	제품 테스트/피드백 제공 ●●●●	맞춤형 뉴스/콘텐츠 제작 ●●●●	자동화 기술 장비 훈련 ●●●●●	사기/낭비/권한 남용 예방 보고 ●●●
성숙기에 접어든 활용사례	개인 예술 창작 & 편집 ●●●	분자 시뮬레이션을 통한 신약 개발 ●●●●	재무 보고 분석/통찰력 제공 ●●●●	소프트웨어 판매, 고객 경험/유지 지원 ●●●●	새로운 단편소설 창작 ●●●●	스마트 공장 용도의 생성형 자동화 ●●●●	연구활동/출처 확인 및 해설 ●●●●
	맞춤형 대화형 소매 경험 ●●●	HR/IT 기능의 셀프서비스화 ●●●●	E2E 자동화 고객 서비스 ●●●●	고객 피드백에서 고객 정서 분류 ●●●●	자동화 코드 디버깅 및 문제 해결 ●●●●	가상 비서를 위한 대화 생성 ●●●●	
수평적	기업 검색 & 지식 관리 ●●●●	메타버스 등 3D 환경 제공 ●●●●	마케팅/판매 콘텐츠 제작 ●●●●	문자↔음성 전환 통한 접근성 지원 ●●●●	자동화 코드 생성 및 완성 ●●●●	플랫폼 전반에 걸친 맞춤형 타겟 광고 ●●●●●	

* 생성형 AI 모달리티(modality, 데이터 유형): ● 문자 ● 이미지 ● 오디오 ● 비디오 ● 3D 모델 ● 코드 ● 기타

출처: 딜로이트

한편 수직적 활용사례는 도메인(domain, 전문 분야 지식), 콘텍스트, 전문성 등을 요구하는 각 산업별 업무 프로세스에 적용하는 것을 목표로 한다.

생성형 AI를 수직적 활용사례로 도입하려면, 파운데이션 모델을 미세조정하거나 아예 새로운 특수 목적 모델을 만들어야 한다. 예를 들어, 위험보상 정보에 기반한 맞춤형 주식 포트폴리오나 환자의 병력과 증상에 기반한 맞춤형 치료 계획을 수립할 때 생성형 AI를 도입할 수 있다. 하지만 이러한 특수 목적에 맞는 수직적 활용사례를 실제로 도입하려면 해당 분야에 대한 깊은 이해가 수반되어야 한다. 일례로 소프트웨어를 만들 때 단순한 프롬프트 입력만으로도 생성형 AI가 코드를 짤 수 있지만, 이 단순한 프롬프트로 명령을 내리는 사람은 효율적 코딩과 코딩 언어, 기술 용어에 대해 깊은 이해와 전문성을 갖추고 있어야 한다.

기업들은 속도, 적합성, 소스의 범위와 같은 모델 성능만으로 벤더를 선택하지 않기 때문에, 도입할 생성형 AI 모델을 고르는 기준이 소비자와 매우 다르다. 생성형 AI 도입의 찬성론자와 반대론자 모두 동의하는 기업의 생성형 AI 도입 기준은 다음과 같다.

- ☑ **용이한 사용:** 시스템과 업무 프로세스에 바로 도입할 수 있거나 로우(low)/노(no) 코드 방식으로 도입할 수 있어, 막대한 IT 비용을 줄이면서도 실무 사용자들의 능력을 증강하는 효과를 내야 한다.
- ☑ **보안 및 개인정보보호:** 데이터보안 표준[예: 미국 데이터보안표준2(SOC-2), 미국 '건강보험 이전과 책임에 관한 법'(HIPAA), 개인정보보호 규정(GDPR) 등]에 부합하고, 기밀 데이터에 대해 역할/개인에 차등을 두어 접근권을 통제해야 한다.
- ☑ **탄탄한 생태계:** 개발된 제품과 서비스 제공자들이 다양하게 포진돼 있어 특수 데이터세트, 활용사례, 애플리케이션을 확장, 맞춤화, 공동개발할 수 있어야 한다.
- ☑ **투명성과 설명가능성:** 모델의 산출물과 대응이 도출되는 방식을 설명할 수 있고, 부정확한 결과가 도출된 근본 원인을 분석할 수 있어야 한다.
- ☑ **유연성과 맞춤화:** 파라미터를 정립하고, 자체 데이터로 학습하고, 임베딩(embedding, 데이터를 연산가능하도록 수치화 또는 벡터 형태로 변환)을 맞춤화함과 동시에 데이터와 튜닝에 대한 개인정보와 소유권을 유지해야 한다.



새로운 활용사례의 등장 속도가 가속화된다 하더라도, 생성형 AI 시장의 발전 경로는 다음과 같은 6가지 길에서 벗어나지 않을 것으로 전망한다.



수평적 활용사례가 수직적 사례보다 먼저 가치를 창출하겠지만, 어떤 자체 데이터를 활용하느냐에 따라 수직적 활용사례가 프리미엄 가치를 창출할 수 있다. 따라서 데이터가 새로운 '통화'로 부상하며, 자체 및 합성 데이터를 확보하기 위한 새로운 경제가 창출될 것이다.



모든 산업이 생성형 AI의 혜택을 보겠지만, 그 중에서도 은행·소매업·숙박업 등 고객 데이터가 풍부하거나 정보서비스 등 제품 자체에서 얻을 수 있는 데이터가 풍부한 산업이 가장 먼저 생성형 AI의 이점을 누릴 수 있을 것이다. 이와 반대로, 법률 및 의료 등 전문가의 판단이 중요한 분야는 생성형 AI의 도입에 더욱 신중할 수밖에 없지만, 이전에 축적된 데이터의 합성이 가속화되면서 결국에는 생성형 AI의 혜택을 누리게 될 것이다.



저금리 시대가 사실상 막을 내리면서 비용에 민감해진 기업들이 ROI가 분명한 용도에만 생성형 AI를 도입하려 할 것이다. 따라서 챗봇과 같이 즉각적인 비용 절감 효과가 있거나, 검색 등 생산성을 즉각 끌어올리거나, 마케팅 문구를 작성하는 등 수익에 즉각적 도움이 되는 생성형 AI는 사람을 완전히 대체하는 도구보다 훨씬 빠르게 도입될 것이다.



텍스트 기반 활용사례가 가장 먼저 상용화되겠지만, 잠재적 비용 절감 및 생산성 향상 효과는 이보다 고도의 업무를 상업화할 때 더욱 클 것이다. 이러한 스킬은 인력을 채용하거나 훈련시키는 데 더 많은 비용과 시간이 들고, 창의적 우뇌와 논리적 좌뇌의 대결 구도에서 볼 때 성공적인 결과물을 판단하는 기준이 주관적일 수밖에 없기 때문이다.



생성형 AI 관련 규제 조치는 속도, 범위, 권한, 보고 규정 등이 지역마다 상이하게 수립될 것이다[예: 미국 'AI 권리장전'(AI Bill of Rights)¹³, 유럽연합(EU) 'AI 법'(AI Act)¹⁴, 중국 사이버공간관리국(CAC, Cyberspace Administration of China)¹⁵ 등]. 따라서 생성형 AI 벤더와 이를 이용하는 기업들은 데이터 품질, 투명성, 공정성, 안전성, 무결성 등 '신뢰할 수 있는 AI'(Trustworthy AI)를 구축하는 핵심 요소들을 선제적으로 강화할 필요가 있다.



생성형 AI가 등장하자마자 일자리가 위협받을 것이라는 윤리적인 우려가 제기됐다.¹⁶ 하지만 이전 세대 AI와 마찬가지로, 생성형 AI도 우선적으로는 사람의 능력을 증강하는 역할을 하게 될 것이다. 분석가들이 사용하는 구글 워크스페이스(Workspace), 코더(coder)들이 사용하는 깃허브(GitHub), 마케팅 전문가들이 사용하는 크리에이티브 클라우드(Creative Cloud) 등 AI는 모든 이들의 업무용 툴킷에서 빠지지 않는 톨이 될 것이다.

생성형 AI의 잠재력은 무한하지만 대규모로 도입되기에 앞서 해결해야 할 문제 또한 무수히 많다. 본고에서 이러한 문제를 자세히 다루고 있지만, 그 외에 또 다루어야 할 사안이 상업적 성공 가능성이다. 다시 말해 생성형 AI가 약속하는 매력적 가능성과 활용사례에도 불구하고, 이를 제공하는 벤더들은 지속 가능한 사업 모델을 구축해야 할 과제를 안고 있다.



03

생성형 AI의 상업성과 시장경쟁 구도

생성형 AI가 약속하는 가치를 선점하고자 하는 전투가 다수의 전선에서 펼쳐질 것이다. 기술 스택의 각 계층에서 저마다의 경쟁 구도가 형성되고 있으며, 규모, 데이터, 브랜드, 종속적 고객기반 등을 누가 먼저 확보하느냐가 경쟁 역학을 좌우하고 있다.

생성형 AI 시장의 경쟁자들은 크게 두 가지 유형으로 나눌 수 있다. 첫 번째는 인프라, 모델, 애플리케이션 중 한 개의 계층에만 주력하는 벤더들이고, 두 번째 유형은 계층을 넘나들며 영역을 확장한 통합 벤더들이다. 한편 가격 구조는 기존 기술과 마찬가지로 소비자 대상인 경우 사용자당 또는 월간 구독료 형식으로 단순하지만, 기업 대상인 경우 콜(call, 요청) 건당 및 시간당 과금부터 수익배분 방식까지 조금 더 복잡하다. 하지만 생성형 AI 시장의 범위를 조기 도입 기업들이나 비주류 활용사례에서 일반적인 기업용 톨로 확대하려면, 가격 구조의 단순화, 예측가능성, 제공 가치 등을 명확히 정립할 필요가 있다.

첫째, 생성형 AI 기술 스택에서 가장 성숙한 단계에 있는 인프라 계층의 경우 하이퍼스케일러들이 시장을 장악하고 있다. 인프라 계층은 이미 수익구조가 입증된 사업모델이 있다. 확장 가능한 컴퓨팅 서비스를 사용량 기반의 투명한 가격 구조로 제공하는 것이다. 하이퍼스케일러들은 생성형 AI의 쓰임새를 '고착화'(sticky)하기 위해, 모델 제공업체들과 계약을 맺는 경우가 많다. 자체 모델을 개발했음에도 불구하고 오픈AI와 손잡은 애저(Azure)¹⁷, 앤스로픽과 손잡은 구글¹⁸, 스테빌리티AI(Stability.ai)와 손잡은 AWS¹⁹가 대표적이다.



클라우드 서비스 제공업체(CSP)들이 추상화 서비스를 제공하는 동안, 인프라 계층에서는 또 다른 하위계층이 빠르게 발전하고 있다. 바로 반도체 부문이다.

이 분야에서는 엔비디아가 각각 학습 및 추론 워크로드 용도로 제작된 암페어(Ampere)와 호퍼(Hopper) 시리즈 GPU와 함께, 훈련 시간을 단축시켜 주는 슈퍼컴퓨터 셀레네(Selene)로 단연 선두를 달리고 있다.²⁰ AMD는 머신러닝 애플리케이션에서 구동되는 엑사 스케일 컴퓨팅 용도로 제작된 CDNA2 아키텍처로 고성능 컴퓨팅 시장의 경쟁 수준을 한 단계 끌어올렸다.²¹

그림 5. 인프라 계층

오퍼링	설명	대표 업체	주 고객			주요 수익구조	
			기업	개발사	소비자	모델	단위
클라우드 서비스 제공업체	하이퍼스케일 및 특수 목적용 서버, 네트워크, 스토리지 제공	·아마존 ·바이두 ·구글 ·마이크로소프트	O	O	X	사용량 기반	·분당 ·CPU/GPU 종류별
생성형 AI 서비스 제공업체	배포 가속화를 위한 특수 서비스 (예: 보안, 모니터링, 테스트, 모델 구별)	·아마존 ·코히어(Co:here) ·구글	O	O	X	사용량 기반	·시간당 ·산출물 건당 ·임베딩 건당
칩 공급업체	GPU/CPU 등 특수 목적용 반도체	·AMD ·엔비디아	O	X	X	1회성 리스 계약	·부품당

출처: 딜로이트



둘째, 모델 계층은 시장이 빠르게 발전하고 있으나 자원 집약적 부문이라 업체들이 파라미터와 임베딩 등 아키텍처를 끊임없이 수정해야 성능을 유지할 수 있다. 따라서 모델의 무결성과 신뢰성을 보장하기 위해 프레임워크, 가드레일, 학습 메커니즘을 설계할 수 있는 아키텍트, 엔지니어, 데이터 과학자 등 AI 인력을 확보하는 것이 관건이다. 또한 **생성형 AI의 워크로드를 소화하려면 고도의 컴퓨팅 성능과 특수 반도체프로세서가 필요하기 때문에 막대한 비용이 든다.**²² 모델 계층에서 활동하는 업체들은 이러한 비용 문제를 타개하기 위해 이미 수익화된 제품과 통합하거나 비용을 받고 이러한 제품에 모델 탑재 계약을 맺는 방식으로 투자금을 회수하기 시작하고 있다. 마이크로소프트 엣지(Edge)에 탑재된 GPT-3.5와 구글 서치에 탑재된 람다(LaMDA)가 대표적이다.

이보다는 덜 활성화된 방법이지만 개발한 모델 아키텍처에 라이선스 계약을 맺거나 모델 개발 플랫폼을 운영하는 것도 또 다른 수익구조가 될 수 있다.

반도체 등 여타 산업에는 ARM이 CPU로, 퀄컴(Qualcomm)이 무선 네트워킹으로 각각 라이선스 수익 구조를 만들어 안정적인 대규모 사업모델을 구축했다.

그림 6. 모델 계층

오퍼링	설명	대표 업체	주 고객			주요 수익구조	
			기업	개발사	소비자	모델	단위
폐쇄형 소스 모델 제공업체	방대한 데이터 코퍼스 (corpus, 문치)로 구축한 모델 호스팅 및 관리	·코히어 ·구글 ·오픈AI	O	O	X	사용량 기반	·토큰당 ·API 콜 건당
오픈소스 모델 제공업체	커뮤니티가 유지하는 파운데이션 모델	·메타 ·스테빌리티	X	O	X	미세조정 모델 또는 모델 허브를 통해 수익 확보	
미세조정 모델 제공업체	파운데이션 모델의 개별 활용사례 맞춤 버전	·코히어 ·C3	O	X	X	·사용량 기반	·토큰당 ·API 콜 건당
모델 허브	모델 마켓플레이스, 커뮤니티, 호스팅 서비스	·깃허브 ·허깅(Hugging) ·페이스(Face) ·레플리케이트 (Replicate)	O	O	X	·구독형 ·사용량 기반 ·수익 배분	·월 사용료 ·시간당
모델 서비스 제공업체	자체 아키텍처, 합성 데이터, 가중치, 임베딩	·코히어 ·모스틀리AI(MostlyAI) ·리얼AI(RealAI)	X	O	X	·1회성 구독 ·라이선스	·임베딩 건당 ·월 사용료 ·사용자당

출처: 딜로이트



셋째, 애플리케이션 계층은 모델과 최종사용자를 연결하는 게이트웨이 역할을 한다.

현재 생성형 AI 애플리케이션들은 대체로 구독형 또는 반복거래를 통해 수익을 얻는 구조로 운영되고 있다. 이는 지속 가능한 형태의 수익구조라 할 수 있지만, 생성형 AI의 특성에 맞게 사업모델을 수정할 필요가 있다.

그림 7. 애플리케이션 계층

오퍼링	설명	대표 업체	주 고객			주요 수익구조	
			기업	개발사	소비자	모델	단위
플랫폼	앱 제작 및 배포에 필요한 소프트웨어 개발 키트(SDK), 프레임워크, 각종 툴	·구글 ·허깅 ·페이스 ·마이크로소프트	O	O	X	·라이선스 ·수익 배분	·사용자당
단독 애플리케이션	워크플로우 수정을 위한 통합 솔루션	·부미(Boomy) ·캔바(Canva) ·렌사	O	X	O	·구독형 ·사용량 기반 ·1회성 계약	·사용자당 ·월간 사용료 ·서비스당
플러그인	생성형 AI 앱의 임무와 워크플로우를 보장하는 추가 기능 및 사양	·이아트(AI Art) ·그래머리 ·재스퍼	O	X	O	·구독형 ·사용량 기반	·사용자당 ·월간 사용료

출처: 딜로이트

애플리케이션 계층의 경쟁은 몇 개 시장 내에서 전개될 수 있다.

하지만 앞으로 다양한 생성형 AI 애플리케이션과 활용사례가 등장하는 것은 필연적이므로, '마이크로 시장'(micro-market)에 주목할 필요가 있다.

전반적으로 오늘날 이미 실현됐거나 앞으로 예상되는 기업용 활용사례는 아래 5가지 범주로 구분할 수 있으며, 이 구분에 따라 시장 경쟁구도가 형성될 것으로 예상된다.



가속화: 결과물 산출 속도를 가속화해 생산성을 향상한다. 그 과정에서 사람의 개입이 배제되지는 않지만, AI 모델이 생산한 고품질 산출물에 기반해 사람이 품질을 고도화할 수 있다.



맞춤화: AI 모델이 개인 데이터를 활용해 콘텐츠를 맞춤화함으로써 과거에는 실행이 어려웠던 개인 맞춤화와 친밀한 관계형성이 가능해진다.



자동화: AI 모델이 비즈니스 및 기술적 업무 프로세스를 수행하고, 일부의 경우 사람을 대체하기도 한다. 자동화는 즉각적 비용 절감 잠재력이 있기 때문에 공급업체들이 생성형 AI의 이러한 특징을 강조하고 있다.



창작: 그 자체로 새로운 예술 형식인 프롬프트를 활용하여 이미지, 비디오, 텍스트, 미디어 등 새로운 콘텐츠를 생성함으로써 지식재산 개발의 한계를 뛰어넘을 수 있다.



시뮬레이션: 프로덕션에 적용하기에 앞서 업무 프로세스와 실험, 경험을 시뮬레이션하여 시간과 비용, 물리적 자원을 절약할 수 있다.

그림 8. 생성형 AI의 기업형 마이크로 시장 표본 추출

	가속화	맞춤화	자동화	창작	시뮬레이션
애플리케이션	이메일 발송	소셜미디어 마케팅	일정 관리/행정 비서	이미지/로고 창작	3D 모델링
	기록 작성	게임 환경 설계	기조연설 기록	광고 문구	마케팅 캠페인
	콘텐츠 마케팅	물리적 상품 설계	고객지원 챗봇	숏폼 영상 생성	의료용 테스트(R&D)
	광고용 영상 편집	자연어처리(NLP) 기반 이메일/앱 응답	콘텐츠 요약	제품 기획/제품요구 조건 작성	화학적 반응
	코드 작성	개인 비서	기초 코드 생성/문서화	음원 작곡	재난 대응 관리
모델	·앤스로픽 ·코히어 ·오픈AI GPT-3	·페이스북 OPT ·GATO ·마이크로소프트 X-CLIP	·빅사이언스 BLOOM ·오픈AI Codex ·탭나인(TabNine)	·오픈AI 달리2 ·사운드파이(Soundify) ·스테이블디퓨전 (Stable Diffusion)	·크레이들(Cradle) ·드림퓨전 (DreamFusion) ·엔비디아 GET3D

출처: 딜로이트

애플리케이션 계층의 이러한 경쟁 구도는 모델 및 인프라 계층의 변화 또한 이끌 것으로 예상된다. 이에 따라 생성형 AI 기술 스택의 하단에 위치한 벤더들은 마이크로 시장의 기반이 되는 특수 목적용 인프라, 모델, 서비스를 구축해야 경쟁력을 유지할 수 있을 것으로 전망된다.

한편 생성형 AI 시장 경쟁 구도에서 두 번째 유형으로 등장한 통합 벤더들은 자체 또는 3P 채널을 통해 한 개의 계층에만 주력하는 벤더들과 달리 수직적으로 통합돼 있거나 다수의 계층에서 활동한다. 이들 통합 벤더들은 번들 가격제, 독점 데이터, 특수목적용 클라우드, 도메인을 넘나드는 전문성 등으로 경쟁우위를 확보한다.

이들 통합 벤더들은 보통 두 가지 방식으로 탄생한다. 우선 처음부터 특수목적용 애플리케이션을 출시한 앤스로픽, 미드저니 같은 벤더들이 있다. 한편 엔비디아와 같이 생성형 AI 기술 스택의 하단에 위치한 기업들이 자체 GPU를 활용해 특수목적에 최적화된 바 이오네모와 같은 모델을 출시해 여러 계층을 넘나드는 경우도 있다.

그림 9. 통합 벤더

오퍼링	설명	대표 업체	주 고객			주요 수익구조	
			기업	개발사	소비자	모델	단위
모델& 애플리케이션	독점적 자체 모델로 구축한 애플리케이션	·앤스로픽 ·코히어 ·미드저니 ·오픈AI	O	X	O	·구독형 ·사용량 기반	·월간 사용료 ·사용자당 ·서비스 건당 ·다운로드 건당
모델& 인프라	통합 관리 인프라 및 서비스형 모델	·구글 ·엔비디아	O	O	X	·사용량 기반	·시간 당 ·API 콜 건당 ·임베딩 건당
반도체칩& 인프라	머신러닝 워크로드를 위한 수평적/수직적 클라우드	·아마존 ·애저 ·구글 ·엔비디아	O	O	X	·사용량 기반	·분당 ·CPU/GPU 종류별
E2E	자체 모델/클라우드로 구축한 애플리케이션	N/A	O	X	O	·사용량 기반 ·구독형	·사용자당 ·월간 사용료 ·시간당

출처: 딜로이트

생성형 AI 시장의 경쟁 역학은 기술적, 상업적 측면에서 동시에 형성되고 있다. 기술적 측면에서는 반도체 칩, 데이터세트, 모델이 갈수록 새롭고 첨단화되는 가운데, 일부 모델은 파라미터가 곧 1조 개를 넘을 것으로 예상된다.²³ 상업적 측면에서는 끊임없는 변화가 진행되며 챗지피티의 프로(Pro) 버전 등 소비자 솔루션이 탄력을 받고 있다. 기업용 솔루션은 아직 대규모 상업화가 이뤄지지 않고 있으며, 막대한 컴퓨팅 비용 문제와 위험 요소들이 해결될 때까지 적극적 도입이 이뤄지기는 힘들 수 있다. 이에 따라 생성형 AI 제품을 사업모델로 확장하기 위해서는 다음을 고려해야 한다.

- ☑ **이제는 아키텍처 혁신으로 얻을 수 있는 수익이 줄고 있기 때문에 학습 데이터의 품질, 사람이 AI 학습을 돕는 휴먼 인터루프(HITL, human in the loop) 훈련, 가이드라인 등이 관건이다. 이에 따라 자체 데이터세트와 귀한 AI 인력을 확보해야 경쟁우위를 점할 수 있다.**
- ☑ **범용 모델은 생성형 AI 시장의 발전에 매우 중요하지만, 가장 큰 가치를 차지하는 분야는 아니다. 대신 수직적 활용 사례와 산업 특화 솔루션을 구축해야 기업용 활용사례를 크게 확대할 수 있다.**
- ☑ **하이퍼스케일러들은 앞으로 코어위브(Coreweave)와 같이 맞춤형 하드웨어와 저렴한 가격으로 무장하고 대안을 제시하는 차세대 CSP들과 경쟁해야 한다. 따라서 기존 업체들은 부단히 혁신을 꾀하고 독립 소프트웨어 벤더(ISV)와 시스템 통합업체(SI) 생태계를 적극 활용해 경쟁우위를 유지해야 한다.**
- ☑ **생태계 조율자 및 시스템 통합자 역할을 하는 업체들의 중요성이 확대될 것이다. 이들은 기업이 데이터를 준비하고, 활용사례 우선순위를 정하고, 현지 규제를 준수하고, 수많은 모델과 애플리케이션을 통합하고, 도입에 따른 위험을 줄이는 데 중요한 역할을 한다.**
- ☑ **통합 벤더들은 한 가지 계층에만 주력하는 업체들보다 ROI를 달성할 가능성이 더 크다. 인프라, 모델, 애플리케이션을 통합한 솔루션을 통해 막대한 컴퓨팅 비용 부담을 줄일 수 있기 때문이다. 하지만 공급 병목현상이 완화되고 새로운 GPU가 등장하면, ISV들의 가치가 다시 높아질 수 있다.**

생성형 AI로 인류의 생산성이 다시 한번 극적으로 향상되는 새로운 시대가 도래할 수 있다. 하지만 지금까지 관찰된 면면을 살펴보면, 과거 몇 차례의 기술 혁신 때와 동일한 특징들이 나타나고 있다. 이를 바탕으로 시장 경쟁에서 승리할 수 있는 청사진을 그려볼 수 있다.

04

생성형 AI 도입 및 상업화 전략

생성형 AI는 사업 모델과 프로세스, 가치가 창출되는 역학뿐 아니라 개인의 업무, 학습, 교류 방식에 대전환을 가져올 수 있다. 여타 파괴적 변화를 가져온 기술과 마찬가지로 생성형 AI에 의한 변화도 처음에는 천천히 발생하다가 속도가 점차 빨라질 것이다.

소프트웨어 개발을 예로 들어보자. 일부 추산에 따르면 전 세계 인구 중 코드를 작성할 수 있는 비율은 1%도 되지 않는다.²⁴ 하지만 소프트웨어는 오늘날 수많은 기업과 사업모델에 필수적인 존재다. 생성형 AI를 전략적으로 사용한다면 코딩을 민주화해, 아이디어가 도출된 후 수익을 창출하기까지 소요되는 시간과 절차를 혁신적으로 줄일 수 있다. 생성형 AI가 제품 요구서를 작성하고, 프롬프트 명령을 받으면 코드를 짜주고, 코드를 검사해 버그를 해결하면서 코드를 최적화하고, 선제적으로 테스트와 활용사례 적용에 최적화된 환경을 제공하기 때문이다.

또 생성형 AI는 E2E 고객 확보 경로를 최적화할 수 있다. 판매 및 마케팅 부문에서 수요를 창출해야 할 때, LLM이 다양한 채널에 배포할 수 있는 마케팅 문구를 작성하고 디지털 마케팅 캠페인을 펼친다. 가트너는 2025년에 이르면 외부 마케팅 콘텐츠의 30%를 AI가 제작할 것이라 추산했다.²⁵ 한발 더 나아가 생성형 AI는 고객 정보를 수집할 수도 있고, 최초 응대용 프레젠테이션 자료를 작성할 수도 있으며, 고객사 임원들에게 토크트랙(talk track)을 제시할 수 있고, 결과와 대응 조치를 문서화 및 추적할 수도 있다. 마지막으로 생성형 AI는 선제적으로 가격 및 할인 정책을 제시하고, 계약서를 작성하며, 고객 및 고객관계관리(CRM) 기록을 업데이트할 수도 있다. 마케팅/판매 담당자들은 이렇게 AI의 도움을 받아 고객과의 관계 구축 및 가격 정책에 대한 의사결정 등 더 높은 가치를 창출하는 활동에 매진할 수 있다.

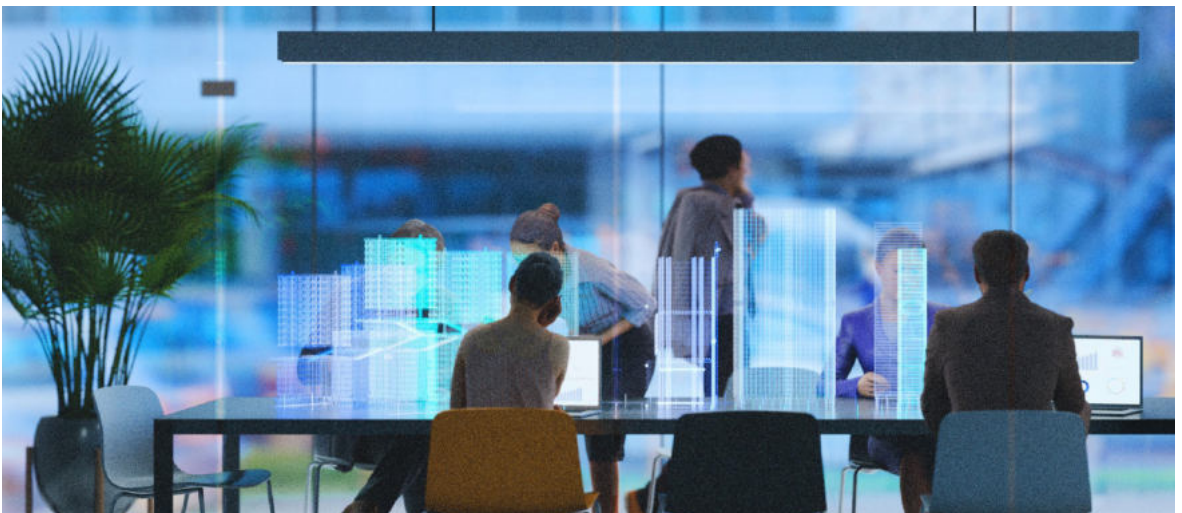
시장 리서치부터 기록 작성, 고객지원 소통 개선까지 앞서 2장에서 각 산업에서 생성형 AI를 어떻게 활용할 수 있는지 다뤘다. 여기서 더 나아가 섹터별 활용사례를 살펴보면, 생성형 AI는 자산관리 전문가들을 위해 맞춤형 재무설계를 제공할 수 있고, 의료인들의 진단을 도울 수 있으며, 미디어 및 엔터테인먼트 부문에서 새로운 세계와 경험을 창출할 수도 있고, 소매 부문에서는 의상 코디네이터 역할을 할 수도 있다. 이처럼 생성형 AI 도입에 따른 이점이 매우 강력할 것으로 예상된다. 아직 초기 단계인 생성형 AI 도입 현황을 살펴본 결과이기는 하지만, 그림 10에서 기업들이 얻을 수 있는 이러한 이점을 3장에서 다뤘던 마이크로 시장 기준으로 분류해 정리했다.

그림 10. 기업들이 생성형 AI 도입으로 얻을 수 있는 이점

	이점	가속화	맞춤화	자동화	창작	시뮬레이션
재무	TAM* 확대		●		●	
	수익 창출		●		●	
	투자 절감	●		●	●	●
	자산 활용	●	●	●		●
인력	실행속도 단축	●		●	●	●
	인력 투입 감축	●	●	●	●	
	비인력 투입 감축	●				●
직원	경험 개선		●		●	●
	스킬 향상		●			●
	관계 구축				●	
	혁신 촉발	●		●	●	●

* TAM은 총시장규모(total addressable market)를 뜻한다.

출처: 딜로이트



생성형 AI에 상업적 관심이 있는 기업들에게 지금 당장 실천에 옮길 수 있는 5가지 행동전략을 다음과 같이 제시한다.

생성형 AI 공급업체

- 1 생성형 AI 모델, 바로 도입할 수 있는 활용사례, 로우코드/노코드 툴링 등을 모두 포함한 플랫폼 접근법으로 수익을 창출함과 동시에 생태계 참여자들과 네트워크를 형성하라.
- 2 앱의 미세조정과 확장을 담당하는 개발자들이 시장 규모 확대에 핵심적 역할을 하는 만큼, 고객과 개발자 모두를 위한 솔루션을 구축하라.
- 3 이윤마진 창출 가능성이 높은 수직적 솔루션을 위한 로드맵을 구축하라.
- 4 2-3세대 생성형 AI가 곧 등장할 것으로 예상되는 만큼, 크로스/멀티 모달 솔루션에 조기 투자하라.
- 5 어크하이어*와 지식재산(IP) 등 새로운 기회를 창출할 수 있는 M&A에 투자해, 정체된 기업가치의 성장을 촉진하라.

생성형 AI 구매/사용 기업

- 생성형 AI의 잠재력과 위험에 대해 경영진에게 충분한 교육을 제공해 공통의 이해와 조율된 행동을 이끌어내라.
- 수평적 사례부터 시작해 활용사례를 파악해 우선순위를 정하라. 특히 자동화의 이점을 아직 누리지 못했던 부분에 초점을 맞춰라.
- 데이터 엔지니어링과 파이프라인, ML옵스 툴, AI에 익숙한 인력 등 명확한 기술전략을 수립하라.
- 자체 데이터 등 경쟁우위가 창출되는 근원을 파악하고 향후 쏟아져 나올 생성형 AI 활용사례에 맞춰 이러한 경쟁우위 요소들을 정비하라.
- 조언자와 파트너로 구성된 생태계에 적극적으로 참여해, 선점자의 우위를 창출하고 유리한 가격 정책 및 계약을 구축하고, 새로운 솔루션을 실험적으로 이행하라.

* 어크하이어(acquihire)는 인수 대상 기업의 인력 가치를 높이 평가해 합병 기업의 직원으로 고용하는 인재 전략을 뜻한다.

생성형 AI는 여러 위험을 수반한다.

시장 규모가 확대되는 과정에서 이러한 위험을 충분히 고려하고 해소하지 않는다면 시장 발전과 신기술의 도입은 더디게 이뤄질 것이다.

우선 생성형 AI 모델은 성능 개선을 위해 지속적으로 훈련을 받아야 하는데, 이로 인해 민감한 데이터, 개인정보보호, 데이터 보안 관련 문제가 발생할 수 있다. 또 생성형 AI의 산출물은 훈련의 품질에 달려 있다. 따라서 대표성 및 샘플링 측면에서 편향적 데이터로 훈련을 받으면 그 결과물도 편향적일 수밖에 없다. 결과물의 지식재산권 결정, 막대한 연산 비용, 역시 막대한 비용이 드는 HITL 강화 학습 등도 해결해야 할 문제다. **생성형 AI의 개발, 사용, 담론, 규제에 관여하는 모든 이들은 다음의 위험 요소들을 해소하려는 노력을 펼쳐야 한다.**

- ☑ **신뢰 약화:** 환각(hallucination), 딥페이크(deepfake), 프롬프트 인젝션(prompt injection)²⁶ 등 악의적 행태뿐 아니라 데이터 출처 미인용 등 무책임한 행태로 인해 공격 노출면이 늘어나고 고객 신뢰가 약화될 수 있다.
- ☑ **데이터 보안 및 위험:** 기업들은 빠르게 변화하는 규제 환경에 선제적으로 대응함과 동시에 필연적으로 멀티테넌트(multi-tenant) 환경에서 구동되는 데이터, 임베딩, 튜닝 기밀을 유지해야 한다.
- ☑ **편향성 및 차별:** 가드레일과 지속적 모니터링이 없다면 생성형 AI는 사람이 만든 데이터의 편향성을 모방해 차별적 행동을 확산시킬 수 있다.
- ☑ **개인정보보호 및 지식재산의 불명확성:** 생성형 AI 모델은 자체 데이터로 학습하며 여기에는 개인정보가 포함되는 경우도 많기 때문에 규제 준수, 노드 격리, 출처 추적가능성 등을 이행해야 한다.
- ☑ **비용:** 생성형 AI를 사용할 때 질의/명령(프롬프트) 비용은 인덱스 기반 질의와 비교해 10배 더 많이 든다.²⁷ 시간이 지나면 결국 비용이 낮아지겠지만, 내부 비즈니스 활용사례와 고객 가격정책에 이러한 비용의 경제학을 고려해야 생성형 AI의 도입을 촉진할 수 있다.
- ☑ **장기적인 근로자 대체:** 현재 ROI가 가장 높은 생성형 AI 활용사례는 업무 프로세스를 증강하고 생산성을 향상하는 것이다. 하지만 모델이 더욱 발전하면서 적절한 업스킬이나 인력 계획이 이행되지 않아 사람의 일자리만 빼앗길 위험을 배제할 수 없다.

결국 생성형 AI는 사람과 기술 간 관계를 인터넷, 스마트폰, 클라우드가 했던 것보다 더욱 심오한 단계로 끌어올릴 것이다.

생성형 AI의 잠재력에 강한 확신을 가지고 그 세계에 뛰어든 준비가 돼 있다면, 향후 1~2년 사이 펼쳐질 다음의 네 가지 추세를 주목하라.

☑ 생성형 AI가 일의 미래를 바꿀 것이다

AI 동료가 필수불가결한 업무 수단이 될 것이며 조직 구성원들이 대부분 AI를 활용함으로써 사람-기계 '공생의 시대'(Age of WithTM) 도래를 가속화하고 새로운 규범을 만들 것이다. 그 과정에서 AI를 활용하지 못하는 사람들은 업무환경에서 도태될 것이다.

☑ 데이터 확보만이 아니라 신뢰 구축도 관건이다

기업들의 생성형 AI 도입이 일반화되면 본격적 검증이 시작될 것이다. 따라서 직관적이고도 실질적 기대에 부합함과 동시에 신뢰를 얻을 수 있어야 생성형 AI의 도입이 확산될 수 있다.

☑ 초개인화가 생성형 AI 시장 성장을 견인할 것이다

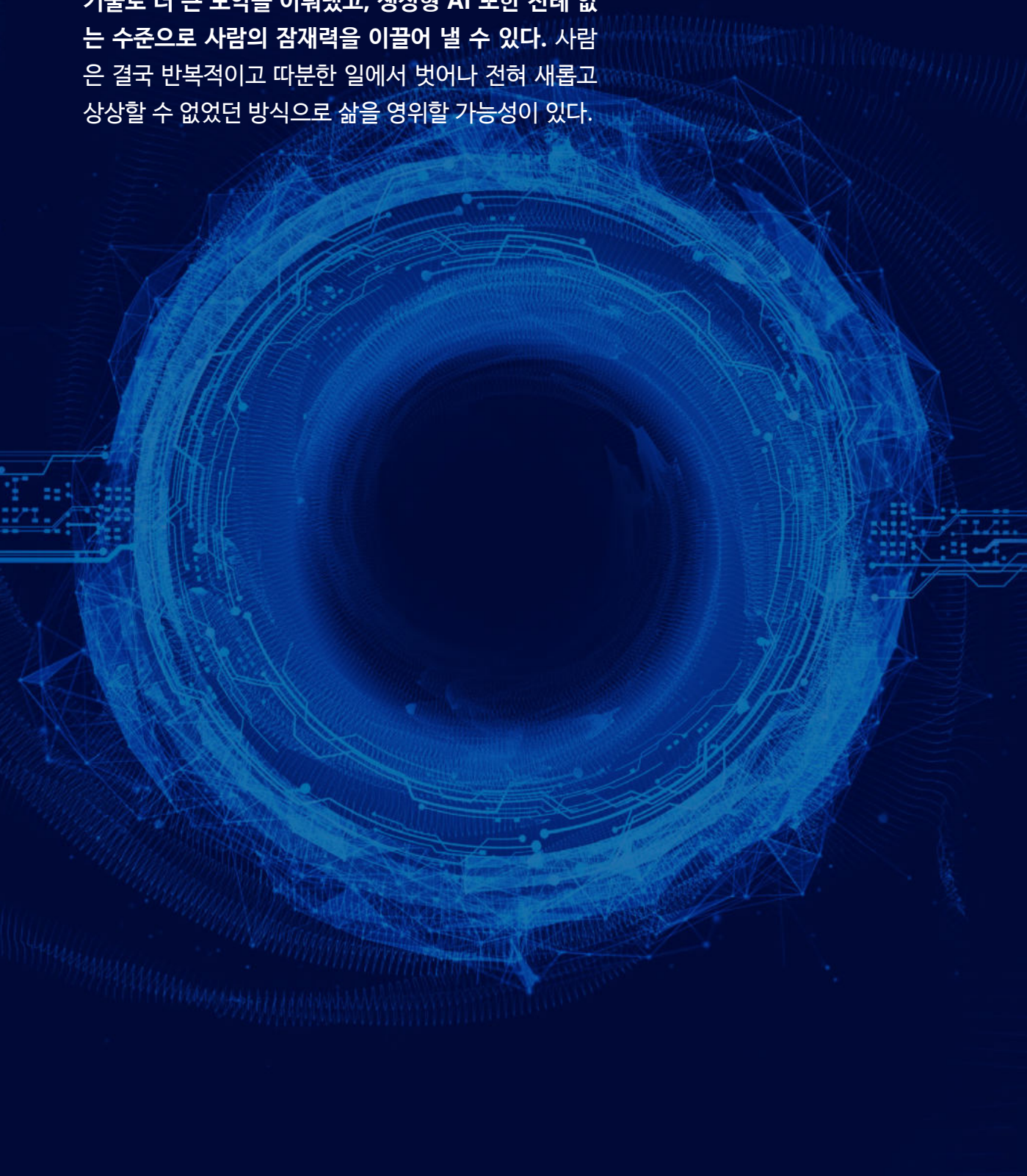
기업들은 대규모 고객 데이터를 분석할 수 있는 생성형 AI의 능력을 활용해 역학적인 맞춤형 경험, 제품, 서비스, 소통을 실시간 제공하게 될 것이다.

☑ LLM이 가장 먼저 범용화될 생성형 AI 모델이 될 것이다

LLM은 텍스트 기반 모델로 멀티모델, 유비쿼터스, 크로스-도메인 AI와는 거리가 먼 것처럼 보이지만, 범용인공지능(Artificial General Intelligence, AGI)이 거대한 뿌리를 내릴 수 있는 씨앗 역할을 하고 있다.



더 먼 미래에 생성형 AI가 인류에 어떠한 변화를 가져올지 아직은 예측할 수 없다. 일, 신뢰, 사람-기계 관계 등 모든 것이 변할 것이다. 하지만 인류는 언제나 새로운 기술로 더 큰 도약을 이뤄냈고, 생성형 AI 또한 전례 없는 수준으로 사람의 잠재력을 이끌어 낼 수 있다. 사람은 결국 반복적이고 따분한 일에서 벗어나 전혀 새롭고 상상할 수 없었던 방식으로 삶을 영위할 가능성이 있다.



딜로이트 AI 연구소(Deloitte AI Institute) 소개

딜로이트 AI 연구소는 기업들이 매우 강력하고 역동적이며 빠르게 진화하는 AI 생태계의 다양한 차원을 연결하도록 도움을 주는 조력자 역할을 하고 있습니다. 우리는 날카로운 통찰력으로 산업 전반에 걸쳐 AI 혁신 기술을 적용하기 위한 담론을 주도하며, '공생의 시대'(Age of With™)를 맞아 사람과 기계의 협력을 고취하고자 합니다.

딜로이트 AI 연구소는 AI를 둘러싼 논의와 개발을 뒷받침하고, 혁신을 촉진하며, AI 도입을 가로막는 장애물을 파악해 해결책을 제시합니다. 학계 연구단체, 스타트업, 기업가, 혁신가, 선도적 AI 제품 생산 기업 등 생태계 내 모든 구성원과 협업해 AI에 대한 핵심 영역들, 즉 관련 리스크와 정책, 윤리, 일과 인력의 미래, 응용 AI 활용 사례 등을 탐구하고 있기에 가능한 일입니다. AI 연구소는 딜로이트가 보유한 심도 깊은 AI 지식과 경험을 기반으로 이처럼 복잡한 생태계의 생리를 정확히 파악하여, 기업들이 AI에 대한 유효한 의사결정을 내려 경쟁에서 승리할 수 있도록 실질적인 조언을 제시합니다.

딜로이트 AI 연구소는 기업 이사회 멤버, 고위 경영진, 현직 데이터 과학자 등 여러분이 어떤 역할을 맡고 있든 또한 여러분이 AI 여정의 어떤 단계에 있든 상관없이, 전 세계 기업들이 AI를 활용해 어떻게 경쟁우위를 확보하고 있는지 배울 수 있도록 생생한 정보를 전달합니다.

딜로이트 AI 연구소(www.deloitte.com/us/AllInstitute)를 방문하시면, 다양한 연구 결과물뿐 아니라 팟캐스트와 뉴스레터, 전문가 담론과 라이브 이벤트 등을 접할 수 있습니다. 딜로이트 AI 연구소와 함께 AI의 미래를 탐구해 보십시오.

주석

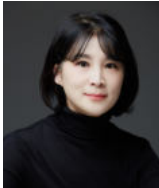
1. [Gartner: Insights for Generative AI](#)
2. [ZDNet: Google makes Contact Center AI generally available](#)
3. [VentureBeat: Nvidia boosts generative AI for biology with BioNeMo](#)
4. [Financial Times: Investors seek to profit from groundbreaking 'generative AI' start-ups](#)
5. [Bloomberg: Microsoft Invests \\$10 Billion in ChatGPT Maker OpenAI](#)
6. [Financial Times: Google invests \\$300mn in artificial intelligence start-up Anthropic](#)
7. [Global NewsWire: Global Generative AI Market Size Value Cap Expected to Grow USD 200.73 Billion By 2032](#)
8. [Precedence Research: Artificial Intelligence Software Market Size & Forecast](#)
9. [Stanford's Human-Centered Artificial Intelligence: HELM—Center for Research on Foundational Models](#)
10. [Google: Attention is All You Need, Vaswani et al.](#)
11. [Tech Crunch: Now anyone can build apps that use DALL-E 2 to generate images](#)
12. [New York Times: Bing \(Yes, Bing\) Just Made Search Interesting Again](#)
13. [The White House: Blueprint for an AI Bill of Rights](#)
14. [European Council: Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights](#)
15. [Tech Crunch: China's generative AI rules set boundaries and punishments for misuse](#)
16. [The Atlantic: How ChatGPT Will Destabilize White-Collar Work](#)
17. [Engadget: Microsoft's OpenAI supercomputer has 285,000 CPU cores, 10,000 GPUs](#)
18. [Anthropic: Anthropic Partners with Google Cloud](#)
19. [AWS: Stability AI builds foundation models on Amazon SageMaker](#)
20. [Forbes: NVIDIA Provides More Details On Selene Supercomputer](#)
21. [Tech Insights: AMD CDNA2 Targets Supercomputers](#)
22. [InfoWorld: The cost and sustainability of generative AI](#)
23. [VentureBeat: Google trained a trillion-parameter AI language model](#)
24. [Evans Data Corporation: Worldwide Professional Developer Population of 24 Million Projected to Grow amid Shifting Geographical Concentrations](#)
25. [Gartner: Beyond ChatGPT: The Future of Generative AI for Enterprises](#)
26. [Business Insider: Breaking ChatGPT: The AI's alter ego DAN reveals why the internet is so drawn to making the chatbot violate its own rules](#)
27. [Reuters: For tech giants, AI like Bing and Bard poses billiondollar search problem](#)

딜로이트 컨설팅 코어테크놀로지 그룹 및 디지털 금융 그룹

딜로이트 컨설팅 코어테크놀로지 그룹 및 디지털 금융 그룹은 테크 전략 설계부터 도입 및 최적화, AI & Data 전문 컨설팅, 클라우드 전환 및 설계, 이행까지의 라이프사이클 전방에서 선도적으로 고객을 지원합니다.

AI & Data 전문 컨설팅 조직은 데이터 표준화, 모델링 및 분석, AI 활용 방안에서 AI 운영 거버넌스 체계 수립 및 ISO 42001 인증 지원까지 엔터프라이즈 고객사의 데이터에 기반한 E2E 서비스를 제공하고 있습니다. 또한 IT에서 재무 및 공급망까지 엔터프라이즈 전반의 시스템과 조직의 특정 기능에 대한 솔루션 투자로 고객이 영향력을 확대하고 가치를 달성할 수 있는 전문 IT 컨설팅 서비스 역량을 보유하고 있습니다.

Contact Point



김우성 파트너

Core Technology 그룹리더

Tel: 02-6099-4670
Email: wooskim@deloitte.com



안상혁 파트너

디지털 금융 그룹 리더

Tel: 02-6676-3625
Email: sanghyan@deloitte.com



최규웅 파트너

Core Technology

Tel: 02-6676-3873
Email: kyuwchoi@deloitte.com



박지숙 파트너

금융 IT, 오퍼레이션 리더 |
딜로이트 컨설팅

Tel: 02-6676-3722
Email: jisukpark@deloitte.com



강기식 상무

Core Technology, chief architect

Tel: 02-6676-2039
Email: gikang@deloitte.com



이성호 이사

Core Technology,
Data 분석 전문가

Tel: 02-6676-3767
Email: sholee@deloitte.com

Deloitte.

Insights

딜로이트 안진회계법인·딜로이트 컨설팅
성장전략본부

손재호 Partner
성장전략본부 리더
jaehoson@deloitte.com

정동섭 Partner
딜로이트 인사이트 리더
dongjeong@deloitte.com

김사현 Director
딜로이트 인사이트 편집장
sahekim@deloitte.com

HOT LINE
02) 6099-4651

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited (“DTTL”), its global network of member firms, and their related entities (collectively, the “Deloitte organization”). DTTL (also referred to as “Deloitte Global”) and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see www.deloitte.com/about to learn more.

Deloitte Asia Pacific Limited is a company limited by guarantee and a member firm of DTTL. Members of Deloitte Asia Pacific Limited and their related entities, each of which are separate and independent legal entities, provide services from more than 100 cities across the region, including Auckland, Bangkok, Beijing, Hanoi, Hong Kong, Jakarta, Kuala Lumpur, Manila, Melbourne, Osaka, Seoul, Shanghai, Singapore, Sydney, Taipei and Tokyo.

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited (“DTTL”), its global network of member firms or their related entities (collectively, the “Deloitte organization”) is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.

본 보고서는 저작권법에 따라 보호받는 저작물로서 저작권은 딜로이트 안진회계법인(“저작권자”)에 있습니다. 본 보고서의 내용은 비영리 목적으로만 이용이 가능하고, 내용의 전부 또는 일부에 대한 상업적 활용 기타 영리목적 이용시 저작권자의 사전 허락이 필요합니다. 또한 본 보고서의 이용시, 출처를 저작권자로 명시해야 하고 저작권자의 사전 허락없이 그 내용을 변경할 수 없습니다.