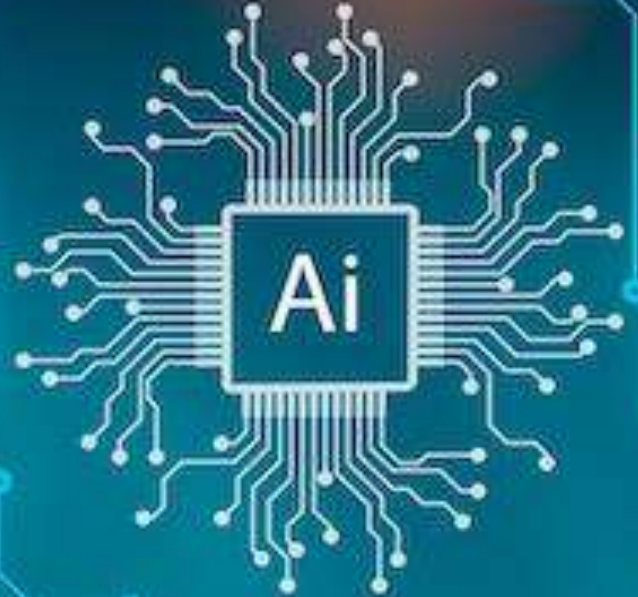


# 생성 AI, Awhile to Infinity

setting  
config



connecting  
connecting



**권예준**

010,9022,8720  
yejoon0309@g.skku.edu  
46<sup>th</sup> member

**김원영**

010,2756,1059  
wykim1059@gmail.com  
46<sup>th</sup> member

**양예지**

010,3615,1527  
yellowyeji@g.skku.edu  
46<sup>th</sup> member

# CONTENTS

PART 0. 생성형 AI 기본기 잡아가기	4
PART 1. 생성형 AI 산업의 현황과 흐름	7
PART 2. 생성형 AI, 왜 주목해야 할까	12
PART 3. 생성형 AI, 어디를 봐야 할까	15
PART 4. AI 응용 산업 맛보기	21
PART 5 개선해야 할 부분은?	30
총목 추천	32



# 용어정리

생성형 AI	대규모 데이터 세트를 기반으로 훈련된 딥러닝 모델을 사용하여 새로운 콘텐츠를 생성하는 일종의 인공지능 기술
파라미터	인간의 시냅스와 유사한 AI의 뇌 기능 담당. 파라미터 규모 커질수록 AI는 대용량의 데이터를 학습하여 인간에 가까운 능력 발휘 가능
멀티모달	단순히 텍스트에만 국한된 것이 아닌, 시각, 청각 등 여러 인터페이스를 통해 정보를 교환할 수 있는 인공지능 모델
트랜스포머	머신러닝 연구의 가장 핵심 모델. 언어 모델링부터 음성 및 컴퓨터 비전, 영상까지 확장
오픈 소스	오픈 소스 소프트웨어로 공개적으로 액세스할 수 있어 설계되어 누구나 자유롭게 확인, 수정, 배포할 수 있는 코드
강화학습	머신 러닝의 한 영역으로 알고리즘이 다양한 시도를 거치며 최적의 결과를 낼 수 있는 조합을 찾아내는 방식
머신러닝	명시적으로 규칙을 프로그래밍하지 않고, 데이터로부터 의사결정을 위한 패턴을 기계가 스스로 학습
딥러닝	인공신경망 기반의 모델로, 비정형 데이터로부터 특징 추출 및 판단까지 기계가 한 번에 수행
대규모언어모델(LLM)	테라바이트 단위의 대용량 말뭉치를 통해 모델을 학습시켜 복잡한 문장을 생성 가능
이미지 생성 모델	텍스트를 입력하면 그에 따라 이미지를 생성하는 AI기술
파인 튜닝	특정 AI 모델을 새로운 작업에 적용하고자 할 때, 기존 모델에 새로운 데이터 셋을 학습시켜 최적화해 나가는 과정
AI 편향성	AI 시스템이 정당하지 않거나, 부당하게 특정 그룹이나 개인을 차별하는 경향을 가진 것
할루시네이션 현상	환각 현상, AI가 잘못된 정보를 마치 사실인 것처럼 제시하는 현상
자연어처리(NLP)	컴퓨터가 인간의 언어를 이해, 해석 및 생성할 수 있도록 하는데 중점을 둔 AI 분야
플러그인	오픈AI의 LLM 역량을 누구나 쉽게 타 소프트웨어나 인터넷 서비스에 연동할 수 있게 하는 것
프롬프트	사용자의 다음 입력을 기다리고 있는 인터페이스를 통칭하는 용어
GPT	Generative Pre-trained Transformer. OpenAI에서 개발한 언어 생성, 이해에 중점을 둔 딥러닝 모델의 일련의 시리즈

# Part 0. 생성형 AI 기본기 잡아가기

## 생성형 AI와 LLM 알아보기

### 생성형 AI to 머신러닝

생성형 AI와 머신러닝

AI (Artificial Intelligence)는 인공지능의 약어로, 인간의 학습능력 및 추론/지각능력, 자연어 이해/처리능력(NLP) 등을 컴퓨터 프로그램으로 실현한 기술을 의미한다. 이때 머신러닝(ML)은 AI를 구현하는 하나의 방법론이자 AI의 하위 개념으로, 대규모 데이터셋과 그 패턴을 학습/분석하여 예측을 수행하고 스스로의 성능을 향상시키는 시스템과 이를 위한 알고리즘을 의미한다. 즉, 입력된 데이터를 기반으로 예측이나 결정을 이끌어내기 위한 특정한 모델을 구축하는 방식이라고 이해할 수 있다. [도표 1]

### 머신러닝 알고리즘에 대하여

머신러닝 4단계

머신러닝은 1) 데이터 수집 및 정리, 2) 특징 추출, 3) 머신러닝 알고리즘 적용하여 학습 수행, 4) 학습한 모델을 적용하는 등의 4단계로 이루어진다. 이때 학습 시스템에 따라 머신러닝 알고리즘을 3가지로 분류할 수 있다.

3가지 머신러닝 알고리즘

- 1) 감독(Supervised) 학습: 입력값과 이에 대응하는 출력값을 mapping 하는 함수를 학습하는 과정
- 2) 비감독(Unsupervised) 학습: 출력값 없이 입력값만으로, 스스로 데이터 내부의 숨겨진 구조를 발견 및 학습하여 모델을 구축하는 과정. 이때 스스로 비슷한 데이터들을 clustering하여 새로운 데이터에 대한 결과를 예측하는 방식
- 3) 강화(Reinforcement) 학습: 학습자가 입력값에 대한 올바른 출력값을 알지 못하는 상태에서 누적보상을 극대화하기 위한 행동을 선택하는 방법을 학습하는 과정.

예를 들어 감독학습은 컴퓨터에 먼저 정보를 가르치는 방식으로 이루어진다. 강아지 사진을 주고, '이 사진은 강아지다' 라고 알려주는 것이다. 이때 컴퓨터는 미리 학습된 결과를 바탕으로 강아지 이미지를 구분할 수 있게 된다. 비감독학습에서는 이러한 배움의 과정이 없다. '이 사진은 강아지'와 같은 과정이 없이 '이 사진이 강아지군'과 같이 컴퓨터가 스스로 학습하게 되는 것이다. 감독학습에 비해 발전된 기술이며, 보다 높은 연산 능력이 요구된다. 강화학습은 한 차원 높은 수준의 기술로, 애초에 입력값-출력값이 정해진 데이터가 주어지지 않는다. 16년 2월 이세돌 9단과 바둑 대결을 펼쳤던 구글의 AI '알파고'에 적용된 학습 기술이 바로 강화학습이었다.

### 딥러닝 to LLM(Large Language Model)

LLM, NLP 수행이 가능한 딥러닝 알고리즘

딥러닝은 머신러닝의 하위분야로, 계층화된 알고리즘 구조를 통해 자체적으로 배우고 지능적인 결정을 내릴 수 있는 인공신경망을 활용하기에 인간의 개입이 필요하지 않다. 즉, 인간의 개입 없이 스스로 학습하고 인간의 지능적 행동을 모방하는 기술이라 볼 수 있다. 여기서 과생되는 기술이 바로 챗 GPT와 같은 대형언어모델 (Large Language Model)이다. LLM은 자연어처리(NLP) 작업을 수행할 수 있는 딥러닝 알고리즘인데, 이때 자연어는 인간이 사용하는 언어를 지칭한다. 즉, 컴퓨터가 자연어의 의미를 분석하고 처리할 수 있도록 하는 기술이다. 이와 같은 LLM 기술이 적용된 가장 대표적인 어플리케이션이 바로 챗GPT와 같은 대화형 챗봇이다. 챗GPT와 대화를 할 때, 우리가 자연어로 대화를 입출력할 수 있는 것도 자연어 처리가 가능한 딥러닝 기술이 적용되었기 때문인 것이다.

# Part 0. 생성형 AI 기본기 잡아가기

## LLM의 성능: 파라미터와 토큰이 결정

LLM 성능,  
파라미터와 토큰으로  
판단 가능

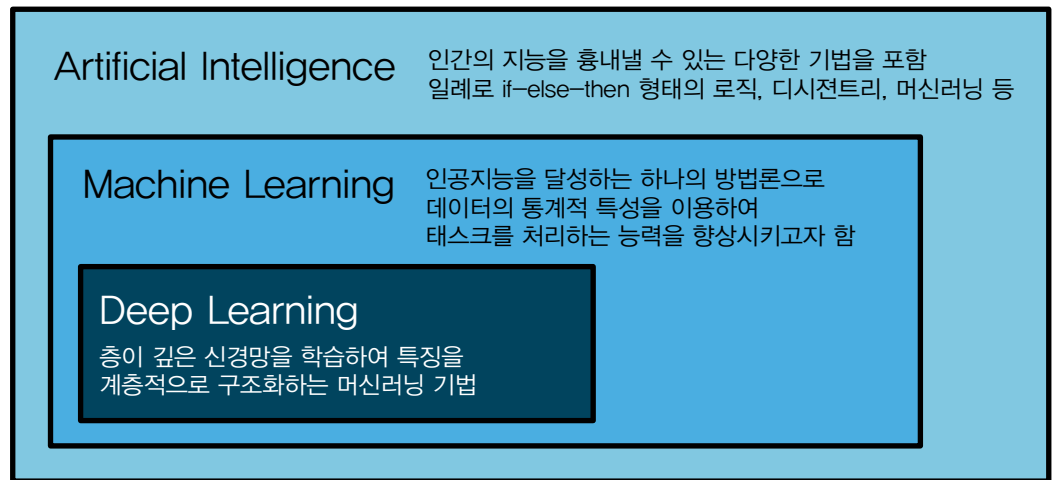
LLM은 대규모 데이터셋을 기반으로 훈련되는데, 학습에 사용된 파라미터(매개변수)와 토큰의 개수를 기준으로 해당 언어모델의 성능을 파악할 수 있다. 파라미터란 인간의 시냅스와 유사한 AI의 뇌 기능을 담당하는 요소로 AI가 학습 가능한 총량을 나타낸다. 즉, 파라미터의 규모가 클수록 AI는 대용량의 데이터를 학습해 정확성이 향상되고, 연산/학습 능력이 개선된다. 파라미터가 생각을 담당하는 뇌의 기능이라면 토큰은 LLM이 인식하는 문자 데이터 단위를 의미한다. [도표 2]

## LLM: 언어의 흐름을 예측하는 알고리즘

LLM,  
언어의 흐름을 예측하는  
알고리즘

LLM은 언어의 흐름을 예측하는 알고리즘으로, 데이터 베이스에서 얻은 대량의 텍스트를 기반으로 개별적인 의미 단위들(단어, 문구, 문장 전체 등) 사이에 가장 흔하게 나타나는 관계를 인식하도록 프로그래밍 되었다. 이에 따라 매우 높은 빈도로, 문맥상 적절하며 언어학적으로 유창하고 사실적인 답변을 생성해 낼 수 있는 것이다. 즉, 대화를 이해하는 것이 아니라, 텍스트 흐름/배열 상 다음에 위치하기에 가장 적절한 단어 및 문장을 예측해낸다는 것이다. LLM은 이를 기반으로 문서 분류, 요약 및 번역, 질의응답, 문장 생성과 같은 다양한 자연어 처리 과제를 수행한다.

도표 1. AI(인공지능)-머신러닝-딥러닝 간의 관계



자료: 인공지능 및 딥러닝 동향, 성균관대학교 금융투자학회 S.T.A.R

도표 2. LLM 별 파라미터 수 비교

	GPT-1	GPT-2	GPT-3	InstructGPT	GPT-3.5
출시일	2018.06	2019.02	2020.06	2022.01	2022.11
인코더수	12개	12/24개 36/48개	96개	96개	96개
파라미터수	1.17억개	15억개	1,750억개	1,750억개	1,750억개
학습데이터	10G	40GB	45TB	45TB	45TB

자료: Open AI, 성균관대학교 금융투자학회 S.T.A.R

# Part 0. 생성형 AI 기본기 잡아가기

## Transformer 등장 → LLM 시장 성장 ↑

트랜스포머의 등장, 대규모 데이터 처리를 한번에!

17년 구글이 발표한 인공신경망 알고리즘인 ‘트랜스포머’는 LLM의 성장을 가속화했다. 앞서 언급한 바와 같이 딥러닝과 여기서 과생된 LLM의 성능은 학습 데이터량과 파라미터 수의 증가를 통해 개선되는데, 문제는 대량의 데이터를 한꺼번에 처리하는 것이 쉽지 않다는 것이었다. 이때 구글의 트랜스포머 알고리즘은 대규모 데이터를 한번에 처리하는 것을 가능하게 했다. 이후 데이터 처리 속도가 개선되고, 언어모델에 사용되는 데이터와 파라미터의 규모가 대폭 증가하며 초거대 언어모델의 등장으로 이어지게 된 것이다. 이처럼 트랜스포머 도입 이후 AI 모델의 연산 능력은 획기적으로 발전했는데, 트랜스포머 이전 AI 모델의 연산능력은 평균적으로 2년 동안 8배 정도 증가했다면, 도입 이후에는 275배의 성장 추이를 보이고 있다. [도표 3]

## What is ‘Transformer’ ?

트랜스포머, 문맥과 의미를 학습하는 신경 모델

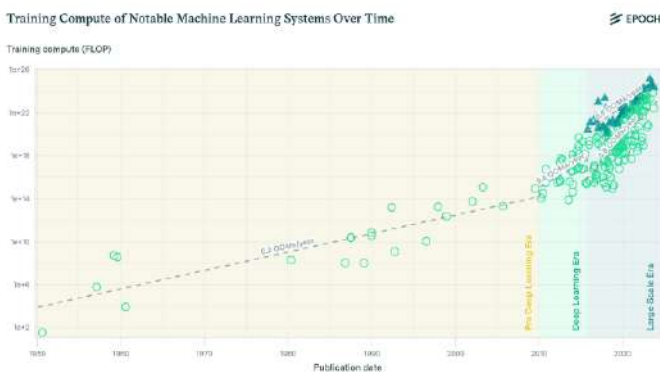
이러한 트랜스포머 모델은 문장 속 단어와 같은 순차 데이터 내의 관계를 추적하여 컨텍스트와 의미를 학습하는 신경 모델이다. 일반적으로 딥러닝 학습에 사용되는 인공신경망 알고리즘은 1) 합성곱 신경망(CNN), 2) 순환신경망(RNN), 3) 트랜스포머로 나뉘는데 이처럼 현재는 트랜스포머 모델을 중심으로 LLM 모델이 발전해 나가고 있다. 해당 알고리즘을 적용한 대표적인 LLM으로는 Open AI의 GPT-3(Generative Pre-trained Transformer-3)가 있다. [도표 4]

## Transformer, 텍스트에서 비디오까지!

트랜스포머, Text to Video 범용성 ↑

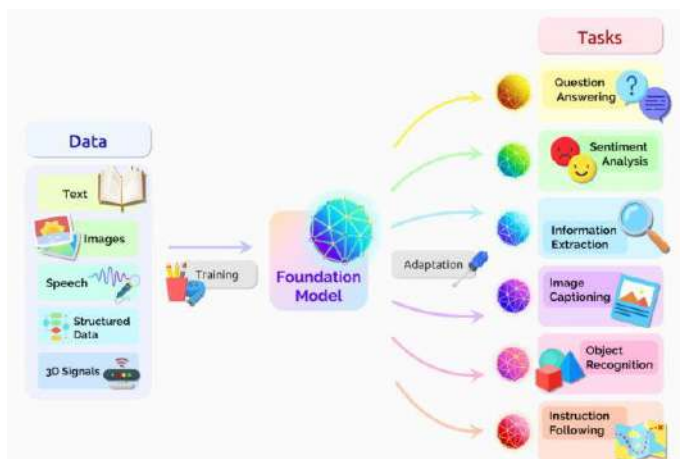
현재 트랜스포머 아키텍처는 텍스트 뿐만 아니라 이미지나 비디오 등과 같은 다양한 형태의 데이터를 학습하며, 더 많은 영역에서 범용성을 갖추어 나가고 있다. 트랜스포머는 특히 서로 떨어져 있는 데이터 요소들의 의미가 관계에 따라 달라지는 부분까지도 감지한다는 점에서 기존 기술 대비 우위를 갖는데, 이러한 경쟁력은 향후 이미지/비디오 모델에서 더욱 부각될 것이라 판단된다. 일례로 기존의 합성곱 신경망(CNN) 모델은 서로 멀리 떨어져 위치한 이미지 간의 연관성을 잘 파악하지 못했지만, 트랜스포머는 멀리 떨어져 있는 이미지를 자연스럽게 처리할 뿐만 아니라 대량의 데이터를 병렬로 연산처리 하면서 작업 시간을 단축시켜 비용이 기존의 1/15로 감소하는데 기여하였다.

도표 3. 컴퓨팅 연산량 계산력 추이 (트랜스포머 17년 등장)



자료: Epochai, 성균관대학교 금융투자학회 S.T.A.R

도표 4. 트랜스포머 모델 구조



자료: 산업자료, 성균관대학교 금융투자학회 S.T.A.R

# Part 1. 생성형 AI 산업의 현황과 흐름

## AI 산업, 여기는 글로벌 빅테크의 격전지

### Open AI의 챗GPT, AI 시장에 불을 지피다!

AI 챗GPT의 흥행과 MS의 투자

22년 11월 30일, Open AI가 공개한 GPT 3.5 의 흥행으로 글로벌 AI 시장은 성장에 박차를 가했다. 챗GPT는 출시 이후 5일만에 사용자 100만명을 돌파했으며, 이듬해 2월 MAU(월 사용자) 1억명을 돌파하며 글로벌 곳곳에 침투했다. SNS인 틱톡과 인스타그램이 MAU 1억명에 도달하는데 각각 9개월, 30개월이 걸렸다는 점을 감안하면, 챗GPT의 높은 성장세를 실감할 수 있다. [도표 5] 이후 23년 1월, Microsoft는 Open AI의 기업가치를 290억 달러로 평가하며, 기존 30억 달러에 더해 100억 달러(약 12조 3,500억 원)의 추가 투자를 결정했다.

Open AI - MS 파트너십의 의의

해당 투자를 통해 두 기업은 파트너십을 체결하며, 각자 사업분야의 성장뿐만 아니라 AI 기술의 대중화를 꾀할 수 있게 되었다. 1) Microsoft는 Open AI의 고급 대형언어모델(LLM)을 확보하여 자사 검색엔진인 Bing에 챗GPT, DALL-E 등의 기능을 추가하고, 자사 클라우드 컴퓨팅 플랫폼인 Azure를 통해 오픈AI API를 제공하며 클라우드 사용량을 높일 수 있다. 2) Open AI 입장에서는 자금 여력이 높아질 뿐만 아니라, Microsoft의 클라우드 및 탄탄한 데이터베이스를 기반으로 기술력 및 성장성을 높일 수 있게 되었다. 이 둘의 시너지는 글로벌 대중에게 친숙한 Microsoft를 통해 Open AI의 LLM 기술을 상용화할 수 있게 했으며, 곧 AI 산업의 발전으로 이어졌다.

### 이어지는 글로벌 빅테크의 LLM 투자

폭발적인 빅테크들의 LLM 투자

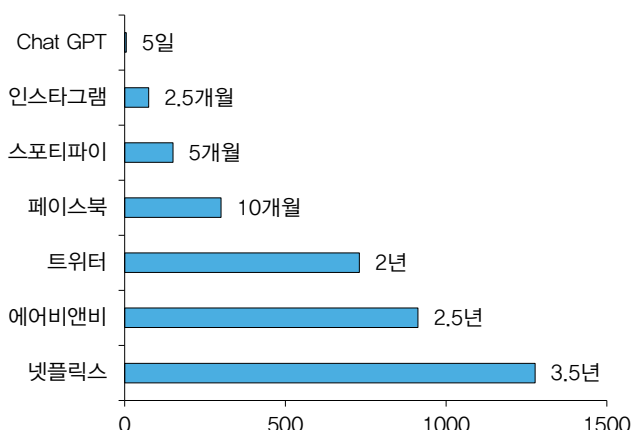
23년 9월, Amazon은 LLM 기업인 Anthropic에 최대 40억 달러를 투자하며 생성형 AI 파트너십을 체결한 바 있다. Anthropic은 Open AI와 같이 생성형 AI LLM 기술 개발 기업으로, 챗GPT와 같은 챗봇 '클로드(Claude)'를 출시한 바 있다. 이에 따라 Microsoft - Amazon 간 생성AI LLM 파트너십을 필두로 한 양강 구도가 세워졌다. 이 외에도 Google은 23년 3월 PaLM2, LaMDA LLM 기반의 AI 챗봇 Bard를 선보였으며, Meta는 23년 2월 자체 개발한 LLM인 LLaMA를 출시한 바 있다. 이처럼 LLM을 중심으로 한 글로벌 빅테크들의 공격적인 자금 투자 및 AI 기술경쟁이 심화되는 가운데, 23년을 기점으로 본격 개화하는 AI 산업에 주목해야 하는 시점이라 판단한다.

### AI 산업으로 몰리는 글로벌 자금, 담보된 성장성

글로벌 AI 시장 규모 Cagr +36.6%

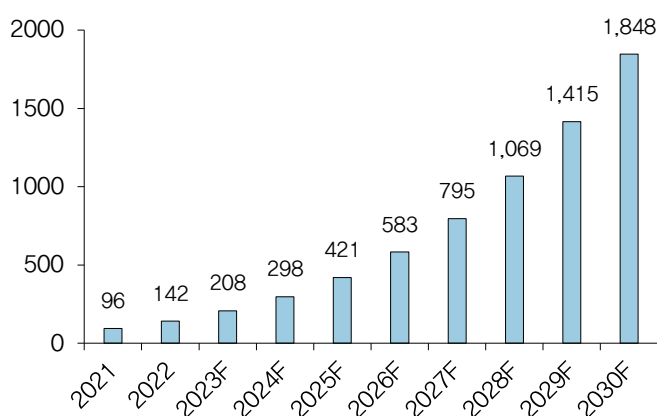
글로벌 AI 시장 규모는 23년부터 연평균 +36.6%로 성장해 28년 1조 달러를 상회할 것으로 전망된다. [도표 6] 이처럼 AI 시장 성장성이 담보된 현 시점, 글로벌 자금은 AI로 몰리고 있다. 최근 22년부터 본격 지속되어 온 고금리 상황에 따라 글로벌 스타트업에 대한 투자는 부진한 실정이다. 일례로 Crunchbase에 따르면 글로벌 월별 스타트업 투자액은 21년 11월 이후 지속적으로 하락하는 추세이며, 23년 1월의 경우 YoY -63% 감소한 180억 달러를 기록한 바 있다.

도표 5. 온라인 서비스별 유저수 100만 달성에 소요된 기간 (단위: 일)



자료: Statista, 성균관대학교 금융투자학회 S.T.A.R

도표 6. 글로벌 AI 시장 성장 전망 (단위: 십억원)



자료: Next Move Strategy Consulting, 성균관대학교 금융투자학회 S.T.A.R

# Part 1. 생성형 AI 산업의 현황과 흐름

생성형 AI 스타트업 투자  
→ YoY 13배 ↑

이런 상황에서도 생성 AI 스타트업에 대한 투자 규모는 증가세를 보인다. 생성 AI 스타트업 펀딩 동향에 따르면, 자금조달 규모는 23년 1월부터 8월까지 총 131억 달러로, 22년의 10억 달러 대비 13배 넘게 증가했다.

## 그 쪽도 생성형 AI를 아세요?

### AI 산업의 Essentials, 생성형 AI

생성형 AI,  
무에서 유를 창조하다

현재 AI 산업의 화두는 '생성형 AI'이다. 기존의 AI는 데이터를 분석 및 활용하는데 주로 사용되며, 인간의 업무를 보완하는 역할에 그쳤다. 그러나 LLM과 같은 생성형 AI는 말 그대로 '기존에 없던 텍스트, 이미지, 영상' 등을 새롭게 생성하며, 인간을 대체할 수도 있다는 시각이 대두되고 있다. 쉽게 생각해 생성형 AI는 텍스트, 콘텐츠, 예술 등을 스스로 생성하는 AI의 하위 집합이라고 볼 수 있다. 일반적으로 생성형 AI 산업의 밸류체인은 [AI 반도체 - 인프라 - 파운데이션 모델 - AI 어플리케이션]으로 구성된다. [도표 7]

### 밸류체인 No.1 AI 반도체

AI 서비스 구현을 위한  
초고속/고효율/고성능  
AI 반도체

AI 반도체는 AI 서비스 구현에 필요한 대규모 연산을 초고속/고효율/고성능으로 처리하는 비메모리 반도체로, AI의 핵심 두뇌 역할을 한다. 대규모 데이터 처리, 신경망 모델 학습 및 추론, 음성인식 등과 같은 AI 응용분야에서 주로 활용되는데, 이와 같은 AI 반도체로는 GPU(Graphics Processing Unit)와 NPU(Neural Processing Unit)가 사용된다. 기존의 CPU(Central Processing Unit)는 데이터를 순차적으로 순차적으로 직렬처리 하기에 대규모 연산을 병렬 처리하는 AI에 최적화되어 있지 않았기에, 병렬 처리의 한계를 극복하기 위해 등장한 GPU와 AI 학습/연산을 위한 전용 반도체 NPU가 사용되는 것이다. 엔비디아, 퀄컴, 인텔을 비롯한 글로벌 빅테크 기업들이 AI 반도체 개발에 주력하고 있지만, 현재 시장은 엔비디아가 90% 정도를 차지하고 있는 독점적 체계에 가깝다.

### 밸류체인 No.2 인프라

생성형 AI  
학습/연산/배포를 위한  
인프라

생성형 AI 인프라는 클라우드, 네트워크 등과 같이 생성형 AI를 학습/연산시키고, 나아가 배포할 수 있는 플랫폼을 가리킨다. 일례로 Open AI는 Microsoft의 Azure 클라우드 인프라를 이용해 GPT와 같은 생성형 AI 기술을 연구하는 동시에, 자사의 챗 GPT를 MS의 클라우드를 통해 제공 및 작동하게 한다. 클라우드와 같은 인프라는 대규모 자본 및 데이터베이스를 기반으로 하기에 글로벌 빅테크 기업들이 장악하고 있는 시장인데, 대표적으로 Google Cloud, Microsoft Azure, Oracle, AWS(Amazon Web Service) 등이 있다.

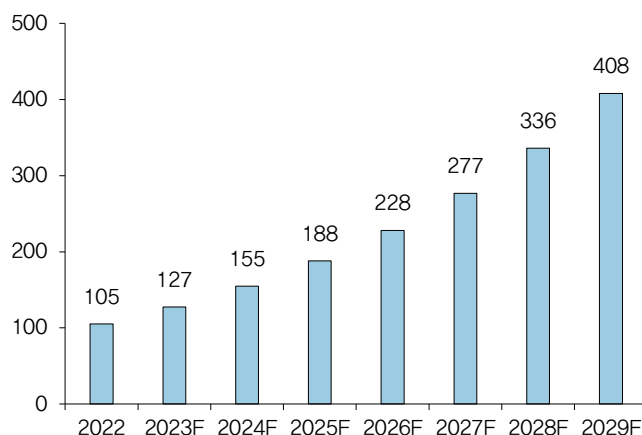
도표 7. 생성형 AI 밸류체인

어플리케이션	B2C: 챗봇, 음성비서, 광고 B2B: 산업 솔루션, 기업 솔루션
파운데이션 모델	GPT(Open AI), PaLM(구글), LLaMA(메타)
인프라	클라우드: AWS(Amazon Web Service), Google cloud 네트워크
AI 반도체	GPU/NPU/TPU/HBM

자료: 산업자료, 성균관대학교 금융투자학회 S.T.A.R

도표 8. 글로벌 LLM 시장 규모 전망

(단위: 억 달러)



자료: Valuates Report, 성균관대학교 금융투자학회 S.T.A.R



# Part 1. 생성형 AI 산업의 현황과 흐름

## 파운데이션 모델, 생성형 AI는 여기서 시작!

### 밸류체인 No.3 파운데이션 모델

파운데이션 모델, 응용/적용이 가능한 범용적 인공지능경망

파운데이션 모델의 정의는 대규모 raw data를 비감독 학습시킨 범용적인 인공 신경망인데, 간단히 말해 LLM과 같이 챗봇 등에 응용 가능한 AI 기본 모델을 의미한다. 파운데이션 모델은 생성형 AI 산업의 핵심이라 볼 수 있는데, 이를 기반으로 다양한 애플리케이션을 구축하여 구체적인 콘텐츠를 생성하고, 전반적인 글로벌 시장 곳곳에 침투하여 무한히 가치를 창출해낼 수 있는 근간으로 작동한다는 점에 기인한다. 일례로 Open AI는 독자 개발한 파운데이션 모델인 GPT-3과 GPT-4를 기업들이 접목해 각종 AI 서비스를 제공할 수 있도록 하는 ‘챗GPT 플러그인’을 출시했는데, 23년 5월 기준 약 823개 파트너사가 입점되어 있다. Open AI의 GPT 파운데이션 모델을 통해 823개의 새로운 서비스가 파생된 것으로, 챗봇인 ‘챗GPT’, 문자 생성 AI 플랫폼인 ‘Jasper AI’, ‘Copy.AI’, 온라인 여행 에이전시 ‘익스피디아’, ‘카약’, 자산관리 서비스를 제공하는 ‘모건스탠리’등에 적용되고 있다. [도표 9]

### 스타트업의 성지, LLM 산업의 높은 성장성

LLM 시장 규모 Cagr +21.4% 성장

현재 Open AI, Google과 같이 우리에게 친숙한 기업뿐만 아니라 Cohere, Anthropic, AI21 Labs, MosaicML 등과 같은 다양한 스타트업들이 LLM 개발에 뛰어들고 있다. 23년 현재 500여개가 넘는 LLM이 개발되고 있는데, Valuates Reports에 따르면, LLM 시장 규모는 22년 105억 달러에서 연평균 +21.4%로 성장해 29년에는 408억 달러에 이를 것으로 전망된다. [도표 8]

도표 9. Open AI API 기반 비즈니스 확장하는 기업 사례

(단위: 십억달러)

기업	상장 여부	시가총액		도입 현황
Microsoft	MSFT US	1,941	검색엔진 마이크로소프트 365	Bing에 GPT 적용해 현재 테스트 중 오피스 제품군에 GPT 모델로 다양한 기능 추가
			코파일럿	협업툴 Teams 내에서 마케팅 및 업무 관련 자료 생성
Duolingo	DUOL US	5	클라우드 Azure	GPT, DALL-E 등의 AI 모델을 클라우드 내에서 바로 활용할 수 있도록 함
			DuolingoMax, Roleplay	언어 학습앱 내에서 AI 챗봇으로 GPT-4 활용
Be My Eyes	비상장	NA	Be My Eyes	시각장애인에게 이미지를 텍스트로 전환해주는 기능에 활용
Stripe	비상장	NA	Stripe DOCS	Stripe 고객에게 요약 및 챗봇 서비스 제공
Morgan Stanley		151	자산관리서비스	자산관리를 위해 방대한 지식 기반을 구성하는데 활용 AI 챗봇 통한 정보 검색
Iceland 정부	정부기관	NA	언어보존 위한 협업	Iceland 언어 보존을 위해 Open AI와 협업
Snap	SNAP US	16	스냅챗	GPT 기반 ‘마이 AI’ 통해 대화 가능
Salesforce	CRM US	183	Slack	주제 및 미팅 요약, 이메일 등에 활용
			CRM	고객관계관리 솔루션에 GPT 활용한 ‘아인슈타인 GPT’ 제공할 계획
Hubspot	HUBS US	19	CRM	Chat Spot 도입해 문서생성기능 제공
Apple	AAPL US	2,414	애플워치	watchGPT 기능 런칭 준비
Grammarly	비상장	NA	GrammarlyGO	문법 도우미 업체로, 3,000만명의 사용자를 대상으로 베타 버전 제공 계획

자료: Open AI, 각 사, 성균관대학교 금융투자학회 S.T.A.R

# Part 1. 생성형 AI 산업의 현황과 흐름

## 네카오도 탑승한 글로벌 LLM 열풍

네이버 HyperCLOVAX  
카카오 KoGPT

국내 업체들도 LLM과 같은 생성형 AI 투자에 공격적인 태도를 보이고 있다. 네이버는 자사의 초 대규모 언어모델인 'HyperCLOVAX' 기술을 기반으로 제작한 대화형 챗봇인 'CLOVA X'를 개발했으며, 카카오는 초대형 언어모델인 'KoGPT'를 출시했다. 'KoGPT'는 Open AI GPT-3의 한국어 특화 모델로, 이를 바탕으로 23년 하반기 챗GPT와 같은 'Ko챗GPT'를 출시할 예정이다. 향후 카카오는 'KoGPT'를 카카오 모빌리티에 접목할 방침으로, 카카오 T를 통해 '단톡방에서 대화를 나누며 언급되었던 모임 장소까지의 추천 경로 제시 및 택시 호출'을 한번에 할 수 있는 생성형 AI 서비스를 제공할 것으로 기대된다.

국내 생성형 AI 투자는  
지속 확대중

한편 23년 8월, 국내 AI 스타트업인 'Upstage'가 개발한 LLM 파운데이션 모델 'SOLAR(솔라)'가 세계 최대 머신러닝 플랫폼인 Hugging face의 '오픈 LLM 리더보드' 평가점수에서 72.3점을 획득하며, GPT-3.5(71.9점)의 성능을 상회하며 글로벌 1위를 차지한 바 있다. [도표 10] '오픈 LLM 리더보드'는 생성형 AI 모델 성능 평가의 기준 지표로 사용되는데, 생성형 AI 모델의 1) 추론 능력, 2) 상식 능력, 3) 언어이해 종합 능력, 4) 환각\* 방지 정도와 같은 4가지 지표를 기준으로 평가한다. 23년 7월 엔씨소프트가 국내 게임사 가운데 처음으로 독자 개발한 '바르코 LLM' 출시, 코난테크놀로지의 기업용 '코난 LLM' 출시, SK텔레콤의 Anthropic 1억 달러 투자 등 국내 생성형 AI 파운데이션 모델 투자가 활발히 일어나는 가운데, 향후 글로벌 AI 산업 성장에 따른 수혜가 기대된다.

## [생성형 AI의 그림자, 환각현상(Hallucination)]

환각현상.  
불확실한 정보를  
사실처럼 제시  
→ 해결 노력중

환각현상: 생성형 AI의 환각현상(Hallucination)은 AI가 사실이 아니거나 불확실한 정보를 마치 사실인 것처럼 제공하는 것이다. 이는 학습과정에서의 인종적/성적 편향성과 더불어 생성형 AI의 윤리적 이슈 중 하나로 거론되고 있다. 이에 따라 LLM 파운데이션 모델 개발 기업들은 환각현상을 해결하기 위해 노력하는데, 일례로 Anthropic은 자사 모델 'Claude 2'에 'Constitutional AI' 라는 기술을 적용했다. 일반적으로 생성형 AI 파운데이션 모델 개발사는 LLM이 유해하거나 비윤리적인 응답을 할 가능성을 낮추기 위해 사람의 피드백을 통해 모델을 강화 학습시키는 RLHF(Reinforcement Learning from Human Feedback)이라는 방식을 사용한다. 이때 Anthropic은 사람이 직접 개입하지 않고, Constitution이라고 불리는 기본 규칙을 제공하기만 하면 AI 모델이 스스로 피드백을 주면서 모델을 개선시키는 기법을 활용한다. 이처럼 사람의 개입이 최소화되며 AI 모델 스스로 지속적으로 환각을 개선할 수 있기에 기존의 RLHF 대비 높은 유용성과 낮은 유해성을 특징으로 한다.

## LLM, Generalist를 넘어 Specialist로 거듭나는 중

최근 특정 목적에  
특화된 LLM이 등장  
→ Hippocratic AI,  
Harvey

한편 최근에는 기존의 범용적인 LLM과는 달리, 특정 목적에 특화된 LLM이 등장하기 시작했다. 일례로 'Hippocratic AI'는 의료 분야에 특화된 LLM을 개발한다. 당사는 AI 연구원, 의료계 종사자들이 설립하였으며, 'HGI(Health General Intelligence)' 실현을 통해 의료 접근성을 높여 평등한 의료 체계를 구축하는 것을 목표로 한다. 해당 모델은 안전을 위해 신뢰도가 높은 의료 데이터베이스를 학습하며, 지속적으로 의료계 종사자들의 피드백을 받으며 강화 학습한다.

'Harvey'는 법률 분야에 특화된 LLM으로, 구글/메타 출신의 AI 연구원, 미국 대형 로펌 출신 변호사가 공동 설립했다. Harvey는 Open AI의 GPT에 추가적으로 판례와 같은 법률 데이터를 학습시킨 LLM 모델로, 1) 법률 문서 검토 및 편집, 2) 유사 판례문 검색, 3) 계약서 작성 및 분석 서비스를 제공하며 법률 업무를 보조한다.

Hippocratic AI나 Harvey의 공통점은, 아직까지는 이들이 직접 진단을 하거나 판결을 내리지 않고 전문가의 업무를 보조 및 보완한다는 것인데 향후 생성형 AI 모델의 정확성과 정교함이 상승함에 따라 이들의 영역이 확장되어 나갈 것으로 판단된다.

# Part 1. 생성형 AI 산업의 현황과 흐름

## IGM, 텍스트를 넘어 이미지/비디오 생성에 박차를 가하다

이미지 생성용  
파운데이션 모델.  
IGM

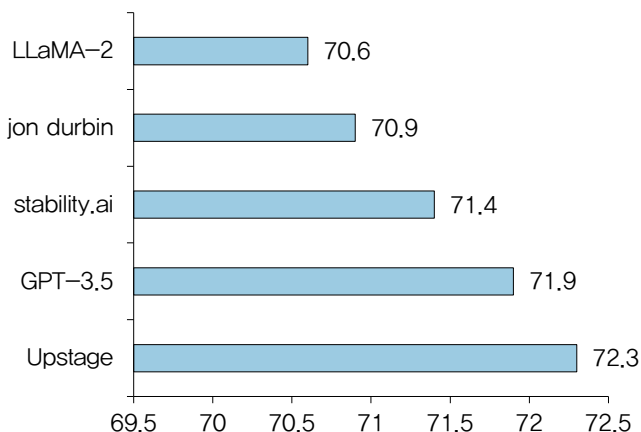
파운데이션 모델에 있어 LLM 다음의 트렌드로 ‘이미지 생성 모델(Image-Generation Model, IGM)’이 부각되고 있다. LLM은 텍스트 명령어 입력을 통해, 질의응답, 문서 번역 및 요약 등의 텍스트 기반 결과물 생성을 주로 했다면, IGM은 ‘Text-to-Image/Video/Motion’ 나아가 ‘image-to-image’와 같이 사용자의 텍스트 명령어에 부합하는 이미지와 영상을 생성하는 AI 모델로 이해할 수 있다. 기술적으로는 LLM에서 파생된 멀티모달(단일 형태가 아닌 텍스트, 이미지, 음성 등 다양한 종류의 데이터를 동시에 처리) 파운데이션 모델이라고 설명할 수 있다. 대표적인 글로벌 IGM으로는 Open AI의 DALL-E, Stability AI의 Stable Diffusion, Midjourney의 Midjourney 등이 있다. 현재 Stable Diffusion은 오픈소스가 공개되어 있기에 파생된 서비스가 빠르게 출시되고 있는데, 이와 같은 IGM 모델은 영화, 광고, 게임 등과 같은 예술 산업 전반에 침투해 나갈 것으로 전망된다.

## 밸류체인 No.4 어플리케이션

어플리케이션,  
파운데이션 모델을  
탐재하고 멀리 나아갈 때

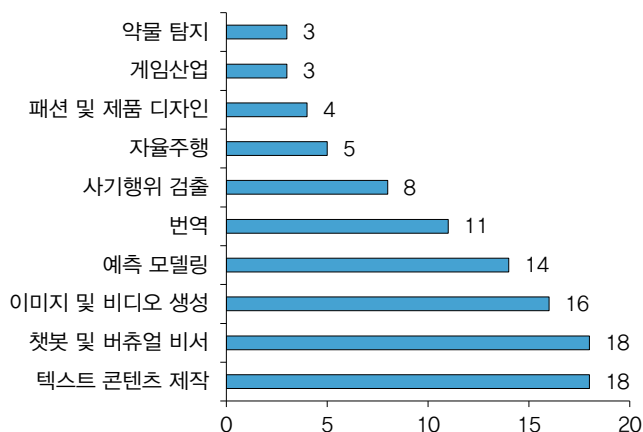
생성형 AI 밸류체인의 정점에 위치한 어플리케이션은 밸류체인 내에서 가장 빠르고, 무한히 성장할 수 있는 영역이라고 판단된다. (여기서의 어플리케이션은 생성형 AI 적용/응용처라는 의미로 이해할 수 있다) 이는 1) 기존의 인프라, 파운데이션 모델을 파인튜닝(활용 목적에 맞게 모델을 변형하는 기술력)하여 서비스를 빠르게 구축할 수 있으며, 2) 비용/기술력/자금력 측면에서 진입장벽이 낮고, 3) 적용처가 무한하고, 응용성이 높다는 점에 기인한다. 이러한 생성형 AI 어플리케이션은 기업이 GPT-3.5, LaMDA, LLaMA와 같은 파운데이션 모델을 자사의 비즈니스, 서비스 등에 맞춤형으로 내재화하는 것에서부터, DALL-E나 Stable Diffusion과 같이 AI 모델을 활용해 새로운 비즈니스 모델을 창출하는 것 등이 될 수 있다. 향후 LLM과 같은 파운데이션 모델이 정교화되고, 접근 비용이 낮아짐에 따라 생성형 AI가 글로벌 산업/기업으로 빠르게 침투해 나갈 것이라 기대된다. [도표 11]

도표 10. 23년 8월 기준 허깅페이스 내 LLM 순위 (단위: 점)



자료: Hugging Face, 성균관대학교 금융투자학회 S.T.A.R

도표 11. 생성형 AI 적용 트렌드 (단위: %)



자료: Nasscom, 성균관대학교 금융투자학회 S.T.A.R

## Part 2. 생성형 AI, 왜 주목해야 할까

생성형 AI는 꺼지지 않을 불꽃

인공지능 전쟁의 최선봉에 있는 Chat GPT가 불러온 “생성형 AI 모먼트”는 그야말로 혁명이었다. Open AI의 Chat GPT는 5일만에 100만 사용자를 돌파하였으며, 월간 활성 사용자 수 1억 명 달성에는 2개월 밖에 소요되지 않았다. 이러한 생성형 AI는 잠깐의 트렌드일까? 그렇지 않다. 생성형 AI는 당분간 반짝하는 테마로 끝나지 않는, 꺼지지 않는 불꽃이 될 것이라 판단된다. 우리는 AI의 1)압도적 생산성, 2)고용시장에 미칠 파급효과와 3)더욱 가속화될 침투에 따라 생성형 AI에 주목해 볼 필요가 있다.

### 생성형 AI의 압도적 생산성

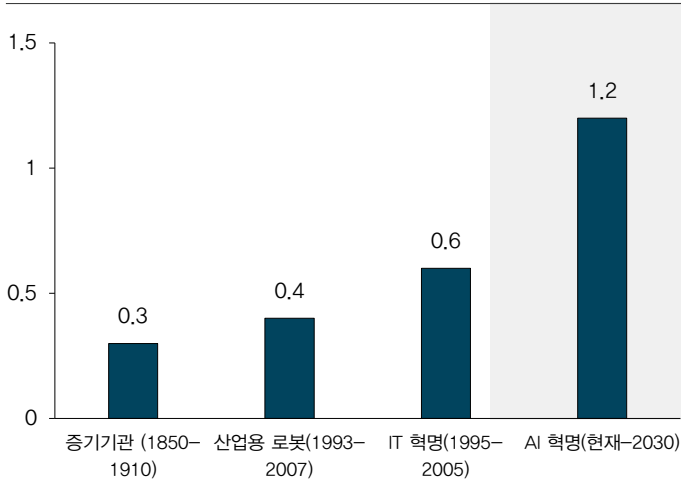
#### 1. 압도적 생산성

생성형 AI에 주목해야 하는 첫번째 이유는 높은 생산성에 따른 경제적 효과에 기인한다. 생성형 AI는 생산성, 편의성, 범용성이라는 세 가지 카테고리를 동시에 충족시킨다. 누구나 손쉽게 검색어 입력을 통해 목적에 맞는 다양한 결과물들을 빠르게 얻어낼 수 있는 것이다. 특히 생성형 AI가 가진 생산성이 가히 압도적인데, 이는 과거의 증기기관, 로봇화보다도 더 높을 것이라 기대된다. McKinsey 조사에 따르면, 로봇화 및 IT혁명과 같은 이전의 혁신은 연간 0.3%~0.6 % 범위의 생산성 혁신을 가져온 반면, AI는 과거 혁신의 2배~4배에 이르는 연간 1.2%의 생산성 향상을 제공할 잠재력이 있다고 평가하였다. [도표 17]

생산성 향상  
→ 인플레이션 억제 효과 존재!

역사적으로 생산성 향상 효과가 인플레이션 억제 효과가 있었다는 점을 고려해본다면, AI가 우리에게 가져다 줄 극적인 변화를 기대해볼 만 하다.[도표 18] AI 기술 도입이 GDP 성장에도 긍정적인 영향을 미칠 것으로 예상되는데, PwC에 따르면 AI 기술의 도입으로 2030년까지 전 세계 GDP가 약 15조 7천억 달러가 증가할 것이라 예상된다.

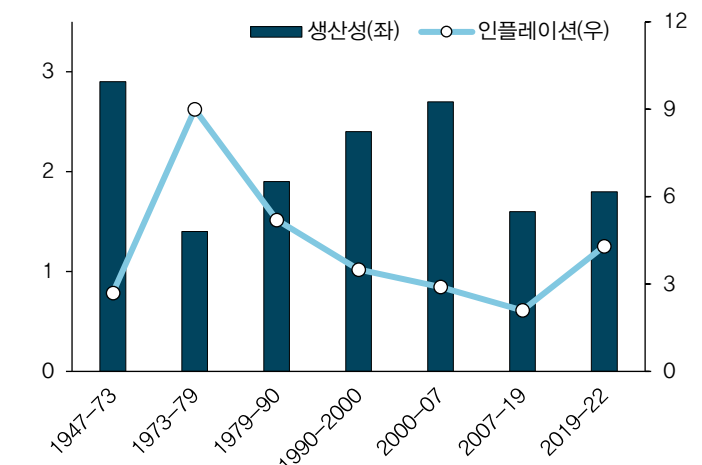
도표 17. 기술 혁신에 따른 생산성 향상 효과



자료: Mckinsey, 성균관대학교 금융투자학회 S.T.A.R

도표 18. 인플레이션과 생산성 간의 역사적 추이

(단위: %)



자료: Bloomberg, U.S. Bureau of Labor Statistics, 성균관대학교 금융투자학회 S.T.A.R

### 고용시장에 미칠 파급효과

#### 2. 고용시장에 미칠 파급효과

AI에 주목해야 하는 두번째 이유는 고용시장에 미칠 파급효과 때문이다. 과거에도 새로운 기술 등장에 따른 일자리 대체와 새로운 일자리 창출이 존재하였다. 다만 AI 기술은 물리적인 노동력 뿐만 아니라 인간의 고유영역으로 여겨지던 인지 및 판단 영역을 대체할 수 있다는 점에서 과거의 기술 혁명들과 큰 차이를 보인다. 인지와 판단능력을 갖춘 AI의 등장에 따라, 현재 사람이 하고 있는 대부분의 직업이 대체 위협에 노출될 수 있는 것이다. World Economic Forum의 연구에서 역시 AI 도입에 따라 2025년까지 전 세계적으로 1,330만 개 이상의 새로운 직업이 생기나, 8,470만 개의 직업이 사라지는 변화가 생길 것이라 밝혔다. 고용시장에 있어 AI가 미칠 파급효과는 그 어떤 기술 진보의 경우보다 거대할 수 있으며, 이에 지속적으로 주목해볼 필요가 있다.

## Part 2. 생성형 AI, 왜 주목해야 할까

### 더욱 가속화될 침투

#### 3. AI의 침투 가속화

마지막으로 더욱 가속화될 AI 침투 가능성에 기인한다. 글로벌 시장조사기관 그랜드뷰리서치에 따르면, 글로벌 생성형 AI 시장 규모는 22년 230억 달러에서 2030년 1,093억 달러로, 연간 34.6%의 높은 성장률을 보이며 성장할 것이라 전망된다.**[도표 19]** 이러한 성장은 1)성능 향상과 응용서비스의 성장, 2)AI 업종에 대한 지속적인 투자와 개발, 3)모델의 학습 비용 감소에 기인한다.

#### 1) 성능 향상과 응용서비스의 성장

##### AI 모델의 성능 향상

생성형 AI 모델의 성능 향상과 응용 서비스의 성장은 생성형 AI의 상용화를 가속화할 것이라 기대된다. 생성형 AI 모델의 성능은 모델의 크기 증가와 함께 지속적으로 향상되고 있다. 23년 3월에 발표한 GPT-4.0 역시 GPT-3.5 대비 데이터 처리 능력이 8배 이상 향상된 업그레이드 버전이다. 거짓된 답변을 하는 hallucination 현상도 40% 이상 개선되었으며, 처리할 수 있는 단어 수도 3.2만개로 기존보다 0.8만개 늘어나 용량이 큰 문서 처리와 깊이 있는 분석을 할 수 있게 되었다. 이미지 입력 역시 가능하며, 총 26개국 언어를 지원하는 다국어 서비스가 개선되어 사용자의 효용을 크게 향상시키고 있다.

##### 응용 서비스의 성장 : 1) 기존 검색 엔진과의 결합

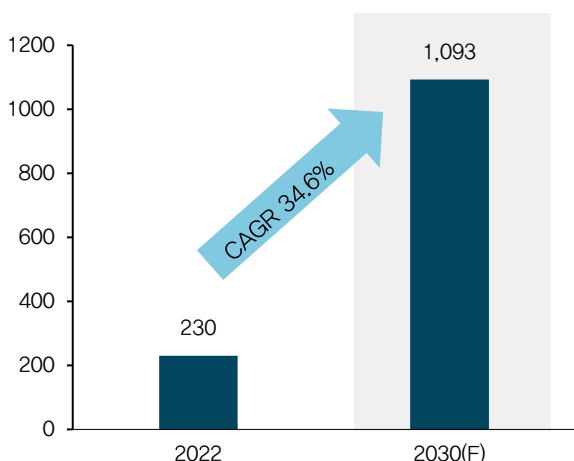
기존 검색 엔진과의 결합은 정확성을 크게 높여, AI의 침투를 앞당길 것으로 판단된다. 대표적으로 마이크로소프트는 자사의 검색 엔진인 Bing과 Chat GPT를 결합시킨 Bing+ 서비스를 출시하였다. 기존의 Chat GPT는 21년까지의 데이터만으로 학습하여 최신 데이터를 활용하기 어려우며, AI 모델 특성상 답변의 정확성을 담보하기 어렵다는 문제가 존재하였다. 다만 자사의 검색 엔진을 통해 실시간 데이터를 처리함으로써 최신의 데이터를 활용할 수 있게 되었고, 가장 연관성이 높은 답변과 이에 대한 출처를 제공함으로써 답변의 정확성을 높였다.

##### 응용 서비스의 성장 : 2) 플러그인 도입

Plugin 도입에 따른 효과 역시 기대해볼 만하다. Plugin이란 chat GPT에 제 3의 어플리케이션을 연동하여, Chat GPT가 최신 정보에 액세스하여 계산을 실행하거나 타사 서비스를 사용할 수 있도록 도와주는 기능이다. 현재 익스피디아(호텔 예약), 피스칼노트(법률 데이터), 오픈테이블(식당 예약), 스픽(외국어 교육) 등의 앱에서 가능하며**[도표 20]**, 향후 연동 앱 개수 확대에 따른 효과가 기대된다.

도표 19. 생성형 AI 시장 전망

(단위: 억 달러)



자료: 그랜드뷰리서치, 성균관대학교 금융투자학회 S.T.A.R

도표 20. Plugin 어플리케이션 예시

어플리케이션	설명
Expedia	호텔 예약과 관련하여 Expedia 정보를 기반으로 답변 제공
Open Table	레스토랑 예약과 관련된 정보 제공
Wolfram Alpha	수학, 물리학, 금융 등 전문 영역에 대한 질의응답 가능
Link Reader	링크를 제공하고, 링크 내용에 대해 요약 등 질문 가능
VokScript	유튜브 링크를 제시하면, 동영상의 스크립트를 제공해줌

자료: 성균관대학교 금융투자학회 S.T.A.R

## Part 2. 생성형 AI, 왜 주목해야 할까

### 2) AI 업종에 대한 지속적인 투자와 개발

스타트업 투자 확대 & 빅테크 경쟁 구도 심화

AI 업종 전반에 대한 투자 확대와 빅테크 기업의 경쟁 구도 역시 AI 침투에 긍정적이다. 글로벌 유동성이 위축되는 상황에서도 생성 AI 스타트업에 대한 투자는 활발하였다. 피치북의 데이터에 따르면 2020년 이후 생성 AI에 대한 투자는 425% 증가하였고, 2022년에는 약 21억 달러 이르렀다. [도표 21] 특히 생성형 AI를 둘러싼 글로벌 빅테크들의 경쟁이 갈수록 뜨거워지고 있다. 그동안 AI 기술 개발에 소극적인 모습을 보였던 애플마저 뒤늦게 합류하며, 글로벌 빅테크들은 시장 선점을 위해 저마다 인프라 구축에 힘을 쏟고 있다. 산업 내에서 우위를 점하기 위한 지속적 투자와 개발은 AI 침투를 앞당길 것으로 기대된다.

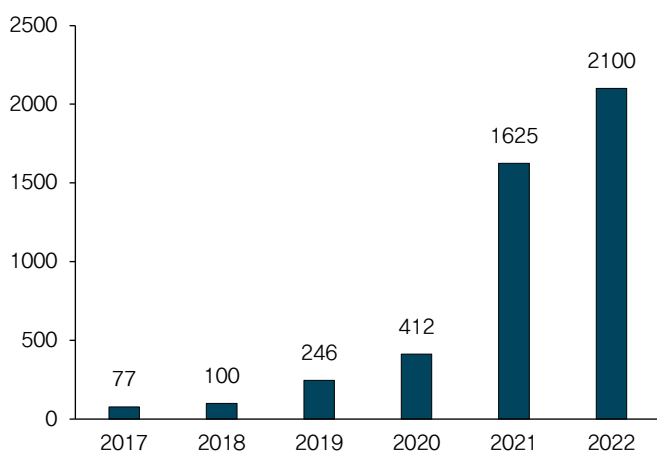
### 3) 모델의 학습 비용 하락

모델 학습 비용 하락

생성형 AI 학습 비용 하락 역시 AI 침투 가속화에 기여할 것으로 기대된다. 이러한 학습 비용 감소는 오픈 소스 모델 이용과 코딩 효율성 개선 등에 기인한다. 20년 Open AI의 GPT-3, 23년 Meta의 LLaMA, 23년 Stanford의 Alpaca의 학습비용 추이에서 감소 폭이 두드러진다. [도표 22] 심지어 2030년에는 AI 훈련에 30달러 가량밖에 발생하지 않을 것으로 기대되며, 이는 산업 전반에서 AI를 받아들이는 속도를 앞당길 것이라 판단된다.

도표 21. 생성 AI 스타트업 투자 금액 추이

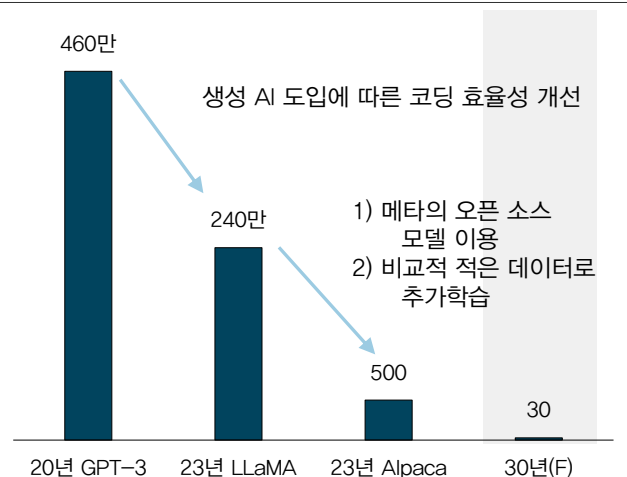
(단위: 백만달러)



자료: CB INSIGHT, 성균관대학교 금융투자학회 S.T.A.R

도표 22. 생성 AI 훈련 비용 추이

(단위: 달러)



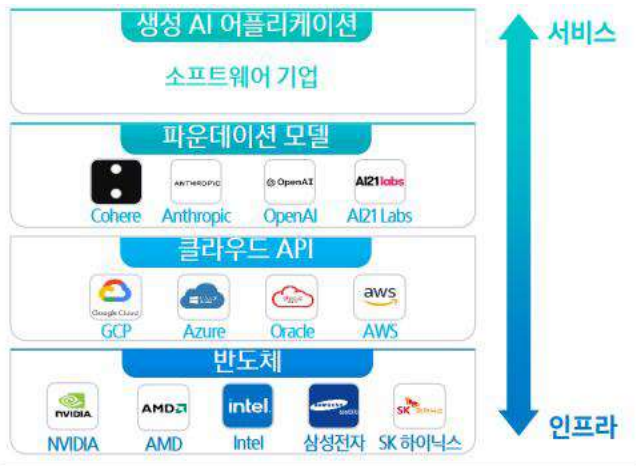
자료: ARK Invest ideas 2023, 성균관대학교 금융투자학회 S.T.A.R

# Part 3. 생성형 AI, 어디를 봐야 할까

## 생성형 AI 밸류체인

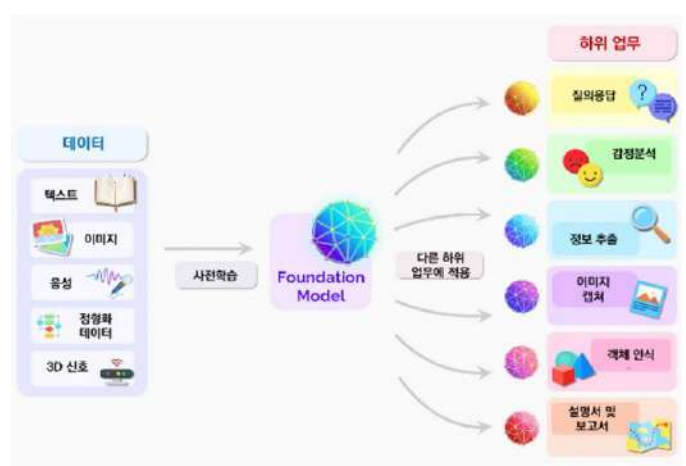
지금까지 생성형 AI의 성장에 대한 가능성을 타진했다면, 이제는 생성형 AI가 본격적인 상용화 단계에 진입하며 발생하는 막대한 부가가치를 누가 가져갈 것인지 살펴보아야 한다. 그러기 위해서는 밸류체인 별 특성에 대한 이해가 선행되어야 한다. 생성형 AI의 밸류체인은 **[반도체-클라우드- 파운데이션 모델- 애플리케이션]**으로 구성되어 있다. [도표 23]

도표 23. 생성형 AI 밸류체인 구조도



자료: 성균관대학교 금융투자학회 S.T.A.R

도표 24. 파운데이션 모델 구조



자료: 성균관대학교 금융투자학회 S.T.A.R

## 1. 파운데이션 모델

### 파운데이션 모델

= 생성형 AI의 핵심

파운데이션 모델은 생성형 AI의 원재료이자, 핵심이다. 이는 생성 AI의 기초가 되는 대형 딥러닝 네트워크이며, 다양한 하위 태스크에 적용이 가능하다. [도표 24] 파운데이션 모델은 방대한 양의 폭 넓은 데이터를 사전에 학습한 이후, 사용자가 원하는 용도나 목적에 맞게 미세 조정되거나 문맥 내 학습 과정을 거쳐 활용된다. 특히 자연어 처리, 이미지 생성, 음성 합성 등의 기능을 가지고 있기 때문에, 개발자는 이를 활용하여 시장의 요구에 맞는 다양한 애플리케이션을 개발할 수 있다. Open AI의 GPT와 구글의 Bert가 대표적인 예시이다.

### 성능

= 주요 파라미터 수로 결정됨

파운데이션 모델에서는 경쟁자를 압도하는 성능을 기반으로 생태계를 확장하는 것이 관건이다. 거대언어모델에서 매개변수 개수가 성능에 큰 영향을 미치는 것으로 알려지면서, 2021년을 기점으로 수천억 개의 매개변수가 사용된 모델들이 줄지어 등장하였다. Open AI의 chat GPT-3.5는 1,750억 개, 구글의 PaLM은 5,400억 개의 파라미터를 가지고 있으며, 수 조 억개가 넘는 모델 역시 등장하게 되었다.

개발 비용이 높아, 주로 빅테크의 영역!

이러한 거대 파라미터는 훈련과정에서 상당한 컴퓨팅 작업이 요구되기에, 일반적으로 막대한 개발 비용이 발생하게 된다. 또한 훈련을 위해서는 대규모의 데이터와 하드웨어 리소스가 필요하기 때문에, 주로 이를 감당할 수 있는 소수의 빅테크 기업과 이들에게 투자를 받은 스타트업에 의해 지배되고 있다. 대표적으로 Open AI, Stability AI, Anthropic, Cohere이 존재한다.

### 최근의 흐름은 ?

1) 최근에 소형 언어모델 확산 중

하지만 최근에는 과도한 비용 발생에 따라, 매개 변수를 줄이는 대신 질 좋은 데이터를 학습시키고 파인 튜닝하는 방식의 소형 언어모델 역시 확산되고 있다. 특정 분야에 한정되어 있다는 단점이 존재하지만, 훈련 비용과 시간을 줄이고 정확도를 높일 수 있다는 장점이 있다. 신속하고 저렴하게 만들 수 있는 데다가 보유한 데이터를 활용하여 맞춤형으로 구축할 수 있기 때문에 기업에 도입하기도 유리하며, 기술 상용화에도 좀 더 용이해질 것으로 판단된다.

2) 멀티모달로 발전 중

멀티모달로의 발전도 주목해볼 만하다. 텍스트나 이미지 등 단일 유형의 데이터만 다루는 기존 AI에 비해 멀티모달은 이미지, 텍스트, 음성 등을 종합적으로 처리할 수 있기 때문에 잠재력이 높은 분야이며, 24년부터 멀티모달 기반의 생성형 AI 경쟁이 더욱 뜨거워질 전망이다.

# Part 3. 생성형 AI, 어디를 봐야 할까

## 2. 반도체

생성 AI의 성장 →  
반도체의 수요 증가

복잡한 AI 소프트웨어를 구현하기 위해서는 높은 성능의 하드웨어가 필수적이다. 그 중 반도체는 생성형 AI의 구동에 핵심적인 역할을 수행한다. 대형언어모델의 학습과 추론에는 무수히 많은 연산이 필요한데, 이를 가능하게 하는 것이 엄청난 수량의 고효율 반도체 칩셋이다. 대표적으로 1,850억 파라미터의 GPT 3.5chat GPT 학습에 엔비디아의 A100 GPU 1만 개 넘게 사용되었으며, 상용화를 위해서는 약 3만 개의 GPU가 요구된다. Chat GPT를 시작으로 다수의 빅테크 기업들이 인공지능 서버와 슈퍼컴퓨터 투자에 전력을 다하고 있으며, 이는 반도체 수요 증가를 주도하고 있다. 이처럼 생성형 AI 시장의 활성화는 AI 반도체 시장의 동반 성장을 견인할 것이라 기대되며, 가트너에 따르면 글로벌 AI 반도체 시장 규모는 23년 534억 달러에서 27년 1,194억 달러로 성장할 것으로 전망된다. [도표 25]

### 생성 AI에서 가장 핵심은 GPU !

WHY GPU 가 핵심 ?  
병렬적 데이터 연산 방식 때문 !

AI가 본격적으로 발전하면서 가장 각광을 받은 반도체는 GPU이다. 이는 생성형 AI 연산에 GPU의 병렬적 데이터 연산 방법이 기능과 속도, 정확성 측면에서 적합하기 때문이다. 대규모의 데이터와 파라미터를 다루기 위해서는 특화된 연산을 수없이 계산해야 되는데, 이를 얼마나 빠르게 수행하느냐가 성능 차이이다. 이러한 상황에서 기존의 단일 코어 중심의 CPU는 대량의 반복 학습을 한꺼번에 처리하기에 속도의 한계가 존재한다. 반면 GPU는 다수의 코어를 통해 한 번에 많은 데이터를 계산하는 병렬 처리 방식을 가졌기 때문에, 빠른 속도로 방대한 양의 데이터를 훈련시켜야 하는 AI 연산에 적합하다. [도표 26]

### Nvidia 가 제일 잘 나가 ~

NVIDIA 가 GPU의 90%  
독점

이러한 생성형 AI와 GPU의 성장에 가장 큰 수혜를 본 것은 Nvidia이다. 이는 엔비디아가 독보적 경쟁력을 바탕으로 AI GPU 시장의 90% 가량을 점유하며, 생성 AI 확대에 따른 수혜를 온전히 받았기 때문이다. Open AI의 Chat GPT 역시 개발과 학습 과정에서 엔비디아의 GPU 수만개를 연결한 슈퍼컴퓨터를 이용하였다.

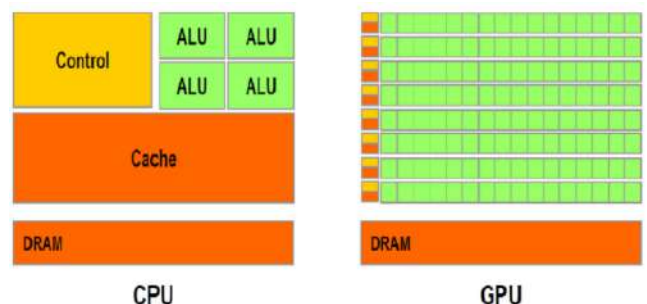
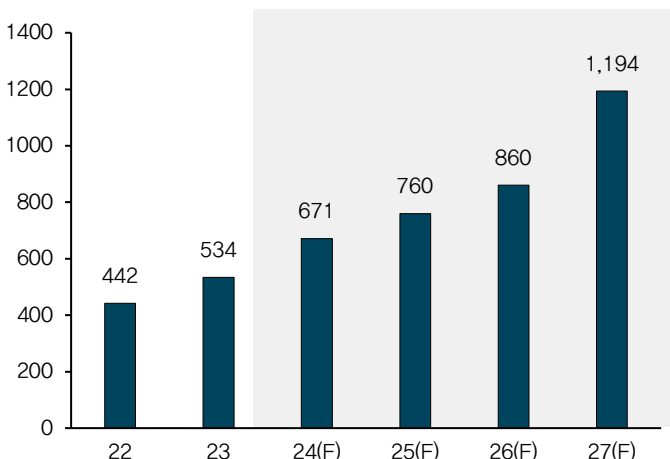
너무 잘나가는 NVIDIA

엔비디아는 AI를 도입하려는 기업을 중심으로 반도체 주문량이 늘면서 시장 침체에도 불구하고 높은 실적을 이어가고 있다. 생성형 AI 시장의 확대에 따라 23년 2분기 기준 매출액 135억 1000만 달러로 전년동기비 100% 이상 증가하였으며, 영업이익 역시 1,263% 늘어난 68억 달러를 기록하였다. 이는 매출과 이익 양 측면에서 시장의 컨센서스를 20-30% 상회한 수치이며, 올해 주식시장에서 주가가 200% 넘게 상승하며 시장의 주목을 한 번에 받았다.

도표 25. AI 반도체 시장 전망

(단위: 억 달러)

도표 26. CPU vs GPU



자료: 가트너, 성균관대학교 금융투자학회 S.T.A.R

자료: NVIDIA, 성균관대학교 금융투자학회 S.T.A.R



# Part 3. 생성형 AI, 어디를 봐야 할까

## NVIDIA의 성공 비결은 ?

무엇이 성공 비법일까 ?

이러한 엔비디아의 압도적 지위는 1) CUDA를 기반으로 한 강력한 생태계 2) 세계 최고 수준의 GPU 설계 기술력 3) 하드웨어/소프트웨어 수직계열화에 기인한다. 특히 엔비디아는 AI GPU에 최적화된 CUDA 소프트웨어를 앞세워 생태계를 장악하고 있다.

일등공신은 CUDA.

CUDA는 C언어 등 다양한 프로그래밍 언어를 이용하여 GPU의 병렬처리 알고리즘을 쉽게 작성할 수 있도록 개발된 프로그램이다. 엔비디아는 GPU 시장 개화 초기 개발자들에게 CUDA 플랫폼을 제공하여, 현재 대부분의 AI 알고리즘은 CUDA 플랫폼을 기반으로 이루어지게 되었다. 이 아키텍처는 엔비디아의 GPU에서만 작동하기 때문에, 일종의 락인 효과가 발생하여 구글 메타가 제공하는 프레임 워크는 물론 각종 라이브러리, 도구 등이 엔비디아의 GPU에 의존하고 있는 상황이다. 이에 따라 단기적으로 AI 모델 학습용 인프라에서 엔비디아 GPU 대체는 사실상 불가능하며, 엔비디아의 독점적 지위는 당분간 이어질 것으로 판단된다.

## 여전히 공급은 부족하고, 가격은 비싸고 ..

여전히 공급이 부족하다 ..

이러한 엔비디아의 독점적 지위는 GPU의 공급부족과 가격상승으로 이어지게 되었다. 생성형 AI 확대에 따른 타이트한 수급 영향으로 엔비디아의 H100의 한대 가격은 4만 달러까지 상승하게 되었다. 그럼에도 여전히 마이크로소프트를 시작으로 글로벌 빅테크 기업들은 엔비디아의 GPU를 조달하기 위해 노력 중이며, 아직까지도 공급이 수요를 따라가지 못하고 있다.

## 새로운 AI 반도체 ,, NVIDIA GPU 넘어설 수 있을까 ?

NVIDIA GPU 대체하기 위해 새로운 NPU들 등장 중 ! 하지만 아직은 ,,

이에 따라 엔비디아의 칩을 대체하기 위해 AMD의 MI시리즈, 인텔의 가우디까지 등장하였다. Google의 TPU (tensor processing unit) 사례와 같이 자금력이 풍부한 업체들은 직접 반도체 설계에 나서기도 하였다. 범용성을 추구하는 엔비디아의 GPU와 달리, 이들의 솔루션은 AI에만 특화되어 있다 보니 일부 데이터의 경우 엔비디아의 GPU보다 뛰어나기도 하다. 주로 NPU는 인공지능 개발에 최적화된 설계를 갖추고 있기에, 동급의 GPU보다 처리 속도도 높고, 우수한 전성비의 장점을 가지고 있기도 하다. 다만 아직까지 부족한 검증 사례, 기존 제품과의 호환성, 에코 시스템, 비용 등의 문제에 따라 급격히 성장하는 AI 시장의 수혜는 범용 반도체 GPU가 우선적으로 누릴 것으로 판단된다.

도표 27. 주요 AI 반도체 특징

반도체	전력 소모	특징	장점	단점
CPU	높음	유연성, 범용성	복잡한 명령 수행	메모리 병목 100개 미만의 적은 코어수
GPU	높음	병렬 연산 효율성	고성능 AI 연산 수천개의 코어로 구성	높은 전력 소모
FPGA	중간	프로그래밍 기능	유연함 하드웨어 수정없이 기능 변경	비싼 가격
ASIC	낮음	목적에 최적화	특정 목적 수행에 최적화	한정된 기능 높은 설계 비용
TPU	낮음-중간	목적에 최적화	TensorFlow에 최적화	google에만 귀속 한정된 호환성

자료: ADLINK, 성균관대학교 금융투자학회 S.T.A.R

# Part 3. 생성형 AI, 어디를 봐야 할까

## 추론용 시장이 개화된다면, 이야기는 달라질 수도 !

추론 시장의 성장 → NPU의 성장

다만 향후 추론 연산 시장 성장에 따른 NPU의 성장은 주목해볼 필요가 있다. 인공지능은 연산 방식에 따라 크게 훈련과 추론으로 나눌 수 있으며, 연산의 방식에 따라 적용되는 반도체의 종류와 구조가 다르다. [도표 28] 훈련은 모델을 구축하기 위해 대량의 데이터를 학습시키는 것을 뜻하고, 추론은 학습된 모델을 바탕으로 유저들에게 답변을 제공하는 것을 의미한다.

WHY?  
추론 단계에서는 많은 연산 필요 없거든 ~ NPU면 충분해 !

훈련 단계에서는 여전히 복잡한 계산과 소프트웨어를 필요로 하기 때문에 범용성이 뛰어난 엔비디아의 GPU가 당분간은 절대적 지위를 차지할 것으로 보인다. 다만 서비스를 유저에게 전달하는 추론단계에서는 상대적으로 요구되는 연산량이 적고, 애플리케이션별로 시장이 세분화됨에 따라 반도체는 특정 분야에 특화된 연산만 수행할 수 있으면 된다. 즉, 추론 단계에서는 GPU의 성능까지 필요하지 않다. 이에 따라 범용성은 떨어지나, 가격 효율성과 전성비가 좋은 NPU가 시장을 확대해 갈 가능성이 존재한다. 구글의 TPU 역시 구글 클라우드 등 자체 AI 서비스에 특화되어 높은 성능 구현하는 추론용 칩의 일환이다.

추론용 시장은 잠재력이 큰 시장

초기에는 머신러닝 학습을 위해 ‘훈련용’의 수요가 높았지만, 장기적으로 AI가 상용화되면서는 학습 데이터를 기반으로 AI서비스를 구현하는 ‘추론용’ 수요가 크게 증가할 것이라 판단된다. 이는 추론용 시장이 이미 학습된 모델을 기반으로 수많은 응용분야에 적용될 수 있어 확장성과 다양성 측면에서 향후 잠재력이 크기 때문이다. 이에 따라 추론용 시장은 학습용 시장을 후행하며 성장할 것으로 전망되며, 25년 이후부터 추론용 칩과 NPU의 본격적 수요를 확인할 수 있을 것이라 판단된다. [도표 29]

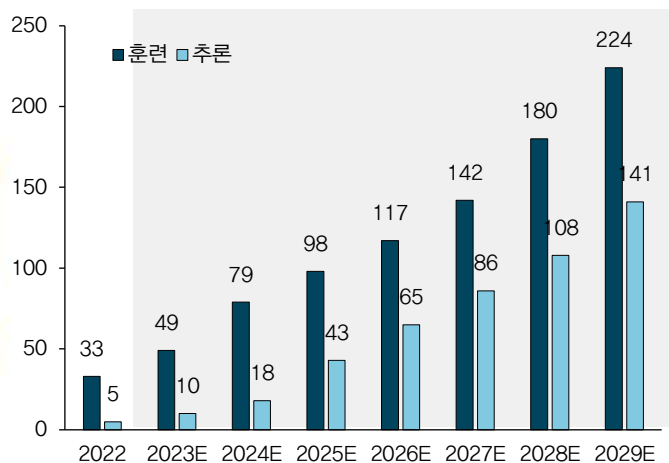
도표 28. 훈련용칩 과 추론용 칩 개념



자료: ETRI, 성균관대학교 금융투자학회 S.T.A.R

도표 29. 훈련용 및 추론용 칩 시장 전망

(단위: 십억 달러)



자료: Boomborg, IDC, 성균관대학교 금융투자학회 S.T.A.R

## HBM 역시 폭발적으로 성장할 것

AI 성장 → HBM 성장

Chat GPT가 불러온 AI 투자 확대는 많은 양의 데이터를 빠르게 처리할 수 있는 HBM 폭발적 수요 역시 유발시켰다. HBM은 여러 개의 D램을 수직으로 연결하여 기존 DRAM 대비 데이터 처리 속도를 대폭 끌어올린 고대역폭 반도체로 [도표 30], 현재 초거대 AI 모델 추론이 HBM의 성능에서 결정지어진다고 해도 과언이 아니다. 현재 HBM은 엔비디아 등 주요 GPU 제조처에 납품되고 있으며, 최근에는 빅테크 기업들이 자체적으로 AI모델과 AI 가속기를 자체 개발하면서 HBM 수요는 다양한 곳으로 확대되고 있다.

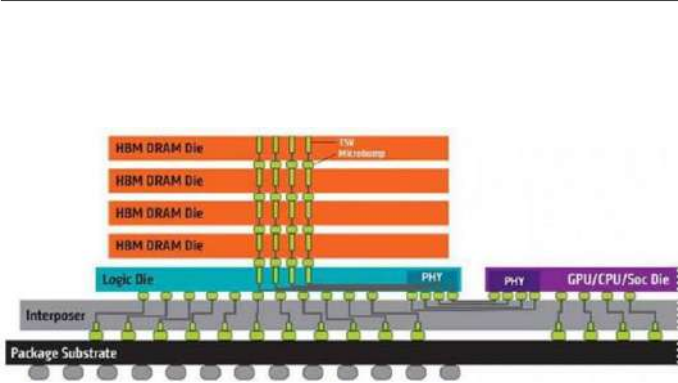
글로벌 HBM 시장은 2022년 23억달러에서 2025년 103억달러로 연평균 64% 성장할 것으로 전망된다. H100의 경우 HBM3 80GB, GH200은 HBM3E 141GB, AMD의 MI300 Max는 HBM3 192GB를 지원하는 등 GPU당 HBM 탑재량이 증가하고 있는 것도 성장 동력으로 작용할 전망이다. [도표 32]

# Part 3. 생성형 AI, 어디를 봐야 할까

## 한국이 HBM 독점 중

HBM은 AI 산업이 성장하는데 필수적인 하드웨어이며, 전세계에서 한국 반도체 기업이 HBM생산의 90%를 차지하고 있다. 현재 글로벌 HBM 수요는 3세대인 HBM2e에서 4세대인 HBM3로 옮겨가고 있는데 [도표 33], 현재 HBM3 양산이 가능한 기업이 sk하이닉스와 삼성전자라는 점에서, 한국의 24년 HBM 시장 점유율 역시 90% 이상에 달할 것으로 전망된다.

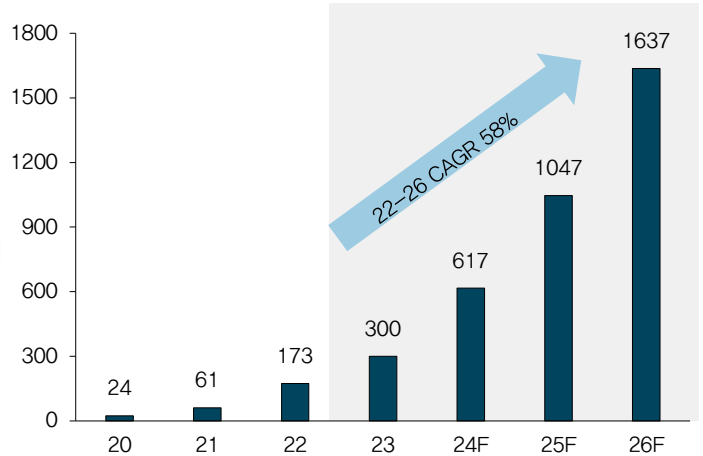
도표 30. HBM 구조



자료: AMD, 성균관대학교 금융투자학회 S.T.A.R

도표 31. HBM 연간 수요 전망

(단위: mGB)



자료: 성균관대학교 금융투자학회 S.T.A.R

도표 32. AI 가속기당 HBM 탑재량

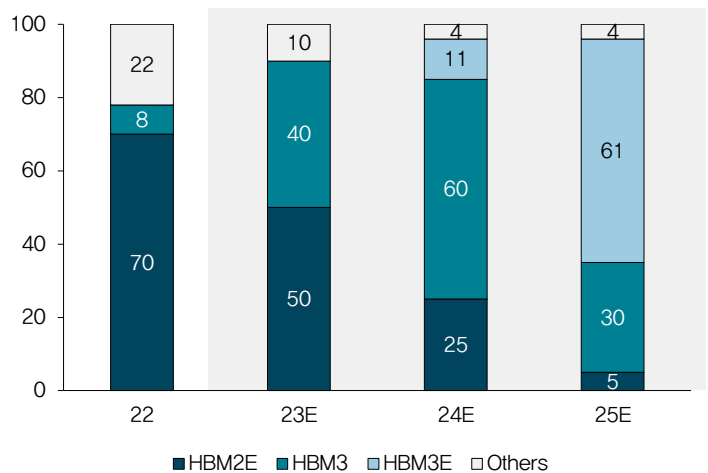
(단위: 명)

AI 칩 공급업체	칩	메모리	메모리 캐파(GB)
Nvidia	GH200	HBM3E	141
	H100	HBM3E	80
	A100	HBM2E	80
	A30	HBM2E	24
AMD	MI300X	HBM3E	192
	MI300A	HBM3E	128
	MI200	HBM2E	128
Intel	Max	HBM2E	128
	Stratix 10	HBM2E	16
Google	TPU v3	HBM2	32
Xilinx	Versal	HBM2E	32

자료: TrendForce 성균관대학교 금융투자학회 S.T.A.R

도표 33. 세대별 HBM 비중 추이 및 전망

(단위: %)



자료: TrendForce, 성균관대학교 금융투자학회 S.T.A.R

## 3. 클라우드

### AI 성장 → 클라우드 성장

AI의 사용증가는 클라우드 컴퓨팅 사용 증가로 이어질 수 밖에 없다. GPU와 TPU가 비싸고 부족한 자원이기에 대부분의 기업들이 대규모 AI모델 구축과 조정 및 실행 작업을 클라우드에서 수행하기 때문이다. 클라우드 인프라를 제공하는 기업은 대표적으로 마이크로소프트(Azure), 알파벳(Google Cloud), 아마존(AWS)가 있으며, 멀티 클라우드 전략이 확산됨에 따라 많은 서버를 보유하고 최적의 서비스를 제공하는 사업자들이 더욱 성장할 것으로 판단된다.

## Part 3. 생성형 AI, 어디를 봐야 할까

### 4. 애플리케이션

정의

애플리케이션은 파운데이션 모델을 활용하여 사용자가 원하는 결과를 더 빠르게 달성할 수 있도록 지원한다. 텍스트, 동영상, 이미지, 코딩 개발, 3D 등으로 분화되어 서비스 이용이 활발하게 진행되고 있으며, 일반적으로 텍스트로 원하는 내용을 주문하면 인공지능이 만들어주는 Text-to-X 형태가 가장 많이 활용되고 있다. 대표적으로 텍스트를 이미지로 바꾸어주는 어도비의 firefly 앱이 있다.

난이도가 낮은 편 so 경쟁이 심해!

애플리케이션 단계의 기업들의 경우 밸류체인 상 인프라 영역과 달리 투자 규모와 원천 기술력에 대한 난이도가 상대적으로 낮은 편이다. 이에 따라 진입 장벽이 낮으며 시장 내 경쟁 구도가 치열한 것이 특징이다. 해당 영역에 경쟁 강도가 높은 만큼 파인 튜닝 능력을 비롯하여 소비자의 니즈를 충족시키는 앱 서비스가 향후 경쟁력이자 핵심일 것으로 판단된다.

### 여전히 인프라, 이제는 애플리케이션까지 보자!

24년에도 인프라는 여전히 중요할 것

AI 테마 초기 가장 강력한 모멘텀을 지닌 것은 인프라 영역이었다. 여전히 GPU 쇼티지에 결정적인 역할을 하고 있는 빅테크들의 수요가 강력하다는 점에서, AI 인프라에 대한 대대적인 투자는 24년에도 지속될 것으로 보인다.

이제는 서비스에 집중해보자

인프라 투자 광풍 이후 시선은 서비스 단으로 이동할 것으로 판단된다. 실제로 지갑을 열어 지속 가능성을 부양해줄 주체는 최종 소비자이기 때문이다. 최종 소비자를 대상으로 한 서비스들이 본격적으로 개화되어야 생성형 AI가 단기적인 테마가 아니라 장기적인 트렌드로 전환될 수 있을 것이다.

이제는 다양한 오픈소스 모델의 등장에 따라 상용화 가능한 수준의 기술 확보가 어느정도 용이해진 상태이기에, 이를 어떻게 활용해서 수익화 할 서비스로 만들어내는지가 중요해지고 있다. 주요 기업들은 기업 가치를 확대하기 위해 경쟁적으로 생성형 AI 제품을 출시하면서 애플리케이션 시장이 급격히 성장할 전망이다. 가트너는 23년 5% 미만에 불과했던 기업의 생성형 AI 애플리케이션 도입이 26년 80%까지 증가할 것이라 전망하였다.

애플리케이션 시장은 가장 빠르게 확장되고 큰 가치 창출 기회를 제공할 것으로 예상되며, 특정한 거나 독점적인 데이터를 사용하여 애플리케이션을 세밀하게 조정하는 기업이 상당한 경쟁우위를 얻을 것이라 전망된다.

## Part 4. AI 응용 산업 맛보기

### AI 응용 산업을 보는 이유

AI는 등장과 동시에 여러 산업 분야에 걸쳐 사용되기 시작하였으며, 이 과정에서 AI는 각 산업이 기존의 틀에서 벗어나 새로운 도약을 하게 해주는 도화선으로서 역할을 충실히 해냈다. 인간이 하는 일을 더 빠르고 정확하게, 하나 더 낮은 가격으로 대체해주는 AI의 근본적인 특징은 여러 산업 내에서 AI가 받아들여지는 데에 결정적인 역할을 했다고 볼 수 있다. 점차 각 산업 군 내에서 AI의 입지가 높아지고 있기에, 최근 생성형 AI의 등장으로 각 산업에는 또 한번 AI로 인한 급격한 변동이 일어날 것으로 전망된다. 심지어는 기존에 AI를 채택하지 않았던 분야로도 생성형AI가 확대될 조짐을 보이고 있다. 이는 본 보고서의 앞 부분에서 언급된 생성형 AI의 높은 편의성과 생산성, 범용성 등의 특징에 기인한다.

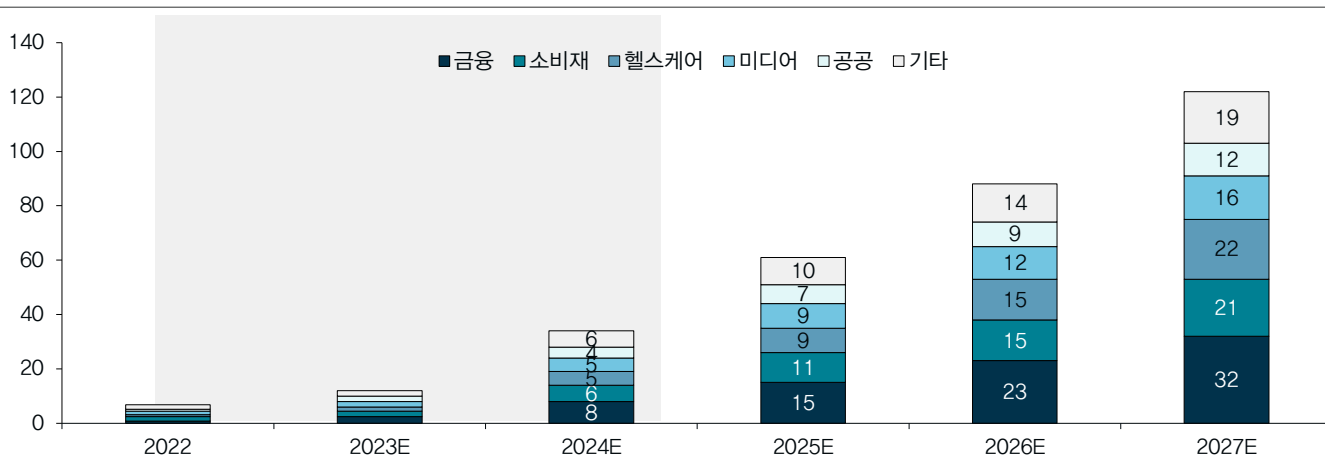
따라서 생성형 AI의 도입 이후 각 산업의 향후 진행방향을 파악하는 것이 중요 해졌다. AI가 적용되는 산업 분야에는 금융, 의료, IT, 운수 및 물류, 에너지, 소매 및 전자상거래, 제조업 등 여러 산업이 존재한다. [도표34] 서비스 분야로의 확장이 기대되는 생성형 AI의 특징상, 일반 소비자들이 주위에서 쉽게 접할 수 있는 산업 전반적에서 생성형AI의 도입으로 인해 변화가 일어나고 있는 것이다.

그러나 이 모든 응용 분야에서의 AI 현황을 살펴보는 것은 현실적으로 불가능하기에, 본 보고서에서는 1) 금융 산업 2) 헬스케어 산업 3) 미디어 엔터/게임 산업내에서 AI가 어떻게 응용되고 있는지에 대해 다루고 있다. 본 리서치팀이 위와 같은 세가지 산업을 선정한 데에는 다음과 같은 이유가 있다. 바로 1) 현재 AI응용 산업 분야 중 가장 큰 시장을 구성하고 있는 산업, 2) AI 응용 산업 중 가장 빠른 성장세를 보이고 있는 산업, 3) 생성형 AI중에서도 차세대 AI 주역으로 떠오를 이미지 생성형 모델의 특징이 가장 돋보이는 산업이다.

생성형 AI의 도입으로 어느때보다 AI 산업에 대한 기대감이 고조되어 있는 지금, AI 뿐만 아니라 AI 응용 산업을 살피는 것은 투자의 방향성을 설정하는 데에 중요한 지표가 되어줄 것으로 예상된다. 따라서 본 리서치팀은 본 보고서 앞부분에서 제시된 직접적인 생성형 AI 밸류체인과 더불어, Part3에서 다루어질 각 산업 내의 AI 발전 현황 및 미래 청사진을 통해 AI 및 생성형 AI에 대해 주목할 것을 제안한다.

도표 34. 생성형 AI 시장

(단위: 10억 달러)



# Part 4. AI 응용 산업 맛보기

## 1. 금융 산업

가장 큰 시장을 가진  
금융 산업

금융 산업이 AI 응용 산업 내에서 차지하는 비중은 약 20%로, AI 산업 내 **가장 큰 비중**을 차지하고 있다. 금융 AI 산업은 CAGR +38%로 성장해 26년 기준 562억 달러 규모의 시장을 이룰 것으로 예상된다. 정보기술 컨설팅 업체인 Capgemini에 따르면, AI를 활용해 지능형 자동화를 이루어냈을 때, 금융 산업 매출 증대 효과가 최대 5,100억 달러에 달할 것으로 추정된다. 그만큼 금융 산업 내에서의 AI가 차지하는 중요도는 높다고 볼 수 있다. 금융 분야의 AI 활용은 우리에게 가장 익숙한 앱 내 챗봇에서부터 시작되며, 현재 AI 기술은 현재 주로 금융기관내부업무를 수행하는데 활용되고 있다.

[도표 35]

### 금융 산업 내 AI 기술의 오늘과 내일

금융 산업내 AI 활용은 1) 고객 서비스 수준 향상, 2) 업무 효율성 향상 및 비용 절감, 3) 금융기관 **고도화**를 이끌어낼 것으로 전망한다.

1) 고객 서비스 수준  
향상

금융 산업은 생성형 AI기술을 활용하여 고객 서비스 수준을 챗봇 및 앱 중심으로 향상하고 있다. 키워드 중심이었던 과거의 챗봇은 생성형 AI의 등장에 따라 자연어 처리 기술을 통해 업그레이드 되었으며, 결과적으로 현재는 소비자의 요청에 즉각적 지원과 문의에 대한 답변을 시의적절하게 제시할 수 있게 되었다. 챗봇은 은행의 짧은 영업시간으로 인한 소비자들의 불편함을 해결할 수 있기에, 소비자의 편의성을 높이는데 도움이 된다. 또한 챗봇 사용성을 크게 높여, 기존에 젊은 세대에 국한되었던 챗봇 이용자에 다양한 연령대를 포괄하고 있다. 그 예시로 Bank of America의 모바일 앱 Erica는 챗봇은 아니나, 음성인식 기술을 사용해 중장년층 소비자의 사용성을 크게 높였다.

2) 업무효율성 ↑  
비용 ↓

생성형 AI는 **반복되는 업무를 자동화 시켜 인력의 효율성을 증가**시킨다. 반복되는 업무는 AI가 맡고, 일반 직원은 민감하거나 중요한 정보를 다루는 일에 집중할 수 있게 되는 것이다. 실제로 Bank of America는 Erica 출시 이후 100만 시간 이상의 직원 시간을 절약하였다. 또한, 주로 오프라인 창구에서 진행되는 고객 대응 업무가 줄어 오프라인 점포를 운영하는데 드는 비용을 아낄 수 있다. 판관비 중 인건비 비중이 높은 금융 산업의 특성상, 생성형 AI의 도입시 영업이익을 상승 시키는데에 큰 도움이 될 것이라 판단한다. [도표 36]

3) 금융기관의 고도화

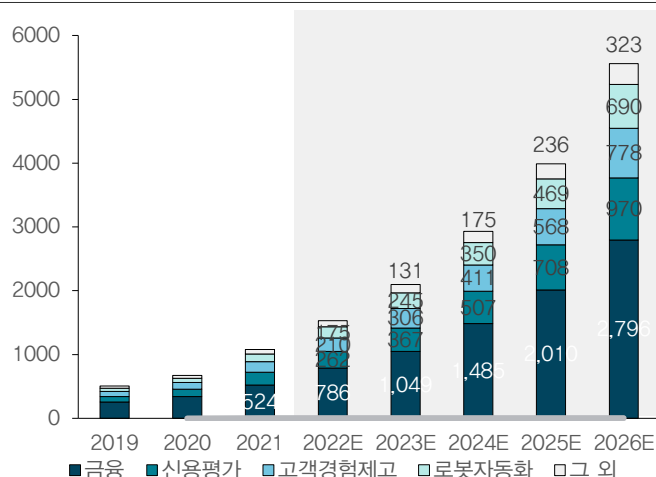
생성형 AI는 고객의 평소 결제 활동 패턴이나, 거래 특징을 파악해 이상 거래 발생 시 보이스피싱과 같은 사기 행위를 미연에 방지할 수 있다. 또한, 복잡한 신용 분석 절차를 간소화 해 대출 관련 의사결정 기간을 줄일 수 있으며, 개인 투자 성향을 파악해 맞춤형 투자 정보를 제공하는 등 금융기관 고도화를 이끌어낼 수 있을 것으로 판단한다.

개인별 투자 설계에  
대한 의식

그러나 AI를 사용해 개인별 투자 포트폴리오를 설계하는 '다이렉트 인덱싱'과 같은 경우에는 투자가 정량적 데이터만을 활용하는게 아닌 시장의 갑작스러운 이벤트나 대중들의 주관적인 반응과 밀접한 관련이 있다는 점에서 그 실효성에 대한 부정적인 의견이 존재하기도 한다.

도표 35. 금융분야 AI 분야별 시장규모

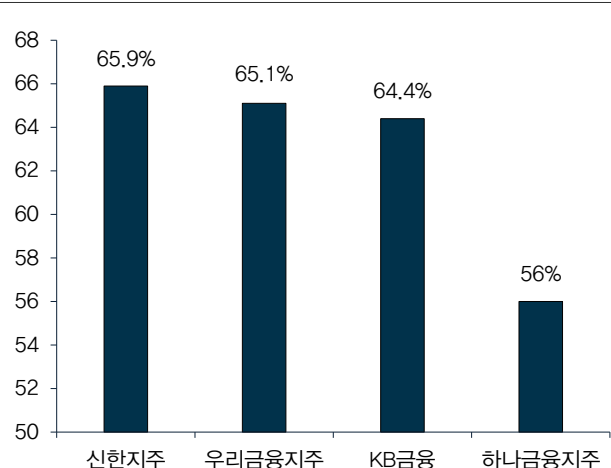
(단위: 백만달러, %)



자료: MarketandMarkets(2021), 성균관대학교 금융투자학회 S.T.A.R

도표 36. 국내 금융사 판관비 중 인건비 비중

(단위: %)



자료: 각 사,성균관대학교 금융투자학회 S.T.A.R

# Part 4. AI 응용 산업 맛보기

## 2. 헬스케어 산업

### 헬스케어 산업에서의 AI 응용이 기대가 되는 이유

가장 빠르게 성장할 헬스케어 AI 산업

헬스케어 산업은 AI응용 산업 중 타 산업 대비 가장 빠른 성장세를 보여줄 것으로 보인다. 23년 기준 의료AI산업 시장 규모는 206억 달러 규모로, 30년까지 CAGR +36.7%로 성장해 약 1880억 달러 규모에 도달할 것으로 예상된다.

헬스케어 - AI : 떨어질 수 없는 사이

헬스케어 AI산업의 급속한 성장은 1) 현재 AI 관련 산업 투자의 15~20% 가량이 헬스케어 AI에서 이루어진다는 점과, 2) 헬스케어 산업 내에서 AI에 대한 관심도가 높아지고 있다는 점, 3) 주요 글로벌 기업들이 의료산업에 진출하고 있다는 점에서 기인한다.

모건스탠리에 따르면, 헬스케어 산업 내 인공지능(AI) 및 머신러닝(ML) 예산 비중은 22년 기준 5.5%에서 24년 10.5%로 증가할 것으로 보인다. 또한 실제로 23년 3분기 기준 전분기 대비 AI 동의어 집단 언급 수 증가율이 75%로 AI 응용 산업내에서 언급 빈도수가 늘고 있다는 점을 통해 여러 산업 중에서도 헬스케어 산업에서 AI에 대한 관심도가 증가하고 있다는 점을 파악할 수 있다. [도표 37]

주요 기업들의 헬스케어 AI 산업 진출

주요 글로벌 기업들의 헬스케어 AI산업 진출은 투자 형태나 플랫폼 개발 형태로 이루어지고 있다. 주목할만한 사례로는 글로벌 반도체 기업인 엔비디아의 AI바이오 업체 '리커전'에 대한 5000만 달러 규모의 투자가 있다. 엔비디아는 리커전에 대한 투자를 통해 AI신약 개발 플랫폼을 고도화할 계획이라 밝혔다. 빅테크 기업의 헬스케어 AI로의 진출도 눈여겨볼만 하다. AWS는 의료용 클라우드를 일반적 클라우드와는 별도로 제공하고 있으며, 대표적 고객사로는 필립스와 루닛 등이 있다. 원격의료 AI 스타트업인 루닛은 제공된 AWS 클라우드를 활용하여 AI 원격의료 영상 분석 소프트웨어를 개발한 바 있다.

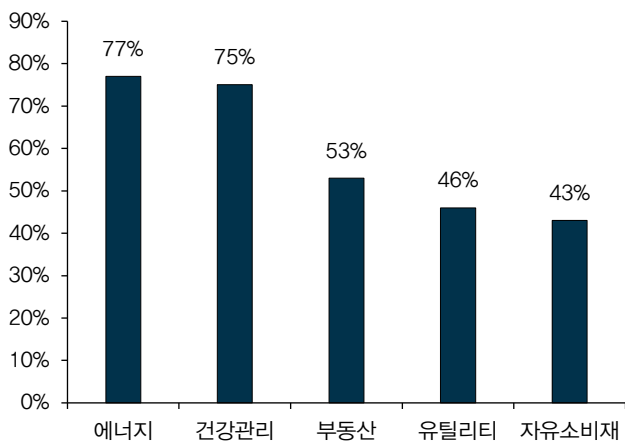
앞으로의 진행 방향

헬스케어 AI 산업은 크게 활용 분야별로 1) 의료용 데이터 2) 진단 의료기기 3) 신약 개발로 분류된다. 특히 보수적인 산업의 특징 상, 헬스케어 산업 내에서 AI기술은 신형 기업의 기술이 독자적으로 사용되기 보다는 빅파마나 기존 대형 헬스케어 기업과의 결합 등으로 확대될 것으로 예상된다. 보수적인 분위기에 불구하고 위에서 언급했던 바와 같이(AI언급 수 증가율) 의료계 내에서 AI에 대한 관심이 증가하고 새로운 기술이 적극적으로 받아들여지고 있다는 것은, 의료계 내부에서도 AI기술이 가격 절감 및 효율성 제고에 도움이 된다는 것을 인정했다는 의미로 이해할 수 있다.

생성형 AI의 헬스케어 산업 내 놀라운 가능성 확인

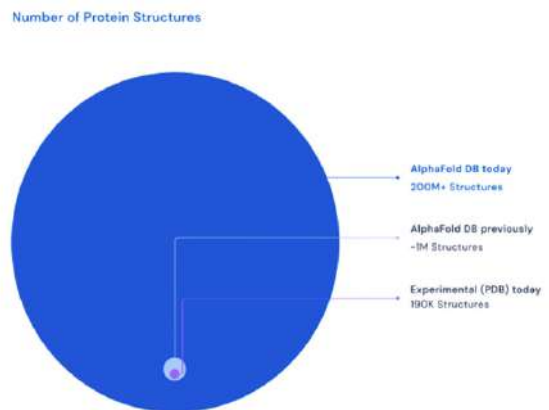
AI기술이 가져온 가장 가시적인 충격은 20년도에 구글 딥마인드사에서 개발한 알파폴드-2가 성공 시킨 단백질 구조 분석이었다. [도표 38] 알파폴드-2는 딥러닝을 통해 헬스케어 분야의 데이터를 학습한 인공지능 프로그램이다. 알파폴드-2는 10여년간 풀리지 않았던 생물학계의 난제였던 단백질 구조 분석을 30분만에 해결했다. 이 성과를 통해 생물학계에서 유전자 변이가 질병에 미치는 영향을 파악할 수 있게 되었다.

도표 37. 전분기 대비 AI 언급 증가율 (단위: %)



자료: Bloomberg, 성균관대학교 금융투자학회 S.T.A.R

도표 38. 알파폴드-2로 훨씬 많아진 단백질 구조 데이터 베이스



자료: 딥마인드, 성균관대학교 금융투자학회 S.T.A.R

# Part 4. AI 응용 산업 맛보기

## 2-1) 데이터 솔루션

헬스케어 데이터 소프트웨어 솔루션의 도입

헬스케어 데이터 분야에서 AI 응용의 영역은 무궁무진하다. 헬스케어 AI 산업 내에서 헬스케어 데이터 소프트웨어 솔루션 산업이 차지하는 비중은 21년 기준 약 39.5%에 달했으며, 가장 고속 성장을 하고 있는 분야로 꼽혔다. 단순히 데이터를 수집하고 저장해 정리할 뿐만 아니라, 의미 있는 정보를 끌어내 정보 수집자에게 제공하는 역할을 해낼 수 있기 때문이다. 의료분야를 예로 들자면, 진찰 데이터를 통해 의료진에게는 최적의 환자 치료 방안을, 환자에게는 개인에게 최적화된 처방 및 향후 질병 관리 방안을 제공하는 것이다. 데이터 소프트웨어 솔루션의 경우, 데이터 그 자체를 다루기 때문에 헬스케어 산업 내 타 분야에 비해 받아 들여지는 데에 거부감이 없다.

(\* 헬스케어 산업 내 데이터를 다루는 산업 분야를 통칭하는 명칭이 정해져 있는 것은 아니나, 본 리서치팀은 해당 산업 분야를 통틀어 헬스케어 데이터 소프트웨어 솔루션 산업이라고 지칭하였다.)

### 지금 주목해야 할 데이터 : 의료 데이터

의료 분야 데이터를 봐야 하는 이유

헬스케어 내 의료, 바이오, 생명과학 등 여러 분야에서 데이터 소프트웨어 솔루션 산업이 활발히 사용되고 있으나 그 중에서도 가장 주목할 만한 분야는 **의료 분야**이다. 이는 1) 주요 국가의 헬스케어 비용 중 의료서비스의 비중이 가장 크다는 점, 2) 의료 데이터의 활용이 아직 활발히 이루어지지 않고 있다는 점, 3) 개인 맞춤형 헬스케어 분야로의 확장이 기대된다는 점에서 기인한다.

헬스케어 내 의료서비스 비중 ↑

OECD 국가 중 헬스케어 관련 지출이 가장 큰 국가인 미국을 살펴보면, 전체 헬스케어 비용 중 약 64%를 의료서비스가 차지한다는 것을 알 수 있다. 의료서비스를 제공하는 대표적인 기관은 병원이다. 대략 4,660여개의 병원들이 의료 데이터 소프트웨어 솔루션 등의 의료서비스에 780만 달러를 지출하고 있으며, 이는 병원의 영업비용인 약 2억 900만 달러의 3%에 달하는 수준이다. 미국 뿐만 아니라 일본, 한국, 중국 등의 주요 국가들 에서도 전체 헬스케어 비용 중 의료서비스 비용이 차지하는 비중이 각 64%, 56%, 33%으로 가장 높은 것을 알 수 있다. [도표 39] 이를 통해 헬스케어 산업 중에서도, 의료서비스에서의 데이터 소프트웨어 솔루션 산업이 가장 Up-side가 클 것으로 추정된다.

아직 많이 남은 데이터

또한, 방대한 의료데이터가 생성되고 있음에도 불구하고 병원에서 생성되는 데이터의 97%는 활용되지 않은 상태로 저장만 되고 있다. 이러한 사실 또한 **AI기술을 접목한 데이터 소프트웨어 솔루션**의 활용 방안에 대한 기대감을 부여한다.

빅테크 기업의 진출

의료 산업 내의 데이터 소프트웨어 솔루션 산업에 대한 기대는 막연한 것이 아니다. 가능성을 엿본 구글, MS와 같은 빅테크 산업들은 이미 각 사의 생성형 AI를 활용해 의료 AI 분야에 진출할 계획을 밝혔다. 구글은 벡텍스 AI서치 플랫폼에 구글클라우드의 헬스케어 API와 검색엔진을 통합하여 의료 검색 기능을 제공할 것이며, MS는 애저 AI서비스에 의료 기관용 서비스를 더한 챗봇 서비스를 제공할 것으로 예상된다. [도표 40]

도표 39. 주요 국가의 헬스케어 비용

(단위: %)

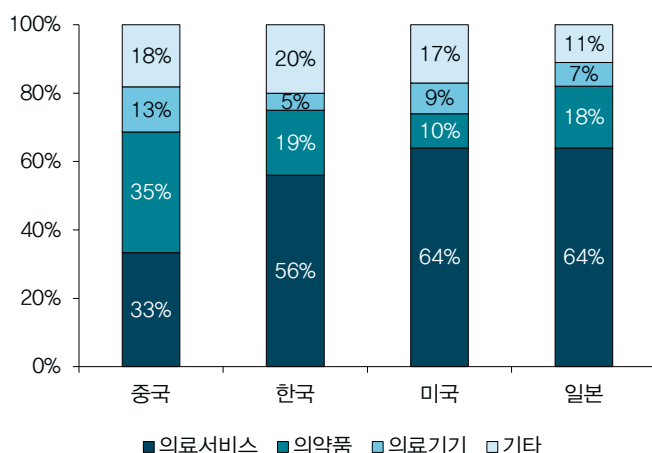
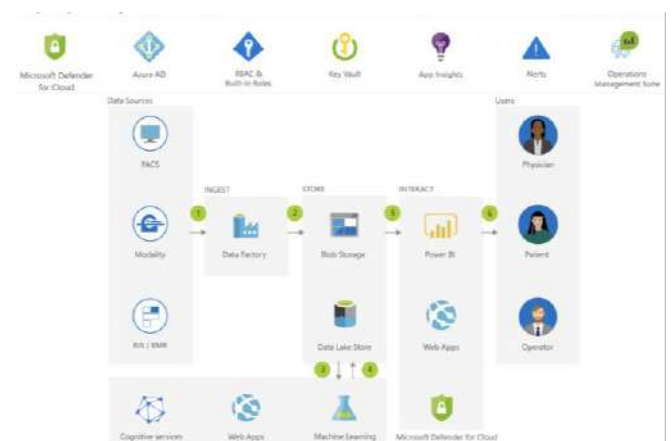


도표 40. MS 의료 데이터 솔루션 구조



자료: 통계청, NIH, CMS, OECD, 성균관대학교 금융투자학회 S.T.A.R

자료: MS, 성균관대학교 금융투자학회 S.T.A.R



# Part3. AI 응용 산업 맛보기

## AI가 의료 데이터 소프트웨어 솔루션에 활용되는 방식은?

의료계 전통적  
데이터 관리 시스템 +  
AI

의료 데이터 소프트웨어 솔루션에는 EHR(전자건강기록), PACS(의료영상저장전송시스템) 등이 활용된다. EHR은 환자의 진료 정보를 전산화 한 것이며, PACS는 영상 진단 장치로 촬영한 환자의 영상 정보를 획득해 관독과 진료기록을 함께 전송 및 관리하는 전통적 시스템이다. 업무 전산화에 불과한 두 시스템에 더해 AI의 도입은 더 편리하고 효율적인 환자 관리 및 치료를 가능하게 할 것으로 전망한다.[도표 41]

구글: Vertex AI

예를 들어, 앞에서 언급했던 구글의 '버텍스 AI'의 경우 의료 검색 기능이 주요 기능이다. 구글은 구글클라우드를 통해 버텍스 AI서치 플랫폼에 API와 헬스케어 데이터 엔진을 통합해 제공하며, 해당 검색 기능의 사용자는 의료용 LLM인 매드-팜2.[도표 42]와 결합해 더 쉽게 전문적인 의료 정보를 찾을 수 있다. 의료진은 환자의 병력에 대해 EHR을 일일이 확인할 필요 없이, '버텍스 AI'에 질문을 던져 환자의투약 기록이나, 타 병원에서의 진료 기록 등 환자에 대한 필요한 정보를 손쉽게 얻을 수 있는 것이다.

: One-step  
: 진단-치료-청구

이 외에도 AI, 특히 생성형 AI와 결합된 EHR은 의료 서비스 내에서 의료진에게 의료 정보를 요약해 보여주거나, 복잡한 임상용어에 대한 자동완성 기능을 제공하고, 음성 인식 기술을 활용해 의료진과 환자 간의 대화를 듣는 것만으로도 최적의 치료 방안을 제시하는 등 다방면에서 활용될 수 있다. 특히 웨어러블 기계나 청구 시스템과의 결합을 통한 부가가치 창출도 기대가 된다. 웨어러블 기계에서 추출되는 데이터를 통해 환자의 평소 건강 관리가 용이해질 뿐만 아니라, 병원 방문 시 웨어러블 기기를 통해 평소 추적 했던 환자의 건강 데이터를 기반으로 출력되는 맞춤형 치료 방안부터 보험금 청구 과정까지 One-stop 형식으로 이루어질 것이 가능해질 것으로 예상되기 때문이다.

## 의료기관의 효율성은 UP, 비용은 DOWN

매출 ↑ 비용 ↓

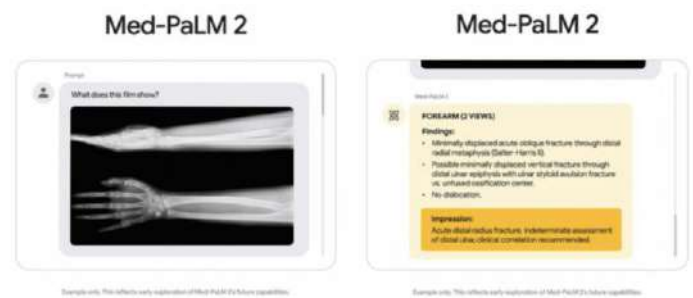
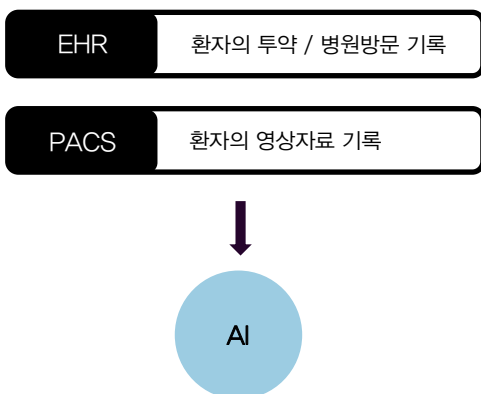
AI를 활용한 데이터 소프트웨어 솔루션은 편리성이나 치료의 효율성 뿐만 아니라, 의료 서비스의 공급자 측면에서 매출 증대 및 비용 절감 효과로 이어질 수 있다. 이는 환자 정보 검색이나 임상 용어 탐색 등의 치료와 직접적인 관련도가 떨어지는 불필요한 행위에 대한 의료진의 노력과 비용을 줄이는 대신, 고부가가치 창출이 가능한 행위에 의료진이 집중한 결과에서 기인한 것으로 추정된다.

가시화된  
AI 데이터 소프트웨어  
솔루션의 효용성

AI 데이터 소프트웨어 솔루션을 활용하는 경우가 늘며 의료기관의 재무 효율성 증가를 실제로 보여주는 사례가 나타나고 있다. 미국의 헬스케어 기업인 카이저 퍼머넌트는 데이터 소프트웨어 솔루션을 도입한 뒤, 검사 비용과 사무실 방문 비용 등 총 10억 달러를 절감한 바 있다. 미국 콜로라도에 위치한 UHealth 병원 또한 수술실 이용 계획을 최적화 해 수술 관련 매출을 4% 증가시켰다. 솔루션을 활용해 단순히 비용 뿐만 아니라, 환자 한 명 당 평균적으로 소요되는 행정 업무 처리 시간인 2시간을 줄일 수 있을 것이다. 이를 바탕으로 부족한 의료인력을 충원하는 효과 또한 나타날 것으로 예상된다. 따라서, 앞으로 의료기관에서 데이터 소프트웨어 솔루션을 사용하는 경우가 늘며 이러한 사례를 찾기는 더 쉬워질 것으로 예상된다.

도표 41. EHR, PACS 와 AI의 결합

도표 42. 매드-팜2 활용 모습



# Part 4. AI 응용 산업 맛보기

## 2-2) 영상 및 진단 의료기기

의료기기의 활용을 돕는 AI 기술

앞서 살펴본 데이터 소프트웨어 솔루션과 달리, 의료기기에 활용되는 AI 기술은 독자적 기술이라기 보다는 이미 쓰이는 의료기기의 활용도를 높이는 데에 중점이 있다. 그리고 이러한 역할을 제일 잘 활용할 수 있는 의료기기 분야는 의료 영상 및 진단 분야이다. AI는 이미지를 잘 분석하는 특징이 있기 때문이다. 따라서 의료 영상 및 진단 AI가 CAGR +58.1%로 성장해 27년에 헬스케어 AI 산업 분야에서 가장 큰 규모를 차지하며, 약 122억 300만 달러 규모의 시장을 이룰 것으로 전망된다.

### 진단형 의료기기 AI의 오늘과 내일

영상 판독이 주기능

영상 및 진단 AI에서 가장 많은 비중을 차지하고 있는 품목은 방사선 관련 AI 기술이다. FDA 인증 AI 및 머신러닝 기반 의료기기 약 521개 중 392개(약 75%)는 방사선학을 기반으로 하는 의료기기이다. [도표 43] 이를 통해 의료기기용 AI는 거의 대부분 직접 의료기기로서 역할을 하기 보다는 영상을 판독하는 기술에 머물러 있음을 파악할 수 있다. AI를 활용하면, 빅데이터를 활용해 영상에 대한 빠르고 정확한 판단이 가능하다는 장점이 있다.

보조도구에서 단독 의사결정까지

현재 영상 및 진단 AI 기술은 현장에서 의료용 데이터를 분석해 의사의 진단을 보조하는 역할(CDSS)로서 대부분 활용되고 있다. 인공지능에게 단독으로 의사결정을 맡기기에는 기술 도입 초기에 머물러 있기 때문이다. 그러나 AI의 단독 의사결정은 빠른 시일 내에 가능해질 것으로 전망된다. 이는 부족한 의료 영상 판독 인력 및 관련 정책의 변동에서 기인한다. 국내 영상전문의는 전체 전문의 숫자의 5% 이내로, 그 수가 매우 부족하다. [도표 44] 이는 다른 주요 국가들도 크게 다르지 않다. WHO는 결핵 검진 및 선별검사에 있어 의사의 엑스레이 판독 대안으로 AI를 단독 사용할 것을 권고했으며, FDA 또한 망막 촬영 카메라에 당뇨병성 망막증을 진단하는 AI기술인 Idx-DR 을 승인했다. 이는 FDA 최초의 AI 단독 진단이 가능한 인허가이다.

### 앞으로 나아갈 방향성은?

병리학 분야로의 확장 기대

앞으로 의료기기 시장에서의 AI는 영상 및 진단 뿐만 아니라 병리학 영역으로의 확장이 기대된다. 병리학 분야로 발전할 경우, 의료기기 AI 시장의 매출 증대와 바이오 산업과의 협업을 기대할 수 있을 것이다.

예시: 루닛

의료 영상 AI 관련 기업인 '루닛'은 실제로 병리학 분야로의 진출을 가시화 했다. AI 영상분석 솔루션을 제공하는 루닛 인사이트 사업부와 AI 바이오 마커 플랫폼을 제공하는 루닛 스코프 사업을 동시에 운영하며, 영상 및 진단 AI 기업이 병리학 분야로 나아가는 선구자로서의 역할을 해내고 있다.

도표 43. FDA 인증 AI/ML 의료기기 현황

(단위: %)

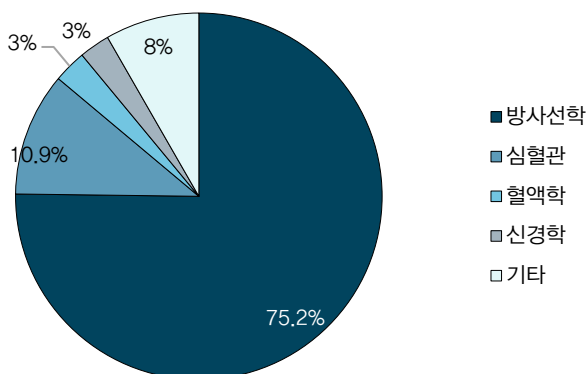
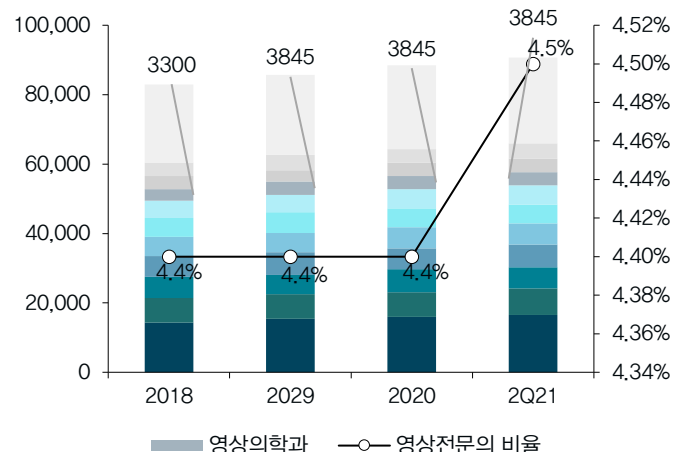


도표 44. 국내 의료인력 및 영상전문의 비율

(단위: 명%)



자료: FDA, 성균관대학교 금융투자학회 S.T.A.R

자료: 산업자료, 성균관대학교 금융투자학회 S.T.A.R

# Part 4. AI 응용 산업 맛보기

## 2-3) 생명과학 / 바이오

신약개발에 소요되는 시간과 비용 감축

AI 신약 개발 시장 규모는 25년까지 CAGR +47%로 성장해 약 59억 달러에 이를 것으로 전망된다. 알파폴드-2의 단백질 구조 분석 성공 등과 같이 AI의 발전은 생명과학 분야에 지대한 영향을 미치고 있으며, 단백질 및 유전체를 대상으로 한 분석 기술뿐만 아니라 신약 개발을 하는데까지의 변화를 일으키고 있다. 특히 신약개발 분야의 경우, AI기술을 활용해 기존에 치명적 단점이었던 평균 10년 이상의 장기간에 걸쳐 진행되는 연구/허가 과정 및 적게는 수백억에서 많게는 몇 조가 소요되는 고비용 문제를 해결할 수 있을 것으로 기대된다. 실제로 알파폴드-2를 신약 개발 과정에 활용할 경우, 전 임상 비용을 기존 비용 대비 20% 수준으로 낮추고 소요되는 기간 또한 절반 가량으로 줄일 수 있을 것이라는 연구결과가 존재한다.

### 신약개발 전반의 효율성을 UP 시켜줄 AI

신약개발 전반적 단계에 관여하는 AI

AI는 신약 개발의 전반적 단계에 대부분 관여한다. [도표 45] AI는 신약을 개발하기 이전 타겟 발굴부터, 신약 연구 단계에서 유효물질 및 선도물질을 발굴하고 연구하며, 전임상 및 임상 단계에서 실험 결과를 예측하고, 시판 이후에는 신약에 대한 추적을 가능하게 한다. 특히 AI 신약 개발 플랫폼을 가진 기업들의 경우, 계속해서 파이프라인을 뽑아낼 수 있다는 점에서 큰 장점을 갖는다. AI 기술을 활용하면 신약 개발 전 단계에서 소요되는 시간과 비용을 평균 1/3까지 단축이 가능하다. 실제로 화이자사는 AI를 활용해 mRNA 코로나 백신 개발을 10.8개월로 단축한 바 있다. (평균 백신 개발 소요기간은 10.7년이다) AI 신약 개발 기업인 'Exscientia' 또한 20년 최초로 AI 플랫폼을 활용하여 찾은 후보물질의 임상실험 단계에 돌입했는데, 임상실험까지의 개발기간단계를 기존 평균 5년에서 채 1년이 되지 않는 기간으로 줄였으며 후보물질 수 또한 20% 정도 수준으로 줄여 비용 또한 절감했다는 평을 받고 있다.

AI 신약개발 기업의 운영형태

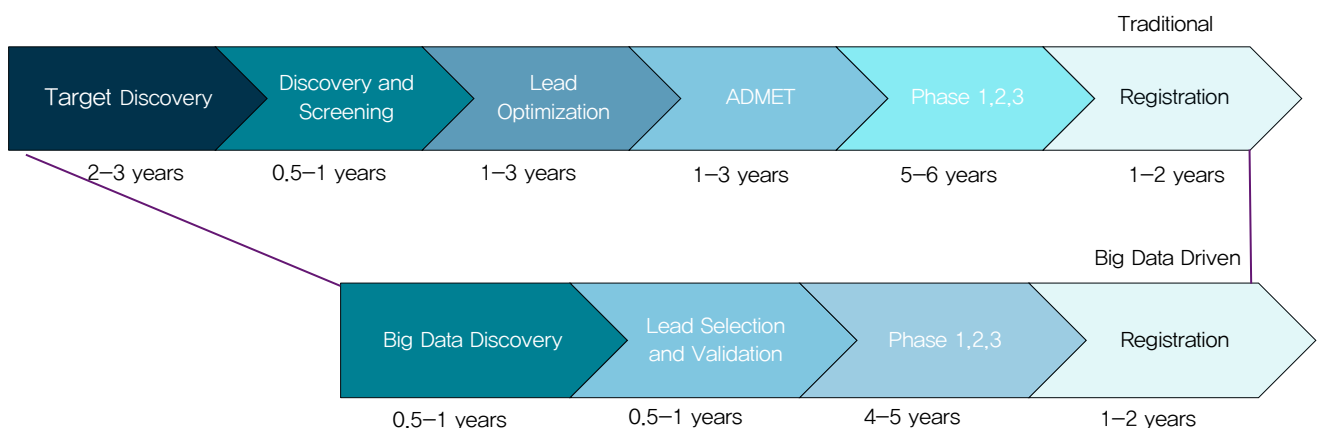
현재 AI 신약 개발 기업은 기존 제약 산업과 거의 유사하게 외주형 및 직접 개발형으로 분류된다. 1) AI플랫폼을 가진 외주형 기업은 제약사에게 연구를 위탁받고, 2) 회사 자체의 머신러닝 기술을 보유한 직접개발형 직업은 신약 개발 전반 과정에 직접 나선다. AI 신약 개발 산업의 성과가 가시화됨에 따라, 글로벌 빅파마들 또한 적극적으로 AI 신약 개발 기업과의 파트너십을 맺고 있으며, 이는 서로에게 안정적 매출 발생처와 연구 효율성을 높일 수 있다는 점에서 win-win로 보여진다. 대표적인 빅파마-AI 신약 개발사 간의 계약으로는 Sanofi와 Exscientia가 22년 체결한 53억 달러 규모의 항암제 관련 계약이 있다.

### 얼마 남지 않은 확인의 시간

1-2년 뒤 보여질 결과

다만 항상 그러하듯, 신약 개발에서의 핵심은 임상 통과인만큼, AI 기술을 활용해 임상 시험 성공률이 얼마나 될지에 대해서는 지켜봐야 할 것으로 예상된다. AI신약 개발 기업들이 이제 막 임상 2상 단계에 진입하기 시작하였기에, 관련 사례는 24-25년이나 되어야 확인이 가능할 것으로 전망한다.

도표 45. 신약개발 단계에서 AI 도입시 기대되는 소요시간 감소 효과



# Part 4. AI 응용 산업 맛보기

## 3-1) 미디어 및 엔터테인먼트

인간과 AI의 경계가 모호해지는 미디어 산업

생성형 AI의 등장으로 인간 고유의 영역으로 여겨졌던 '창작'의 영역에 대한 구분이 모호해졌다. 특히 이미지 생성형 AI가 AI의 차세대 주역으로 꼽히며, 이미지나 비디오 등의 콘텐츠를 창작해 활용하는 미디어 및 엔터테인먼트 분야에서 AI의 도입이 활발해질 것으로 예상된다. 미디어 및 엔터테인먼트 분야 AI 산업은 21년 기준 179억 달러 규모의 시장에서 CAGR +31.89%로 성장해 28년에는 약 44억 달러 규모의 시장에 달할 것으로 예상된다.

### 미디어 업계 전반으로 확장 중인 생성형 AI

영상 산업에서의 기술 활용

AI 기술은 **영화나 드라마 등 영상 산업**에서 적극적으로 활용되고 있다. 촬영 이후 편집 과정에서 AI 기술을 활용해 편집 시간과 비용을 줄이고, AI 보이스 클로닝이나 페이스 디에이징 효과 등을 도입하고 있는 것이다. 뿐만 아니라 각본 제작 과정에서도 AI를 활용한다. 대표적 예시로는 디즈니가 있다. 디즈니는 엔비디아가 개발한 이미지 생성 알고리즘을 활용하여 배우의 얼굴을 디에이징 시키거나 에이징 시킬 수 있는 FRAN(Face Re-aging Network) 기술을 적극 활용하고 있다.

영상 플랫폼 산업에서의 기술 활용

최근 영상 중심의 플랫폼이 인기를 끌며, 영상 편집 도구에 대한 수요가 증가하는 경향을 보였다. 이에 콘텐츠 편집 플랫폼의 전통적 강자인 어도비사는 사용자의 명령에 따라 이미지와 비디오를 생성하거나 편집하는데에 활용되는 생성형 AI 기반 모델인 파이어플라이를 베타버전으로 출시한 바 있다. 숏폼 플랫폼의 강자인 틱톡 또한 텍스트를 입력하면 관련 이미지를 생성해주는 생성형 AI '그린스크린' 기능을 제공하기 시작했다. 비디오 플랫폼에서 AI를 통해 개인 맞춤형 콘텐츠를 추천 받던 시대에서, 생성형 AI를 활용해 스스로 비디오를 만들기가 쉬워진 시대에 들어선 것이다.

엔터 산업에서의 기술 활용

엔터 분야에서도 생성형 AI의 등장으로 **버추얼 휴먼**에 대한 관심이 급증했다. 기존에는 AI 기술을 활용해 인간 외양만을 닮은 버추얼 휴먼을 출시했다면, 생성형 AI 기술의 도입 이후 등장하는 버추얼 휴먼은 데이터 학습을 통해 사람들과의 자연스러운 소통이 가능해진 것이다. 따라서 팬들과의 소통을 중시하는 아이돌이나 인플루언서, 심지어는 쇼호스트 역할을 수행하는 버추얼 휴먼의 등장이 증가할 것으로 예상된다. 버추얼 아이돌로 가장 화제성을 불러일으켰던 예시로는 SM 엔터테인먼트의 '나비스'가 있으며, 버추얼 인플루언서로는 신한 라이프 광고에 등장후 많은 사람들의 관심을 받았던 '로지'가 있다. **[도표 46]** 버추얼 아이돌이나 인플루언서들은 실제 아이돌이나 인플루언서가 갖는 생활 영역에서의 리스크나 여러 법적 제약에서 자유롭다는 점에서 광고계에서 선호되고 있다.

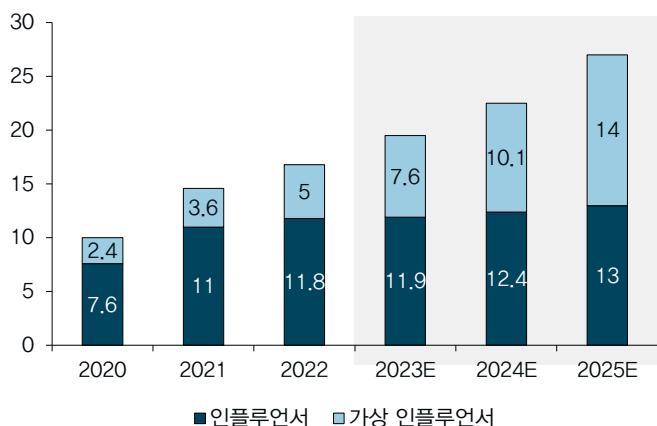
### [도표47] 주의해야 할 점

기술 활용 시 주의점

다만, 실제 인물의 목소리나 얼굴을 AI 기술을 통해 합성하거나 편집하는 것은 가짜 뉴스나 저작권 등의 문제에 직면할 가능성이 높다. 따라서 여전히 미디어 업계에서의 AI 규제에 대해서는 논의가 필요한 상황이다.

도표46. 글로벌 버추얼 인플루언서 시장 전망

(단위: 조 원)



자료: 하이프오디터, 성균관대학교 금융투자학회 S.T.A.R

도표 47. 버추얼 인플루언서 로지



자료: 성균관대학교 금융투자학회 S.T.A.R

# Part 4. AI 응용 산업 맛보기

## 3-2) 게임

생성형 AI의 도입으로 판도가 바뀔 분야: 게임

게임 산업은 생성형 AI의 도입으로 가장 혁신적인 변화를 경험할 산업이 될 것이라고 예상된다. 이는 게임 산업의 기술 및 노동집약적인 성격에서 기인한다. 게임의 3대 요소는 기획, 디자인, 프로그래밍이다. 스토리 라인을 구성하고, 그래픽 및 사운드를 프로그래밍한 게임에 입히는 과정에는 막대한 비용이 소요된다. 특히 각 과정에서 전문적 기술을 요구하므로 게임 산업 내에서 인건비로 지출하는 비중은 매출대비 약 1/3로 타 산업 대비 높은 수치를 보인다. 따라서 AI 도입 시 가장 기대되는 부분은 제작비 감소이며, 또한 코로나 이후 잠시 침체되었던 게임 산업에 활기를 불어 넣어줄 것으로 기대된다. 또한, 게임 업계에서는 기존에 NPC 모델링이나 채팅 등에 AI 기술을 도입해왔었기 때문에, 생성형 AI 또한 게임산업에 자연스럽게 녹아들 것으로 예상된다.

### 비용 감소가 불러올 게임 산업의 혁명

비용절감  
1) 프로그래밍  
2) 디자인

게임 공급자인 게임 기업의 경우 생성형 AI 도입을 통해 1) 프로그래밍 분야와 2) 디자인 분야의 비용을 줄일 수 있을 것으로 판단한다. 게임 개발 비용 중 약 65%는 프로그래밍 및 디자인이 차지한다. [도표 48] 트리플 A급의 게임을 제작하는데에 적어도 국내에선 100억원, 글로벌 기업에선 1,000억원 이상이 드는 것을 고려하면 큰 비용을 아낄 수 있게 되는 것이다.

AI의 도입으로 훨씬 간편해지는 프로그래밍

게임 산업 내에서 프로그래밍은 노코드 개발 툴이 보편화 될 것으로 예상된다. Github의 Copilot이 등을 활용하면 단순한 코딩 같은 경우에는 AI가 수행하고, 프로그래머는 AI가 해결할 수 없는 고난이도의 프로그래밍에 집중화 될 것으로 예상된다. 프로그래밍에 있어서의 AI 기술 도입은 22년 이후 빅테크 기업 및 게임산업에 붙었던 프로그래머에 대한 구조조정 이후 본격화 된 것으로 판단된다.

AI의 도입으로 디자인 분야의 비용을 획기적으로 감소

게임 산업 내의 디자인 분야에서도 특히 생성형 AI의 도입으로 인해 필요한 인력이 줄어들고, 관련 비용을 크게 감소할 수 있을 것으로 예측한다. 게임 산업 내 DALL-E나 스테이블 디퓨전 등 이미 생성형 AI의 도입으로 원하는 디자인적 요소를 쉽게 구축할 수 있게 되었다. [도표 49] 이러한 AI 기술을 활용해 게임 이미지를 생성하면, 일러스트레이터에게 의뢰하는 것보다 약 수십만 분의 일로 비용을 아낄 수 있다. 또한 소요되는 기간이 1~2주에서 3~4일로 줄어든다. 따라서 디자인을 맡던 일러스트레이터들의 역할이 이미지를 처음부터 생성하기보다는 생성형 AI를 통해 출력된 이미지를 리터칭하는 역할로 바뀌고 있다. 이에 더해, 사운드 영역에서는 성우의 목소리를 녹음해 사용하기보다 AI기술을 적용시키는 경우가 많아질 것으로 보인다. 넥슨 인텔리전 랩스의 경우에는 NPC에게 성우의 목소리가 아닌 디지털 사운드를 입힌 게임 세계관 속 AI 페르소나를 창출해내고자 노력하고 있다.

규모와 상관없이 성장할 게임 기업

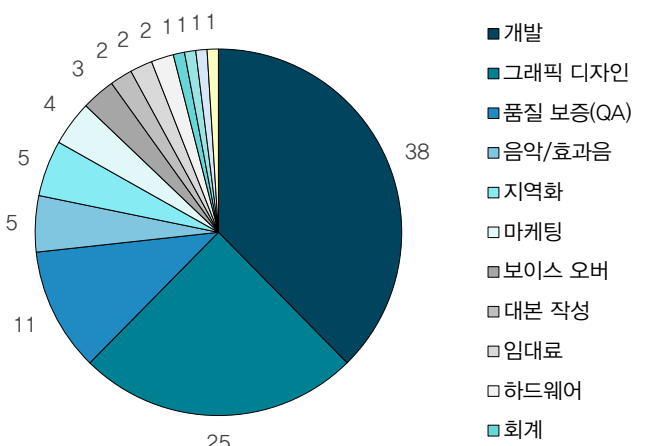
따라서 IP를 소유하고 있는 대형 기업 뿐만 아니라, 비용에 대한 부담이 줄어든 소형 기업의 활약 또한 기대되는 바이다. 생성형 AI 기술 도입은 게임 산업 전반에 편의성을 향상시키고 더 높은 퀄리티의 게임을 출시할 수 있는 기회를 제공할 것으로 판단한다.

도표 48. 이미지 생성형 AI를 활용해 만든 게임속 그래픽



자료: 성균관대학교 금융투자학회 S.T.A.R

도표 49. 게임 제작 시 소요되는 비용의 분야별 비중 (단위: %)



자료: Game Developer, 성균관대학교 금융투자학회 S.T.A.R

## Part 5. 생성형 AI, 개선해야 할 부분은?

생성형 AI는 아직 초기 단계의 기술로서, 향후 상용화 과정에서 다양한 이슈들이 발생할 가능성이 있다. 대표적으로 1)윤리적, 2)수익성, 3)법적 측면에서 문제가 발생할 가능성이 존재한다.

### 윤리적 측면

#### 1) 윤리적 측면

생성형 AI는 특정한 데이터셋의 비중이 높거나 폭력이나 인종차별 등의 내용을 학습하게 되면, 공격적이거나 편향적인 텍스트를 생성해낸다는 문제가 존재한다. 실제로 메타가 22년 11월에 공개한 과학용 대규모 언어모델 Galactica 데모 버전은 인종 차별적이며 부정확한 문헌이 생성되는 것을 막지 못하여, 3일만에 비공개 처리되었으며, Chat GPT에서도 인종, 성별 등에 관한 데이터를 편파적으로 처리하였다는 사례가 다수 보고되고 있다.

#### 개선책

이에 따라 Open AI에서는 사람이 직접 피드백을 주고 이를 통해 모델을 강화 학습시키는 RLHF(Reinforcement Learning with Human Feedback)을 통해 정확도를 개선시키고 있으며, Anthropic 역시 제작된 이미지가 사용자 의도에 부합하는지 여부를 AI가 판단해 피드백하고 이를 다시 학습시키는 RLAI(Reinforcement Learning with AI Feedback)를 통해 개선 중이다.

### 수익성 측면

#### 2) 수익성 측면

생성형 AI가 유망한 산업이라는 것과 별개로, 여전히 수익성 확보에 대한 의문이 제기되고 있다. 이는 기술 개발과 서비스 운영에 천문학적인 비용이 투입되는 한편, 빅테크 기업조차 뚜렷한 수익 모델을 가지고 있지 않았기 때문이다. 대표적으로 Chat GPT의 유지비용은 하루 9억원, 연간 3300억원에 달하고 있으며, 서비스 고도화에 따라 운영비는 더욱 늘어날 가능성이 높다. 지속가능한 생성형 AI 비즈니스를 위해 수익모델 확보는 최대 과제로 꼽혀왔다.

#### 개선책

하지만 마이크로소프트, 구글, 아마존 등의 빅테크 기업들을 필두로 자사 서비스에 생성형 AI 기능을 접목한 서비스를 내놓으며, 유료화를 통한 수익화 행보를 강화하고 있다. [도표 50] MS는 자사 구독형 오피스 소프트웨어 'MS 365'에 생성형 AI기반 업무 보조 도구 'Copilot'을 결합한 'MS 365 copilot'을 내놓았으며, 구글 역시 8월 말 cloud Next 행사에서 Duet AI 유료화를 예고하였다. 무엇보다도 이미지, 비디오 등 세분화한 영역에서 킬러 앱이 등장한다면 이러한 수익화 문제가 크게 개선될 것이라 판단된다.

도표 50. 빅테크 생성형 AI 수익화 전략

빅테크	수익화전략
Open AI	1) API 가격 인하, 성능 강화로 모델 2) Chat GPT 구독 서비스
MS	1) 자사 서비스 운영체제에 생성형 AI 접목 2) 생성형 AI 서비스 탑재에 클라우드 경쟁력 강화
구글	1) 생성형 AI 빌더 플랫폼 '버텍스 AI'로 B2B 공략 2) 워크스페이스 등 자사 서비스에 AI 탑재
AWS	베드록, 렉스, 세이지메이커 등 기업형 생성형 AI 관리 서비스 출시
네이버	개별 기업, 정부 기관에 특화하고 보완성 높은 AI 제공

## Part 5. 생성형 AI, 개선해야 할 부분은?

### 법적 측면

#### 3) 법적 측면

생성형 AI가 만든 결과물에 대한 저작권 논란 역시 존재한다. 생성형 AI는 이미 존재하는 데이터를 기반으로 학습하여 패턴을 파악한 뒤 새로운 결과물을 생성해낸다. 이에 따라 기존 작품과의 유사성을 지닐 위험이 존재하며, 원데이터에 대한 저작권 침해와 보상에 대한 문제가 발생하게 되었다. 23년 9월 Open AI와 유명작가들 간의 소송을 포함하여, AI가 만들어낸 결과물에 대한 저작권 인정 여부와 그 저작권을 소유할 주체가 누구인지에 대한 논쟁이 지속적으로 발생하고 있다 [도표 51]

#### 개선책

이에 따라 일부 국가에서는 저작권 관련 규제 도입에 박차를 가하고 있다. EU에서는 21년부터 AI 규제를 위한 법안의 초안 작업을 진행하였으며, 23년 6월 해당 법안의 협상안이 가결되었다. 최종 절차가 무리 없이 진행된다면 26년부터 생성형 AI 운영 기업은 AI학습에 사용한 원데이터의 저작권을 의무적으로 공시해야 한다. 국내 역시 5월, 과학기술정보통신부에서 ‘인공지능 산업 육성 및 신뢰 기반 조성에 관한 법률’ 제정을 지원하겠다고 밝혔으며, 7월 기획재정부에서는 저작권법 개정을 통해 AI 데이터 활용에 대한 저작권 침해 면책 요건과 근거를 마련할 계획이라고 밝혔다.

도표 51. 생성 AI 관련 법적 분쟁 사례

원고	피고	내용
Getty Image	Stability AI	스테이블 디퓨전 학습에 자사 이미지가 무단 사용되었다고 주장
Sarah Anderson외 아티스트 다수	Stability AI, MidJourney, DevianArt	동의 없이 AI학습에 자신들의 이미지 사용 주장
집단 소송	Github(Microsoft)	Github Copilot 학습을 위해 다른 사람들이 만든 코드를 사용했다는 이유로 소송

자료: 언론 종합, 성균관대학교 금융투자학회 S.T.A.R

### 방향성은 변하지 않는다.

생성형 AI는 아직 개선해야 할 부분이 많지만, 다양한 방면에서 개선책을 마련 중이며, 이제 막 개화하기 시작한 기술이라는 점에서 기회요인이 더 많은 시장이라고 판단된다. 현재 진행중인 소송들의 결과와 법률적 가이드נס에 주목하며, 생성형 AI를 공정하고 윤리적으로 사용하는 방법에 대해 고민해볼 필요가 있다. 산업의 방향성 자체를 의심할 필요는 없다고 판단된다.

## N/R

목표주가	N/A
현재주가	N/A
상승여력	-

Consensus	2022	2023F
매출액(백만 원)	27	N/A
영업이익(백만 원)	(24,883)	N/A
영업이익률(%)	-	N/A

### S.T.A.R Research

#### 권예준

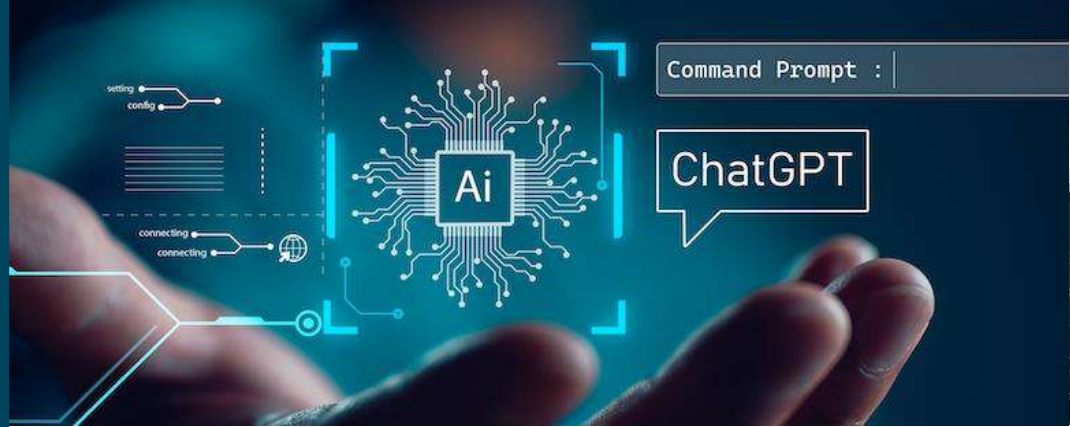
010.9022.8720  
yejoon0309@g.skku.edu  
46<sup>th</sup> member

#### 김원영

010.2756.1059  
wykim1059@gmail.com  
46<sup>th</sup> member

#### 양예지

010.3615.1527  
yellowyeji@g.skku.edu  
46<sup>th</sup> member



# 스탠다임 (비상장 기업)

## 숨겨진 AI 신약 개발사, 이제 그 진가를 드러낼 때

### AI 신약 개발 업계에서는 이미 유명인사

동사는 22년 기준 매출액 약 2737만원(YoY-50.23%), 영업이익 248억원(YoY-30%)를 기록했다. 수치적으로는 유의미한 실적을 내고 있다고 보기 어렵다. 그럼에도 불구하고 동사는 국내 제약사 4곳과 영국, 미국의 제약사와도 신약개발 협력을 성공적으로 마쳤다.

뿐만 아니라 동사는 영국의 제약바이오 전문 투자기업 DPI가 선정한 글로벌 AI 신약개발 기업 300개 중 상위 30개 그룹 안에 드는 등 국내외에서 그 가능성을 인정받고 있다. 이러한 가능성을 바탕으로 23년 시작될 자체 임상에 대한 기대감까지 높아지면서 가까운 시일 내에 상장을 목표로 하고 있다.

### 대기업을 등에 업고 상승 준비 완료

동사가 제공하는 AI 신약 솔루션은 크게 세가지 플랫폼으로 나뉜다. 1) 스탠다임 인사이트와 2) 스탠다임 베스트 3) 스탠다임 에스큐이다.

동사는 신약 후보 물질 탐색과 전임상 디자인까지 도출하는 플랫폼인 '스탠다임 베스트'를 통해 항암, 비알콜성지방간(NASH), 파킨슨병 등 다양한 분야에서 파이프라인을 개발했다. 현재도 약 42개의 신약후보물질을 개발중이며, 20년 기준 파이프라인과 관련하여 동사가 소유한 특허는 약 50개에 달할 것으로 추정된다.

동사는 제약 전반에 AI 기술을 도입하는데 선구주자로서 역할을 하고 있다. AI 관련 플랫폼인 신규 타겟 발굴 AI 프로그램인 '스탠다임 에스큐'를 구축해 스탠다임 베스트와 결합시켜, 타겟발굴부터 선도물질 확보까지 7개월밖에 소요되지 않는 워크플로우 AI를 생성했다.

AI 솔루션을 활용하여 3배 빠르고, 10배 저렴하게 신약개발과정에서 후보물질을 개발할 수 있다는 점을 내세워 차세대 바이오 기업으로 꼽히는 점과, SK C&C, 한미약품 등 국내 대형 제약사와 꾸준히 개발협력을 유지해오고 있다는 점, 곧 상장을 앞두고 있다는 점에서 본 리서치팀은 동사의 가치를 다시 판단해야 할 때라고 제안한다.

투자지표	2019	2020	2021	2022	2023F
매출액 (백만 원)	38	33	55	27	N/A
영업이익 (백만 원)	(3,899)	(8,264)	(17,417)	(24,883)	N/A
영업이익률 (%)	-	-	-	-	N/A
순이익 (백만 원)	(6,227)	(45,713)	(55,053)	7,173	N/A
순이익률 (%)	-	-	-	266	N/A





S.T.A.R Research

Company Report 23.11.11

Team 2' s Opinion

N/R

목표주가	N/A
현재주가	N/A
상승여력	-

Consensus	2022	2023F
-----------	------	-------

S.T.A.R Research

권예준

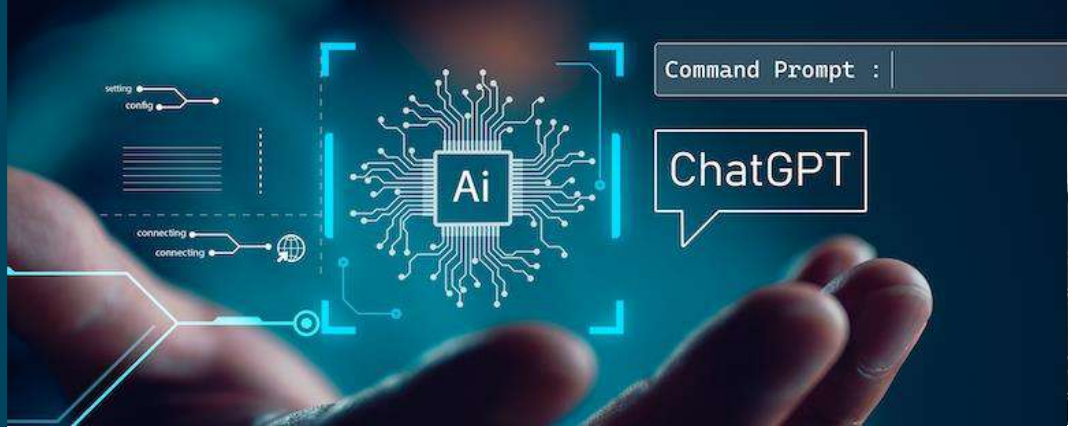
010.9022.8720  
yejoon0309@g.skku.edu  
46<sup>th</sup> member

김원영

010.2756.1059  
wykim1059@gmail.com  
46<sup>th</sup> member

양예지

010.3615.1527  
yellowyeji@g.skku.edu  
46<sup>th</sup> member



# 트웰브랩스 (비상장 기업)

## 글로벌 빅테크와 손잡은 영상이해 초거대 AI 개발 스타트업

### 국내 최초 엔비디아의 투자를 받은 AI 스타트업!

동사는 20년 3월 설립된 영상이해 초거대 AI 개발 기업으로, 멀티모달 신경망을 기반으로 영상 분야 AI 모델을 자체 개발하는 생성형 AI 기업이다. 20년 기준 매출 7억 6,363만원, 영업이익 4억 6,310만원을 기록하였고, 이후 공시된 실적은 없다.

동사는 23년 10월 엔비디아, 인텔과 같은 글로벌 빅테크 기업으로부터 총 1,000만 달러(한화 약 140억 원) 규모의 전략적 투자를 유치한 바 있다. 이는 엔비디아가 한국 AI 스타트업에 투자한 최초 사례로, 동사는 향후 글로벌 반도체 업체들과의 긴밀한 협력 관계를 통해 탄탄한 기술력/자금력을 확보해 나갈 것으로 판단된다.

### ‘페가수스’ 타고 멀리, 더 높이 날아갈 준비 완료!

23년 11월 9일, 동사는 영상을 이해하는 멀티모달 AI 모델인 ‘페가수스(Pegasus-1)’를 공개했다. 이는 동사가 자체 개발한 매개변수 800억 규모의 모델로, VLFM(대형 영상 언어 모델, Video Language Foundation Model)이다. 동사는 해당 모델 개발을 위해 글로벌 최대 규모의 영상-텍스트 쌍 3억 여개로 구성된 데이터셋 중 약 10%인 3,500만개를 자체 구축하고, 학습에 활용했다. 이에 따라 페가수스는 대용량의 영상을 1) 정확하고 정교하게 이해하고, 2) 텍스트로 요약할 수 있으며, 3) 질의응답이 가능한 최초의 모델이다.

최근 Open AI는 ‘image-to-text’, ‘text-to-text’ 기능이 있는 멀티모달 LLM GPT4-V를 출시한 바 있는데, 이에 비해 동사의 모델은 ‘video-to-text’가 가능하다는 점에서 경쟁력을 갖는다.

동사는 국내 스타트업이지만, 주로 미국에서 활동하고 있다. CB 인사이트는 동사의 기술력과 잠재력을 인정해 ‘세계 100대 AI 기업’과 ‘세계 50대 생성형 AI 스타트업’에 선정된 바 있으며, BI(비즈니스 인사이더)가 발표한 ‘2023년 최고 유망 AI 스타트업 34’에 선정되기도 했다. 향후 비디오 기반의 생성형 AI 기술을 기반삼아 글로벌 선도 AI 업체로 도약할 동사에 대한 주목이 필요한 시점이라 판단된다.