

AI

AT TURNING POINT

전환점 혹은 변곡점에서

성장산업 조태나

02)368-6174_ terna@eugenefn.com

인터넷/게임/우주 정의훈

02)368-6170_ uihoon0607@eugenefn.com

Contents

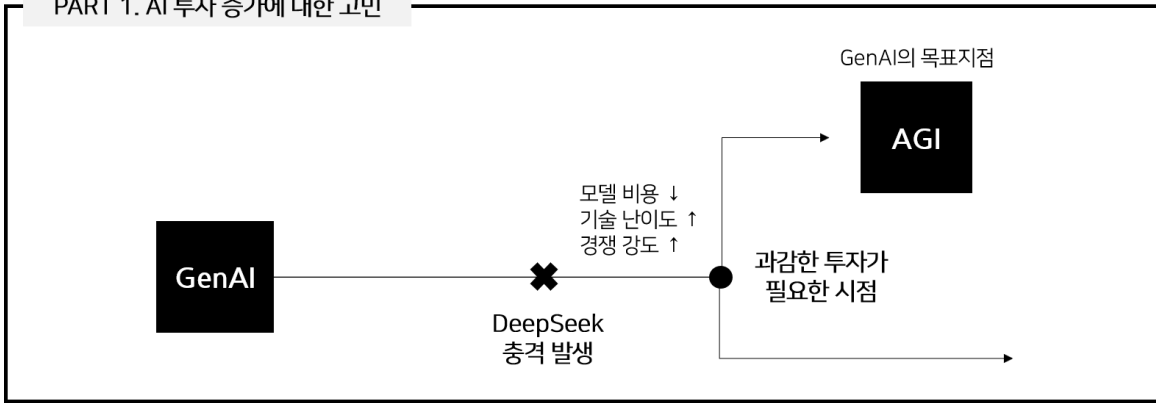
[Part1] 우리는 어느 지점에 서 있나	4
AI Hype Cycle	4
생성형 AI 이해하기	6
AI 글로벌 투자 증가에 대한 고민	11
최근 모델 비용 하락을 어떻게 봐야할까	11
경량화를 이끈 최신 기술 변화	13
비용이 낮아져도 우리는 투자를 늘리겠다	21
여전히 대규모 투자가 필요한 이유	25
[Part 2] 네카오의 AI 전략 분석	33
네이버	33
카카오	42
Game with AI	49
인간 챔피언을 이긴 최초의 AI, 딥블루	49
딥러닝의 알파고	51
언어모델은 게임을 어떻게 바꿀까	53
기업분석	63
NAVER(035420)	
카카오(035720)	
크래프톤(259960)	

AI at Turning Point

Executive Summary

By 조태나, 정의훈

PART 1. AI 투자 증가에 대한 고민



PART 2. 국내 소프트웨어 플랫폼 AI 전략 분석

	Phase 1 (AI 플랫폼)	Phase 2 (초기 언어모델)	Phase 3 (서비스 기반 언어모델)	Phase 4 (현재)
NAVER	CLOVA	HyperCLOVA	HyperCLOVA X	On-Service AI
KAKAO	KAKAO i	KoGPT1.0	(KoGPT2.0)	AI Orchestration
GAME	DeepBlue	AlphaGo	LLM의 부상	AI NPC

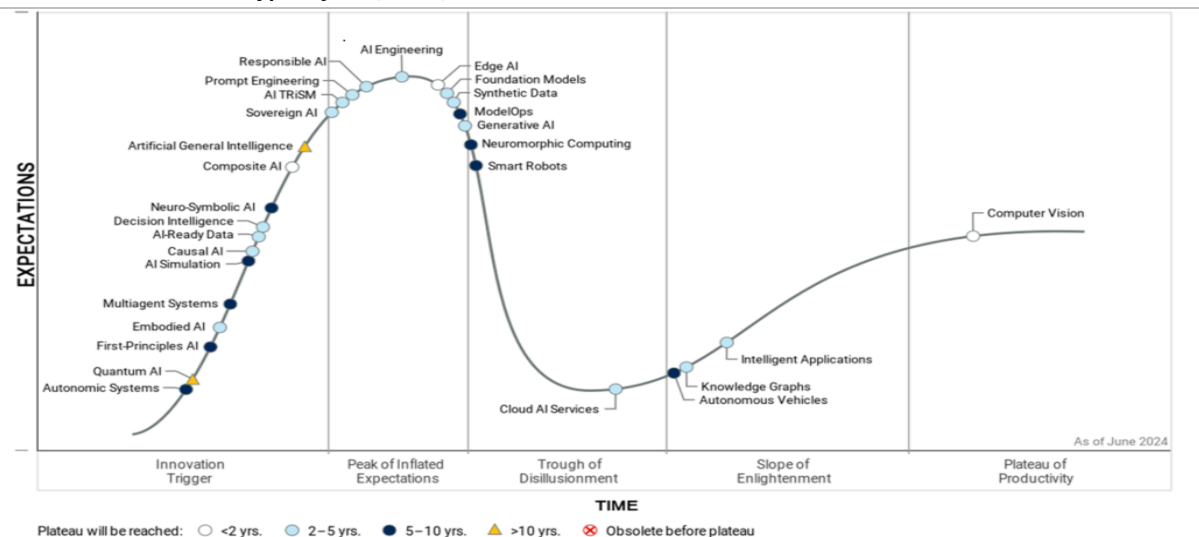
[Part 1] 우리는 어느 지점에 서 있나

AI Hype Cycle

모든 기술이 일정한 발전 단계를 거치는 것처럼 AI 도 하이프 사이클이라는 개념을 통해 기술의 부상과 성숙도를 평가할 수 있다. 하이프 사이클은 새로운 기술이 시장에서 어떻게 도입되고 정착하는지를 나타내며, 이를 통해 현재 AI 기술이 어느 단계에 있는지 살펴볼 수 있다.

하이프 사이클은 다섯 개의 주요 단계로 나뉜다. 첫째 혁신촉발 단계(Innovation Trigger)에서는 새로운 기술이 등장하고 연구가 활발히 이루어지지만, 실질적인 상용화는 이루어지지 않은 상태이다. 둘째, 과대 기대의 정점(Peak of Inflated Expectations)에서는 시장의 기대감이 극도로 높아지며 투자와 관심이 몰린다. 그러나 아직 기술적으로 해결해야 할 문제가 많아 과장된 기대가 형성되기도 한다.

도표 1. Gartner – AI Hype Cycle (2024)



자료: Gartner, 유진투자증권

셋째, 환멸의 골짜기(Trough of Disillusionment)에서는 기대에 비해 실질적인 성과가 부족해 관심과 투자가 감소하는 시기이다. 넷째, 계몽의 경사(Slope of Enlightenment)에서는 기술이 개선되고 현실적인 기대 수준으로 조정되며, 실용적인 사례들이 나타나기 시작한다. 마지막으로 생산성의 정점(Plateau of Productivity)에서는 기술이 산업에 정착하고 실질적인 가치를 창출하게 된다.

현재 생성형 AI 는 과대 기대의 정점(Peak of Inflated Expectations)에 위치하고 있다. 많은 기업들이 생성형 AI 를 도입하고 있으며, 이 기술이 다양한 산업에서 혁신을 가져올 것이라는 기대가 형성되어 있다. 그러나 여전히 해결해야 할 모델의 신뢰성, 비용 문제, 데이터 프라이버시, 규제 이슈 등이 존재한다. 이러한 이유로 생성형 AI 는 향후 몇 년 안에 환멸의 골짜기(Trough of Disillusionment)를 거칠 가능성이 높으며, 기대했던 성과 대비 실질적인 한계가 부각될 수 있다. 그러나 장기적으로는 생성형 AI 가 실용화 단계까지 도달하며, 콘텐츠 생성, 비즈니스 프로세스 자동화, 연구 및 데이터 분석 등의 분야에서 핵심 기술로 자리 잡을 전망이다.

생성형 AI 의 목표지점, AGI 의 실현 가능성

AGI(범용 인공지능) 역시 현재 과대 기대의 정점(Peak of Inflated Expectations)에 있으며, 아직 구체적인 실현 가능성이 검증되지 않았다. 기존의 AI 모델들은 특정 작업을 수행하는 데 최적화되어 있지만, 인간처럼 다양한 과제를 수행할 수 있는 범용 인공지능의 실현은 기술적으로 매우 어려운 문제로 남아 있다. 현재 AGI 는 최소 10 년 이상 소요될 것으로 예측되지만, 최근 기술 발전 속도가 빨라지면서 예상보다 조기 실현될 가능성에 대해서도 생각해 볼 필요가 있다.

특히 o3, DeepSeek R1 등의 최신 모델들은 기존 모델들 보다 훨씬 뛰어난 추론 능력을 보여주고 있다. 추론을 고도화하는 다양한 기술 연구가 진행되면서 AGI 실현 가능성을 높이는 요소로 작용하고 있다.

그러나 생성형 AI 가 목표치에 도달하려면 여전히 많은 투자가 필요하다고 생각한다. LLM 을 고도화하는 연구는 정점을 지나고 있지만, LLM 의 궁극적 목표에 도달하기에는 아직도 많은 과제가 남아있기 때문이다. 투자를 통해 AGI 실현 시점을 얼마나 앞당길 수 있을지가 앞으로의 관건이 될 것이라고 본다.

생성형 AI 이해하기

생성형 AI와 LLM

생성형 AI(GenAI, GenerativeAI)는 새로운 콘텐츠(텍스트, 이미지, 음악, 코드 등)를 만들어내는 인공지능 기술을 뜻한다. 즉, 기존 데이터를 바탕으로 완전히 새로운 결과물을 생성하는 AI 이다. 예를 들어 ChatGPT 처럼 글을 쓰는 AI 가 있는가 하면, Midjourney, DALL-E 처럼 그림을 그리는 AI 가 있다. 뿐만 아니라, 작곡을 하거나 프로그래밍 코드를 생성하는 AI 도 있다. 생성형 AI는 기존의 데이터에서 패턴을 학습하고 이를 바탕으로 새로운 결과물을 만들어낸다는 점에서, 단순한 정보 검색 AI(예: 챗봇)와는 차이가 있다.

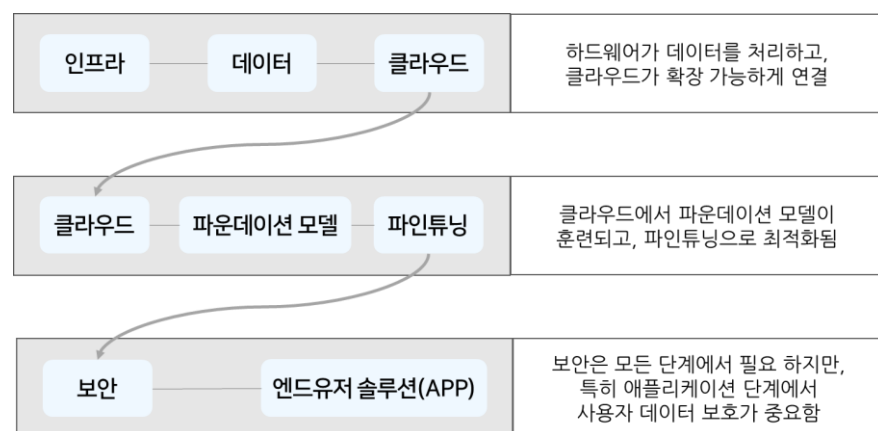
반면 LLM(Large Language Model)이란 엄청난 양의 텍스트 데이터를 학습한 인공지능 모델로, 사람처럼 자연스러운 문장을 생성하고 이해할 수 있는 기술이다. 쉽게 말해, 많은 책과 문서를 읽고 훈련된 AI 가 사람의 질문에 답하거나, 글을 쓰거나, 대화를 나누는 것이다. LLM은 문장을 예측하는 방식으로 작동한다. 예를 들면 “오늘 날씨가” 라고 입력하면, 그 다음에 올 가능성이 높은 단어를 예측하여 “좋은 것 같아”와 같은 문장을 만들어낸다. 이런 방식을 반복해서 적용하면, 질문에 대한 답변을 하거나, 스토리를 만들어낼 수도 있다.

쉽게 비유를 하자면 생성형 AI 는 예술가(창작자) 전체를 의미하고 LLM 은 그 중에서도 작가라고 볼 수 있다. LLM은 생성형 AI의 일부이며, 언어를 다루는 AI로서 질문에 답하고, 문서를 요약하고, 스토리를 창작하는 역할을 한다.

생성형 AI 밸류체인에 대한 이해

생성형 AI의 밸류체인은 AI 기술의 개발부터 최종 사용자 애플리케이션까지의 과정을 단계별로 나누어 이해할 수 있다. 인프라, 데이터, 클라우드 컴퓨팅 아키텍처, 파운데이션 모델, 파인튜닝, 보안, 엔드유저 솔루션(APP) 총 8 단계의 밸류체인으로 구성되어 있다.

도표2. 생성형 AI 밸류체인 간 연계성



자료: 유진투자증권

인프라는 AI 모델을 개발하고 실행하기 위한 하드웨어 기반을 제공한다. 여기에는 GPU, TPU 와 같은 고성능 컴퓨팅칩, 데이터 센터 설비, 네트워크 장비 등이 포함된다. 이 단계는 AI 의 연산 능력을 뒷받침하는 물리적 토대다. 대표적인 기업으로는 엔비디아, 브로드컴, TSMC, 삼성전자, SK 하이닉스 등이 있다.

데이터는 AI 모델 훈련의 원료로, 이 단계에 포함되는 기업은 데이터 수집, 처리, 정제, 라벨링 등을 한다. 데이터의 품질과 양이 모델 성능에 직접적인 영향을 미치기 때문에 생각보다 중요한 작업이다. 특히 DeepSeek 는 데이터 필터링 및 정제 과정을 거쳐 높은 품질의 데이터를 학습시켰고, 모델이 불필요한 정보를 학습하지 않도록 최적화하여 연산 비용을 절감했다. 데이터 카테고리의 대표적 기업으로는 데이터 라벨링과 정제 서비스로 LLM 을 지원하는 Scale AI, 데이터 웨어하우스 플랫폼(데이터 저장 및 분석) Snowflake, LLM 훈련 데이터를 관리하는 Databricks, 클라우드 기반 데이터 분석으로 대규모 데이터셋을 처리하는 BigQuery 가 있다.

클라우드 컴퓨팅 아키텍처 단계에서는 데이터와 인프라를 연결하여, AI 모델 훈련과 배포를 위한 확장 가능한 컴퓨팅 환경을 제공한다. 멀티 클라우드, 하이브리드 클라우드, GPU 클러스터 관리 등이 포함되며, 대표적으로 아마존 AWS, 마이크로소프트 Azure, 구글 클라우드, CoreWeave 등이 있다.

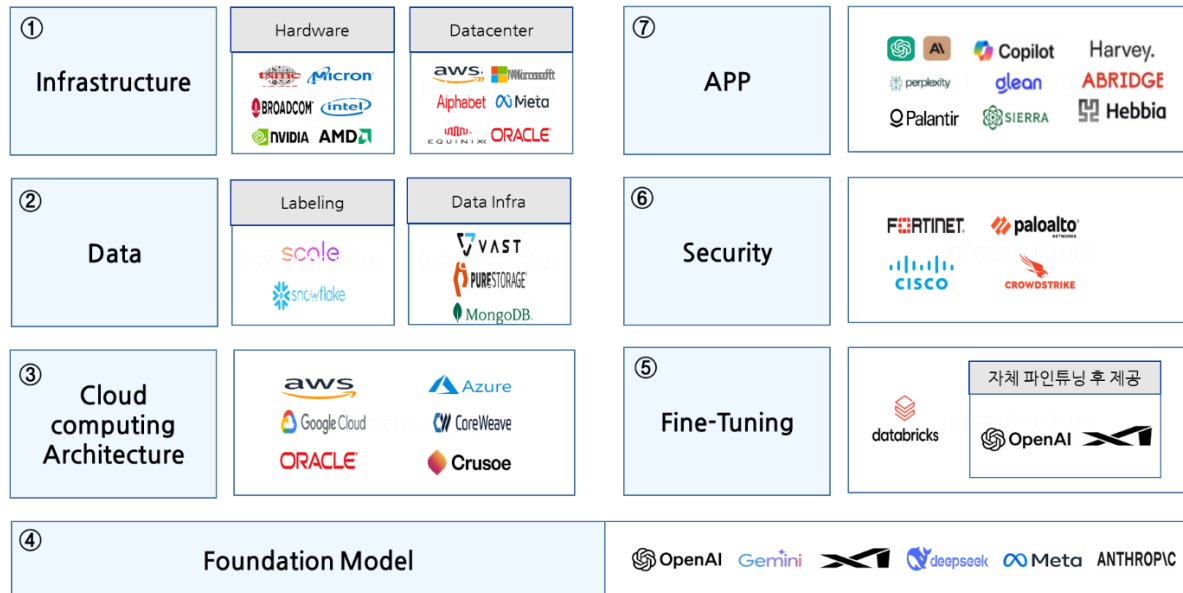
파운데이션 모델의 경우 생성형 AI 밸류체인의 핵심으로 볼 수 있다. 범용적인 AI 모델을 개발하며, 대규모 데이터로 사전에 훈련된 모델을 제공한다. 모델은 이후 특정 작업에 맞게 파인튜닝되는 단계를 거친다. 파운데이션 모델 기업에는 본 보고서의 주요 분석 대상인 OpenAI, Google, Meta AI, xAI 등이 있다. 파인튜닝은 파운데이션 모델을 특정 도메인이나 작업에 맞게 조정한다. 이를 통해 모델의 정확도와 효율성이 증가하게 된다. 오픈소스 모델과 파인튜닝 도구(Transformers)를 제공하는 Hugging Face, Mosaic AI로 파인튜닝을 지원하는 Databricks 등이 있다.

보안은 AI 시스템의 데이터, 모델, 배포 환경을 보호하고 프라이버시와 규제 준수를 보장한다. 데이터의 유출, 모델 도난, 악성 공격 방어가 보안 섹터에 포함된다. 암호화와 접근제어 시스템, 모델 보안 테스트 등이 핵심 기술이다. 대표적인 기업으로는 CrowdStrike, Palo Alto Networks가 있다.

엔드 유저 솔루션, 즉 어플리케이션 부분은 최종 사용자에게 AI 기능을 제공하는 소프트웨어 및 서비스를 개발한다. 여기에는 챗봇, 이미지 생성, 산업별 솔루션 등이 포함된다. 이미지 생성 AI인 Midjourney, CRM에 AI를 통합한 Salesforce 등이 있다.

AI 밸류체인은 인프라에서 시작해 데이터와 클라우드를 거쳐 모델 개발, 보안, 그리고 최종 앱으로 이어지는 연속적 과정이다. 각 단계는 독립적이면서도 상호 의존적이며, 대표 기업들은 해당 분야에서 기술적 우위를 점하고 있다.

도표 3. 생성형 AI 밸류체인 정리



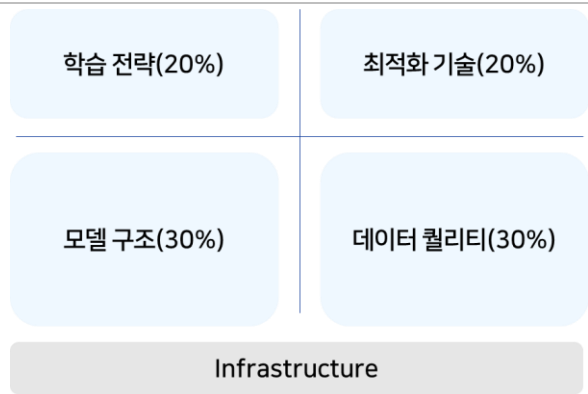
자료: 유진투자증권

모델 성능에 영향을 미치는 요소들

환경에 따라 상황이 크게 달라질 수 있지만, '아주 단순하게' 생각한다면 모델에 영향을 미치는 요소는 크게 4 가지 정도로 분류를 할 수 있다.

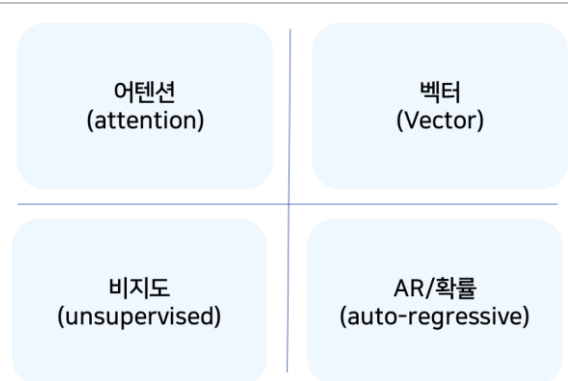
첫째, 모델의 기본 설계가 얼마나 효율적이고 강력한지, **모델의 구조가 중요하다**. 모델의 전반적인 아키텍처는 성능에 약 30% 정도 기여한다. 두 번째로 중요한 것은 **데이터의 질**이다. 데이터가 깨끗하고 다양하며 대표성을 띠 수록 모델 성능 향상에 긍정적이다. 이 역시 30% 정도 기여한다. 세 번째는 **학습전략**으로, 사전 학습, 파인튜닝 등과 같은 학습 방법들이 모델이 특정 작업을 수행할 수 있도록 한다. 모델에 20% 정도 영향을 끼친다. 마지막으로 **최적화 기술** 역시 20% 정도 영향을 미치는데, 연산 효율, 희소 활성화(Sparse Activation), 하드웨어 최적화 등이 성능에 영향을 미친다(DeepSeek 는 이 부분을 극단적으로 향상시켰다). 이 비율은 매우 단순화된 가정이며, 실제로는 모델 개발 과정에서 각 요소가 긴밀한 상호작용을 하기 때문에 크게 달라질 수 있다.

도표4. 모델 성능에 영향을 미치는 요소들



자료: 유진투자증권

도표5. LLM 알고리즘 함수



자료: 유진투자증권

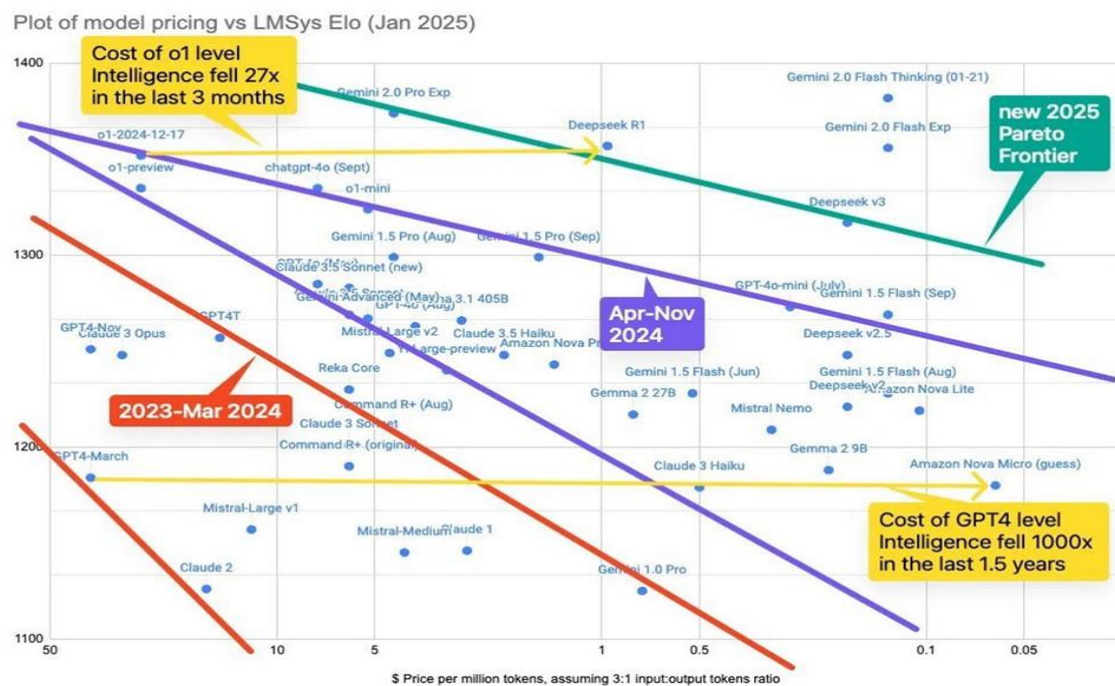
AI 글로벌 투자 증가에 대한 고민

최근 모델 비용 하락을 어떻게 봐야할까

DeepSeek 의 비용 효율화에 온 세상이 놀랐다. 그런데 잘 살펴보면 R1 의 비용 감소는 생성형 AI 발전 속도를 보았을 때 추세적으로 크게 벗어난 것이 아니다. 다리오 아모데이 역시 이에 대해 언급했는데, '추론 비용 붕괴는 원래 AI 개선의 특징'이라는 것이다.

"알고리즘 최적화로 인해 10 배의 비용 개선을 가져올 수 있습니다. DeepSeek 는 7 - 10 개월 더 오래된 미국 모델과 비슷한 성능의 모델을 훨씬 저렴한 비용으로 내놓은 것뿐이고, AI 개발비용 경제학을 근본적으로 바꾸는 독특한 게 아니라 단지 추세에 맞는 수준입니다"

도표 6. 가격 대비 성능 최적화와 파레토 프론티어의 이동



자료: Latent Space, 유진투자증권

가격 대비 성능 최적화와 파레토 프론티어의 이동

AI 모델의 비용 대비 성능 변화를 분석한 결과, 2023년부터 2025년까지 파레토 프론티어(Pareto Frontier)가 빠르게 이동하고 있다. 2023년 3월부터 2024년 3월까지의 비용이 크게 감소했으나, 일부 모델들은 여전히 높은 가격을 유지했다. 그러나 2024년 4월부터 11월까지 더 많은 AI 모델들이 등장하면서 가격이 더욱 낮아지고 성능이 최적화되었다. 특히, DeepSeek R1 과 Gemini 2.0 Flash Thinking 모델이 2025년 파레토 프론티어를 형성하며, 가격 대비 성능이 우수한 모델로 평가되고 있다.

AI 모델의 비용 대비 성능은 기하급수적으로 개선되고 있으며, 이는 AI의 접근성과 경제성을 획기적으로 높이는 결과를 가져오고 있다. **GPT-4 수준의 인공지능 운영 비용이 1.5년 만에 1,000배 감소하면서, AI 활용이 점점 대중화되고 있다.** 2025년의 파레토 프론티어는 2023~2024년보다 훨씬 최적화된 AI 모델들이 등장했음을 보여준다. 향후 AI 시장에서 저비용·고성능 모델을 활용한 업무 자동화와 비즈니스 혁신이 본격화되면서, 기업과 개인 모두 AI 모델을 더욱 효과적으로 활용할 수 있는 환경이 조성되고 있다. 결과적으로, AI 모델의 가격이 지속적으로 낮아지고 성능이 향상되면서, AI 시장은 폭발적인 성장세를 지속할 것으로 예상된다. AI 기술의 경제성이 더욱 향상됨에 따라, 기업들은 AI를 적극적으로 도입하여 운영 효율성을 높이고, 새로운 비즈니스 기회를 창출할 가능성이 크다.

경량화를 이끈 최신 기술 변화

최근 AI 모델의 학습 방식과 최적화 기법이 급격히 변화하고 있다. 기존에는 휴먼 피드백을 기반으로 한 강화 학습(RLHF, Reinforcement Learning from Human Feedback)을 활용하여 모델 성능을 개선하는 방식이 주를 이루었다. 그러나 RLHF 방식은 높은 비용과 피드백 데이터의 품질 문제로 인해 한계에 부딪히기 시작했다. 이에 따라 최신 AI 모델들은 타임스케일링 기반 자기학습(Test-Time Scaling) 기법을 적용하고 있다. **Grok-3와 DeepSeek R1과 같은 최신 모델들은 테스트 단계에서 보다 깊이 사고할 수 있도록 설계되어 있으며, 이는 추가적인 RLHF 없이도 스스로 학습하고 사고의 깊이를 증가시키는 방향으로 발전하고 있다.** 이로 인해 향후 AI 모델들은 외부 데이터 없이도 자기 학습(Self-Learning)을 수행할 가능성이 높아지고 있다.

결론적으로, AI 모델의 학습 방식과 최적화 기법이 RLHF에서 자기학습으로, MoE에서 MLA로, 단일 모델에서 멀티 에이전트 구조로 변화하고 있으며, 학습 데이터 활용 방식 역시 RAG로 전환되고 있다. 또한, 장기적인 문맥 유지 기능이 강화됨에 따라 AI의 기억력이 보다 발전할 것으로 보인다. 이러한 변화들은 AI의 성능과 효율성을 높이며, 앞으로 더욱 강력하고 실용적인 AI 시스템으로의 발전을 예고하고 있다.

도표 7. LLM 최신 기술 변화

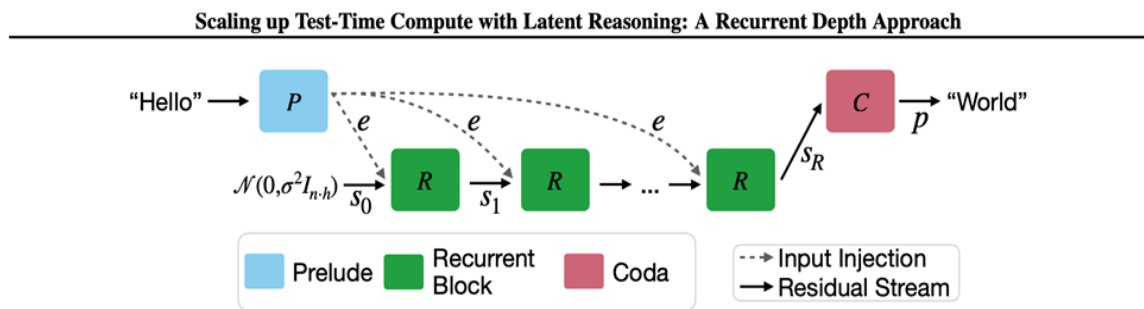
기존 방식	변화된 방식	핵심 모델 & 특징
파인튜닝 + RLHF	타임스케일링 기반 자기학습	Grok-3, DeepSeek R1
MoE(Experts 시스템)	MLA (뉴런 단위 최적화)	DeepSeek R1
단일 거대 모델(Dense)	멀티 에이전트 AI	Grok-3, Gemini
사전 학습 데이터 기반 응답	RAG 실시간 검색 활용	Grok-3, GPT-4o
짧은 컨텍스트 윈도우	AI Memory 장기 기억	Claude 3, Grok-3

자료: 유진투자증권

1) Test-Time Scaling 기반 자기 학습: 스스로 생각하는 모델의 등장

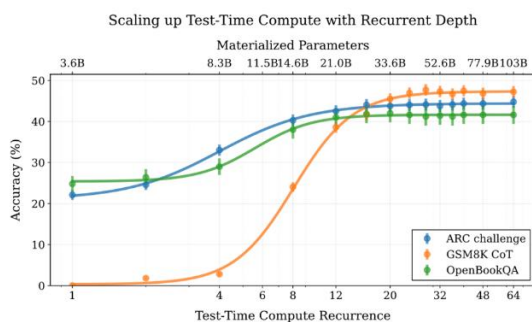
최근 공개된 한 연구 논문 *Scaling up Test-Time Compute with Latent Reasoning: A Recurrent Depth Approach*(17 Feb 2025)에서는 대형 언어 모델을 경량화 할 수 있는 흥미로운 방법을 제시했다. 이 접근법의 핵심은 모델 자체를 거대화 하지 않으면서도 필요한 순간에 추론 능력을 강화하는 것이다. 이 연구는 모델이 답을 내기 전에 내부적으로 여러 번 “생각”할 시간을 주는 기법을 사용했다. 일반적인 언어 모델은 한 번에 한 토크(token)씩 출력하면서 답변을 만들어내지만, 이 모델은 출력 전 모델 안에서 여러 차례 연산을 돌린다. 쉽게 말해, 작은 두뇌에게 생각할 시간을 더 주어서 어려운 문제도 풀 수 있게 한 것이다.

도표 8. 모델이 답을 내기 전에 스스로 여러 번 생각하게 함



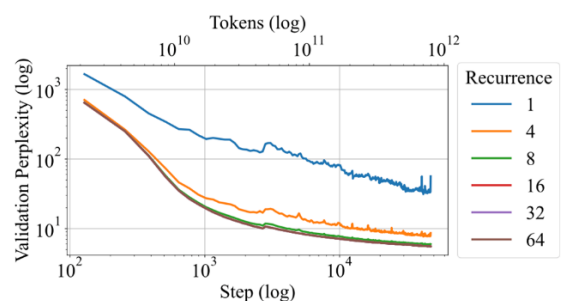
자료: Scaling up Test-Time compute with Latent Reasoning(2025) 논문 발췌, 유진투자증권

도표 9. 재귀 연산 횟수 증가할수록 정확도 증가



자료: Scaling up Test-Time compute with Latent Reasoning(2025) 논문 발췌, 유진투자증권

도표 10. 반복 횟수에 따라 모델 성능 향상



자료: Scaling up Test-Time compute with Latent Reasoning(2025) 논문 발췌, 유진투자증권

그 결과, 파라미터 35 억개 규모의 비교적 작은 모델임에도 불구하고, 복잡한 추론이 필요한 벤치마크에서 마치 500 억개 규모 모델에 준하는 성능까지 향상되는 모습을 보였다. 다시 말해, 모델에 추론(reasoning) 회로를 넣었더니 작은 모델도 큰 모델 수준의 성능을 일부 달성한 것이다. [도표 9]는 이 개념을 실험으로 증명한 것으로, 재귀 연산 횟수가 늘어날수록(가로 축) 모델의 정답 정확도(세로 축)가 향상되어, 충분히 반복했을 때는 수십 배 큰 모델에 근접함을 보여준다.

이 접근법의 장점은, 특별한 훈련 데이터 없이도 이러한 능력을 얻어냈다는 점이다. 기존에는 모델이 중간추론을 할 수 있도록 하기 위해 CoT(chain-of-thought) 방식으로 사람이 중간단계를 설명한 예시들을 많이 훈련시키곤 했다. 하지만 이 논문의 방법은 일반 텍스트 데이터만으로 모델을 훈련하고, 대신 훈련 단계에서 모델이 랜덤하게 여러 번 숨겨진 연산을 하도록 만들어 주었다. **이렇게 추론의 깊이를 가변적으로 활용하는 모델은, 추론할 때 간단한 질문에는 몇 번만 반복하고 어려운 질문에는 여러 번 반복하는 식으로 유연하게 계산량(연산)을 조절할 수 있다.** 이는 곧 동일한 모델이 상황에 맞게 계산 자원을 조절함으로써, 실용적인 효율을 높일 수 있음을 뜻한다. (예를 들어, 쉬운 질의 응답은 빠르게 처리하고, 난해한 문제만 추가 연산을 쓰는 식)

2) MLA(Multi Head Latent Attention): DeepSeek 최적화의 핵심

MoE 에서 MLA 로의 모델 최적화 방식 변화 역시 크게 눈에 띈다. MoE(mixture of Experts)는 하나의 거대 모델 안에 여러 개의 전문가(Experts) 모델을 두고, 입력 데이터에 따라 일부 전문가만 활성화하여 연산하는 방식이다. 반면, MLA 는 MoE 와 비슷하지만 정밀한 어텐션 기법을 사용해서 연산을 최적화하고, 뉴런단 위까지 세분화 해서 모델을 활성화한다.

기존에는 Experts 중 일부를 활성화하여 연산을 최적화하는 MoE 기법이 사용되었다. 그러나 MoE 방식은 라우터(Router) 선택 과정에서의 비효율성과 특정 experts만 활용되는 문제로 인해 최적화에 어려움이 있었다. 이에 대한 대안으로 DeepSeek R1 등에서는 MLA 기법을 적용하고 있다. **MLA 는 뉴런 단위에서 필요한 부분만 활성화하는 방식으로 보다 정밀한 최적화가 가능하며, MoE 보다 효율적인 연산 최적화를 실현할 수 있다.** 이에 따라 앞으로의 AI 모델들은 MoE 보다 더욱 세밀한 최적화를 통해 연산 비용을 절감하는 방향으로 발전할 것으로 예상된다.

도표 11. MoE 와 MLA 비교

	MoE (Mixture of Experts)		MLA (Multi-Head Latent Attention)	
기본 개념	Experts 를 분리하여 일부만 활성화		Attention 기법 활용해 필요한 뉴런만 활성화	
연산 최적화 방식	라우터가 Experts 선택		Attention 메커니즘이 개별 뉴런을 활성화	
활성화 방식	전문가 단위(예: 2~4 개만 활성화)		뉴런 단위(필요한 뉴런만 선택적으로 활성화)	
추론 속도	★	불필요한 Experts 차단	★★	더 정교한 연산 최적화
메모리 사용량	★	모든 Experts 를 쓰지 않음	★★	필요한 뉴런만 활성화

자료: 유진투자증권

도표 12. DeepSeek MoE, MLA 구조

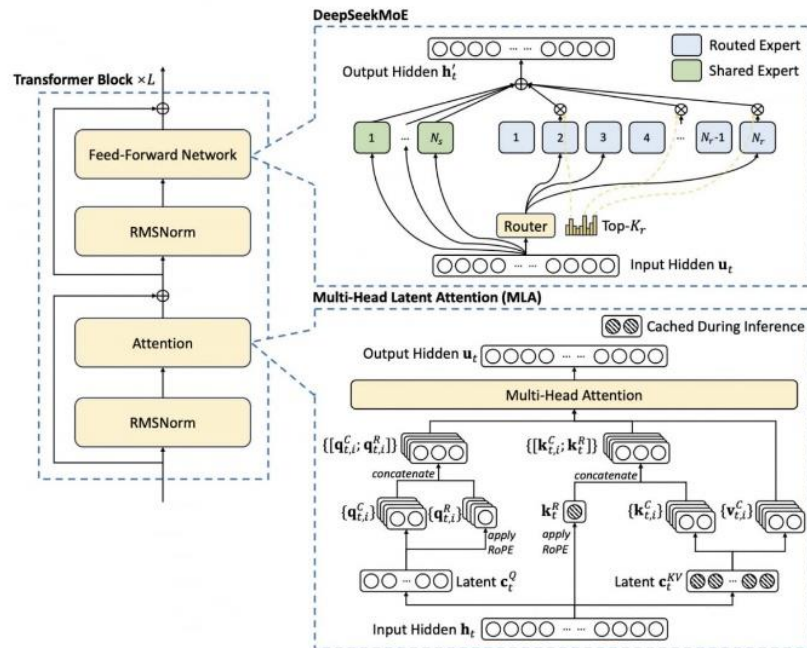


Figure 2 | Illustration of the basic architecture of DeepSeek-V3. Following DeepSeek-V2, we adopt MLA and DeepSeekMoE for efficient inference and economical training.

자료: DeepSeek, 유진투자증권

3) RAG 와 AI Memory: 정교함을 한 스푼 더

AI 의 학습 데이터 활용 방식도 변화하고 있다. 기존에는 사전 학습된 모델을 기반으로 응답을 생성했으나, 이러한 방식은 새로운 정보가 등장할 경우 즉각적인 업데이트가 어렵다는 단점이 있었다. 이를 해결하기 위해 **RAG(Retrieval-Augmented Generation)** 기법이 도입되고 있으며, 이는 AI 가 외부 데이터베이스나 실시간 검색을 활용하여 최신 정보를 반영할 수 있도록 돕는다. 예를 들어, Grok-3 는 DeepSearch 기능을 활용하여 실시간 검색을 수행하며, GPT-4o 와 Gemini 도 외부 데이터 검색을 통한 최신 정보 반영을 시도하고 있다. 이에 따라 향후 AI 모델들은 사전 학습된 정보만을 활용하는 것이 아니라, 외부 데이터를 실시간으로 검색하고 활용하는 방향으로 발전할 가능성이 높다.

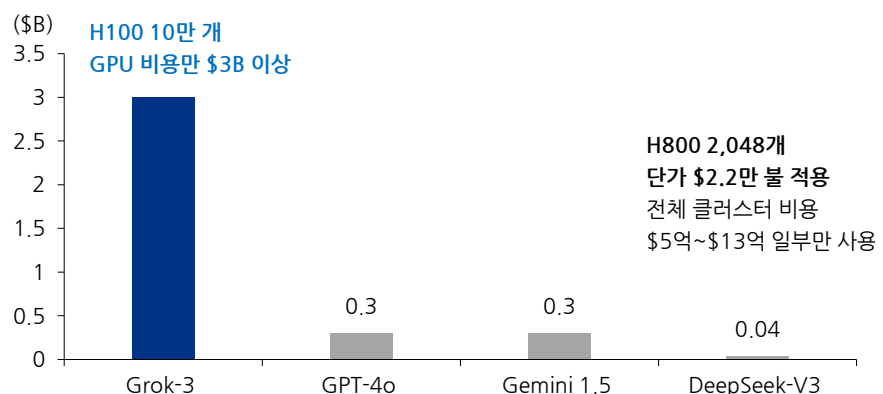
마지막으로, AI 의 장기 기억(AI Memory)이 중요하게 대두되고 있다. 기존에는 컨텍스트 윈도우가 제한되어 있어 AI 가 대화 내용을 장기적으로 유지하기 어려웠다. GPT-4o 와 Gemini 1.5 는 128k~100 만 토큰을 지원하지만, **실제로는 대화 맥락을 완벽하게 유지하는 것이 쉽지 않았다.** 이에 따라 AI 가 대화 이력을 장기적으로 기억하고 활용할 수 있도록 하는 장기 메모리 기술이 개발되고 있다. Grok-3와 Claude 3 는 이러한 장기 기억 기능을 지원하는 방향으로 발전하고 있으며, 향후 AI 는 단기 컨텍스트가 아닌 장기적인 대화 맥락을 유지하고 활용하는 방식으로 변화할 가능성이 높다.

최적화에 집중하는 DeepSeek, 그에 대항하는 금수저 xAI,

최근 LLM 의 최대 화두는 비용이기 때문에 모델별 비용 비교를 중심으로 글로벌 모델들의 전략을 확인해볼 수 있다. xAI Grok-3 스케일이 단연 돋보이는데, 10 만 대의 엔비디아 H100을 사용한 것으로 추정된다. H100 대 당 가격이 \$3만 수준임을 감안하면, GPU 에만 약 30 억 달러를 지불한 것이다. xAI 는 최적화 보다는 막대한 인프라를 이용해 성능을 이끌어내고 있다. (H100 과 더불어 Dojo 슈퍼컴퓨터 역시 활용했을 가능성이 있다) Grok-3 의 정확한 파라미터 수는 공개되지 않았지만, 2.7 조 개 이상일 것으로 추정되며(Grok-3 는 자신의 파라미터 수를 5,000 억 개라고 주장하지만) 매우 큰 Dense 모델일 가능성이 크다. GPT4 가 약 1.76 조+ 파라미터로, Grok 이 다른 모델들 보다 훨씬 더 크기 때문에 몇 배 더 많은 연산이 필요했을 것으로 추정된다.

DeepSeek 의 경우 H100 대신 다운 그레이드된 H800 을 써야 했지만, 소프트웨어 최적화와 비용 절감 전략을 통해 주어진 자원으로 높은 성과를 냈다. 훈련비용에 대해 많은 논란이 있었지만, 결과적으로 DeepSeek 가 소프트웨어 최적화를 통해 학습 비용을 극단적으로 줄인 것은 사실로 드러났다. 앞으로도 모델 경량화와 최적화에 대한 전략은 바뀌지 않을 것으로 보인다. Google은 자체 AI 칩(TPU)과 Google Cloud를 활용하여 훈련 및 추론을 최적화하고 있는데, TPU는 구글의 자체 가속기칩으로, 성능당 비용이 GPU 대비 낮지만 Gemini 개발 시 엔비디아의 H100도 함께 사용했다는 정보가 있다. OpenAI의 경우 MS와의 전략적 제휴를 통해 Azure 클라우드 기반으로 운영되었으나, 최근 CoreWeave와 119억 달러 계약(5년)을 체결하여 기반 인프라에 많은 변화가 있을 예정이다.

도표 13. LLM 훈련 비용(GPU 비용) 추정치



자료: 언론종합, 유진투자증권 추정치

투자 비용을 통해서도 알 수 있듯, 글로벌 LLM 들의 비용 및 시장전략에는 크게 차이가 있다. OpenAI 와 Google 은 API 제공 및 기업 맞춤형 AI 서비스 판매에 집중하고 있으며, xAI 는 실시간 검색과 머스크 생태계(Tesla, X, SpaceX 등)와의 연계를 통해 자체 활용 가능성이 높다. DeepSeek 는 오픈소스를 통해 LLM 생태계를 확장하고, 경량화된 모델로 기업용 컨설팅 및 맞춤형 모델 서비스로 수익을 창출할 것으로 보인다. 생각보다 OpenAI 의 위치가 매우 빈약한 상황이다.

도표 14. 모델 별 훈련 비용 및 인프라 비교

	GPT-4o	Gemini 1.5	Grok-3	DeepSeek-V3
훈련 비용 (하드웨어)	3 억 달러+	3 억 달러+	30 억 달러+	4,000 만 달러+
GPU/TPU	A100 2~2.5 만 H100 1~1.5 만	A100 2.5 만개, TPU 사용가능성↓	NVIDIA H100 10 만개	NVIDIA H800 2,048 개
사전학습 데이터량	10 조+ 토큰 (비공개)	10 조+ 토큰 (구글 데이터 포함)	12.8 조+ 토큰 (실시간 검색 포함)	14.8 조+ 토큰 (중국어 특화)
API 사용 비용	\$2.50 / 1M (입력), \$10 / 1M (출력)	\$2 / 1M (입력) \$8 / 1M (출력)	약\$12~\$18 / 1M	\$0.27 / 1M (입력), \$1.10 / 1M (출력)
운영인프라	MS Azure	자체 인프라 활용	Tesla Dojo 활용 가능	중국 클라우드 및 자체 데이터센터 운영
특이점	API 경제 (구독/사용량 기반),	Google 제품과 연계 활용	X (트위터)와 통합, 실시간 검색 및 AI 에이전트 기능 특화	완전 무료 사용 가능 (오픈소스 제공)

자료: 언론종합, 유진투자증권 추정치

비용이 낮아져도 우리는 투자를 늘리겠다

경이로운 속도와 규모의 xAI의 데이터센터

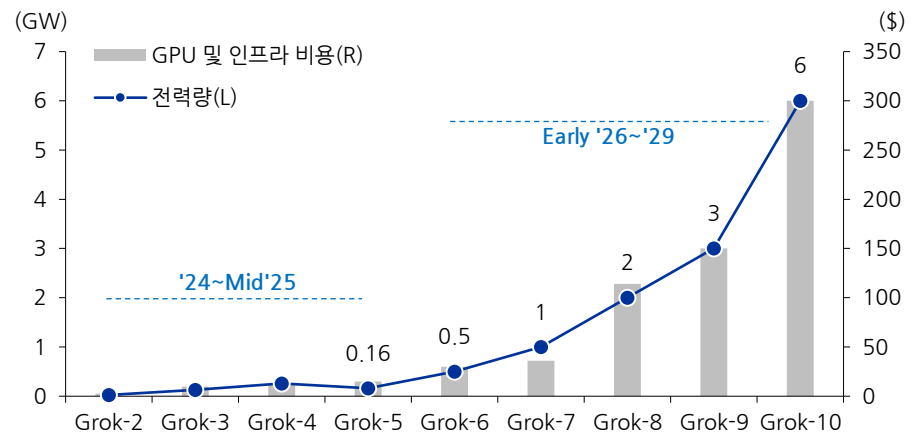
일론 머스크의 xAI는 AI 개발에 있어 다른 기업들과 차별화되는 중요한 전략을 채택하고 있다. 대규모 데이터 센터를 운영하며, 이를 통해 AI 모델의 훈련과 추론에 필요한 막대한 컴퓨팅 자원을 공급한다는 점이다. xAI의 데이터 센터 규모는 단순히 기술적 요구 사항을 충족하는 것 이상으로, 회사의 경쟁력을 결정짓는 핵심 요소로 작용하고 있다.

테네시주 멤피스에 위치한 'Colossus' 데이터센터는 전체 20만 개의 GPU를 보유하고 있다. 'Cortex' 역시 GPU를 5만개 이상을 보유한 클러스터로, xAI는 슈퍼 클러스터 두 곳을 기반으로 빠른 성장을 이어가고 있다. 특히 젠슨 황이 "콜로서스는 단일 클러스터로는 세계에서 가장 빠른 슈퍼컴퓨터"라고 언급하였으며, 같은 규모의 데이터 센터를 구축하려면 일반적으로 3년이 소요되지만, 콜로서스는 단 122일만에 완공하는 놀라운 실행력을 보여주었다.

무엇보다 주목해야 할 점은 xAI가 슈퍼 클러스터를 운영하기 위한 전력 및 냉각 시스템을 구현한 데 있다. 콜로서스를 운영하기 위해서는 최소 120 MW(메가와트)의 전력이 필요하고, 향후에는 훨씬 많은 전력을 필요로 할 것으로 보이는데, 콜로서스는 테네시 밸리 당국(TVA)과 멤피스 전력, 수도, 가스(MLGW)로부터 120 메가와트 이상의 전력을 공급받기 어려웠다. 그 임시 해결책으로 xAI는 데이터센터 한 칸에 트레일러들을 줄 세워 임시 전력을 확보했다. 향후 클러스터 옆 부지를 활용하여 자체 천연 가스 발전소를 건설하고 추가 전력을 확보할 예정이다.

대규모 클러스터에서는 전력 부하가 빈번하게 일어나는데, 전력 변동 때문에 출력도 함께 불안정해지는 문제가 있다. 이러한 문제들을 해결하기 위해 xAI는 테슬라의 메가팩 배터리를 이용해 전력을 안정적으로 공급할 수 있도록 시스템을 구축하였다.

도표 15. Scaling xAI: Colossus 인프라 비용 및 필요 전력량 추정



자료: Nextbigfuture, 유진투자증권

도표 16. 122 일 만에 GPU 10 만 개 → 92 일만에 20 만 개로 확장



자료: xAI, 유진투자증권

궁지에 몰린 OpenAI, 컴퓨팅 자원 확보가 절실

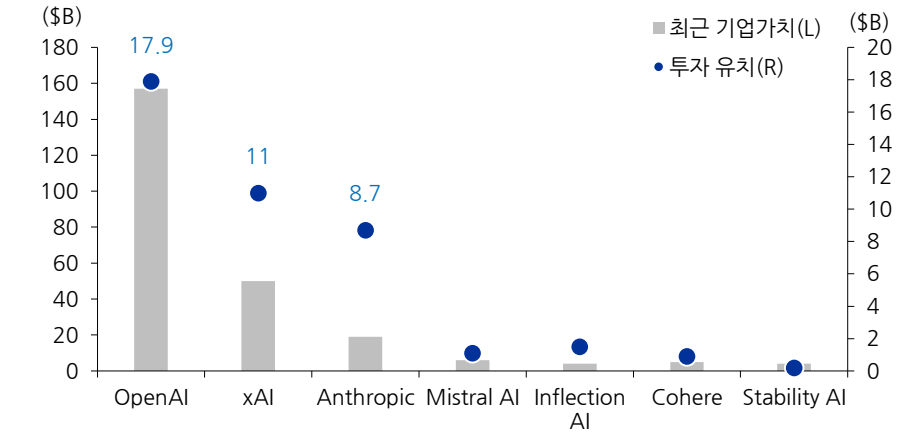
OpenAI는 초기 자금과 컴퓨팅 자원이 부족하여 MS와 협력을 해야만 했다. MS는 2019년에 10억 달러를 투자하며, Azure 크레딧을 제공했고, OpenAI는 이를 이용하여 AI 연구를 지속할 수 있었다. MS가 OpenAI를 위해 만 개의 GPU를 포함한 상위 5위 슈퍼컴퓨터를 구축했으며, 정확한 사용량이나 금액이 공개되지는 않았지만 데이터 센터 사용 비용이 포함된 MS의 총 투자금액은 130억 달러 수준으로 추정된다.

그러나, 최근 OpenAI는 CoreWeave와의 추가 계약(119억 달러, 5년 간)을 통해 컴퓨팅 자원을 다각화하고 있다. 이로 인해 MS와의 독점적 클라우드 제공 관계는 종료되었지만, 두 기업 간 파트너십은 여전히 유지되고 있다. 배경에는 OpenAI가 MS의 제공 속도에 불만을 제기했으며, 그 사이 CoreWeave가 경쟁력 있는 가격과 빠른 배포를 제안한 것으로 보인다. 참고로 CoreWeave는 2024년 말 기준 25만 개 이상의 엔비디아 GPU를 보유하고 있다. xAI가 단 1년 만에 놀라운 성능의 모델을 선보였으며, DeepSeek가 비용효율을 이루어 내며 OpenAI의 마음은 급해지고 있다. Stargate의 프로젝트가 구체적으로 어떻게 진행될지가 매우 중요한 시점이다.

스타게이트 프로젝트는 OpenAI, MS, 소프트뱅크, 오라클, 엔비디아 등이 참여하는 최대 5천억 규모의 AI 인프라 프로젝트로, OpenAI의 컴퓨팅 자원 확보와 AI 모델 개발에 큰 변화를 가져올 가능성이 높다. 스타게이트를 통해 OpenAI는 수십만개의 GPU(GB200)를 이용할 수 있게 된다. 텍사스 애빌린의 첫 데이터 센터에만 6.4만 개의 GB200 칩이 탑재될 예정이라는 보도가 있었다. 이를 통해 OpenAI는 더 큰 모델(GPT-5)를 훈련하거나, 멀티모달 AI를 동시에 처리할 수 있는 능력이 강화된다. 현재 GPT-4 훈련에 2.5만 개~3만 개의 GPU가 사용된 것으로 추정되기 때문에 훈련 속도와 모델 복잡성을 개선시킬 수 있을 것이다.

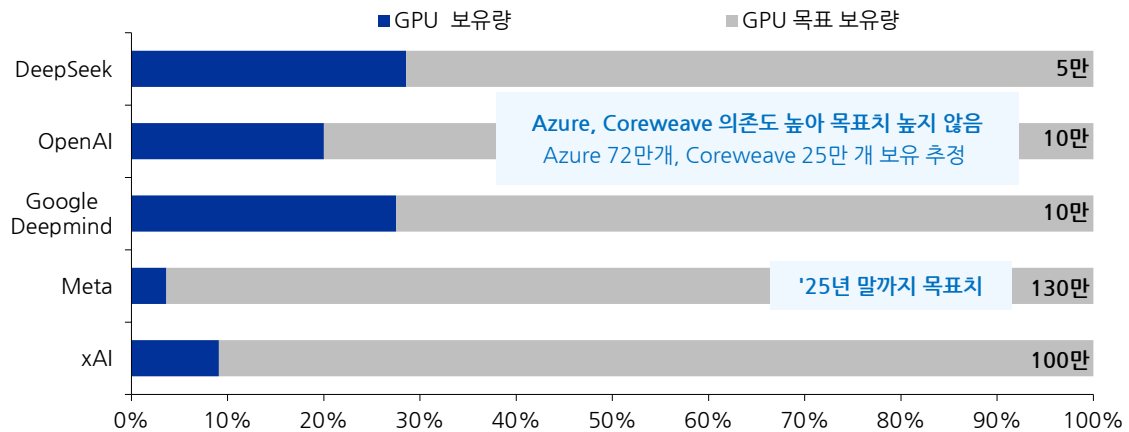
스타게이트 프로젝트는 궁극적으로 AGI 구현이라는 목표가 있다. AGI 수준에 도달하려면 극단적으로 GPT-4의 100배 수준의 파라미터가 필요하고, 5GW(기가와트)급 전력이 필요할 수 있는데, 스타게이트 프로젝트로 AGI 구현이 조금 더 가까워질 수 있다. 그러나 소프트뱅크 및 기타 초기 투자자들의 현금 여력이 있을지에 대해서는 의문인 상황이다.

도표 17. 최근 기업가치와 누적 투자 유치 규모



자료: 언론종합, 유진투자증권

도표 18. 기업별 현재 GPU 보유량과 목표치



자료: 각 사, 언론종합, 유진투자증권 추정치

도표 19. 기업별 현재 GPU 보유량과 목표치

	GPU 보유량(추정)	주요 GPU 종류	비고
OpenAI	10,000~50,000	A100, H100	차세대 클러스터 10만 개 GB200 목표
xAI	200,000+	H100	GPU 1M 개로 확장 예정
Deepmind	26,000~50,000	A100, H100, TPU	TPU v5 백만 개 이상 보유 추정
Meta	49,152	H100	25년 말까지 GPU 1.3M 보유 목표
Anthropic	50,000~100,000	A100, H100	AWS 의존도 높음
DeepSeek	20,000	H800, A100	수출 제한 영향 有

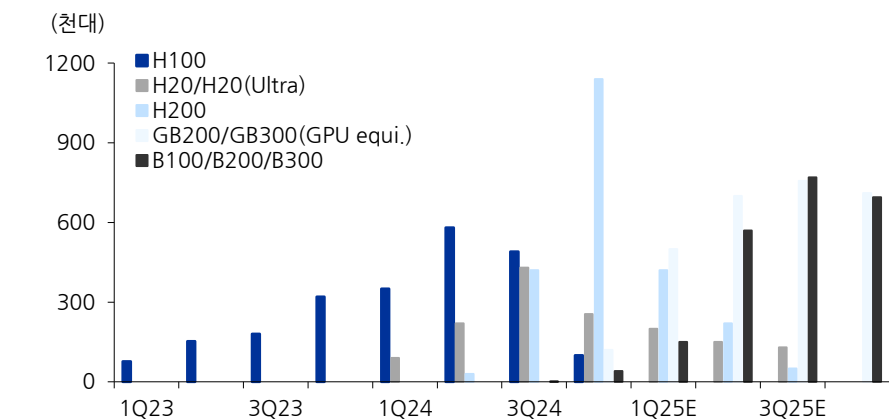
자료: 각 사, 언론종합, 유진투자증권 추정치

여전히 대규모 투자가 필요한 이유

AI 를 다양한 산업에서 사용하며 점차 수요가 늘어난다고 해도, 비용이 이렇게 급격히 감소한다면, 솔직히 더 이상의 대규모 투자는 불필요하다고 생각했다. 하지만 스타게이트 프로젝트를 시작으로 머스크의 대규모 클러스터 조성까지 투자가 증가하는 모습이 영 앞 뒤가 맞지 않는다.

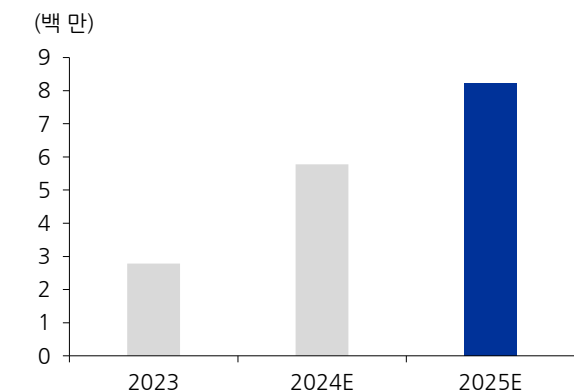
단순히 '범용화' 라고만 생각하기에는 천문학적인 금액이 투입되고 있다. 아무리 AI 사용처가 늘어난다고 해도, 비용이 낮아진다는데 투자는 왜 증가하는 걸까? 첫째는, 지금이 AGI 를 향해가는 변곡점에 위치해 있기 때문에 출혈 경쟁이 심해질 수밖에 없으며, 둘째는 AGI 에 도달하기 위해 파라미터수-모델 크기가 압도적으로 커지는 시점이기 때문이다.

도표 20. 엔비디아 GPU 모델별 출하량 추이 및 전망



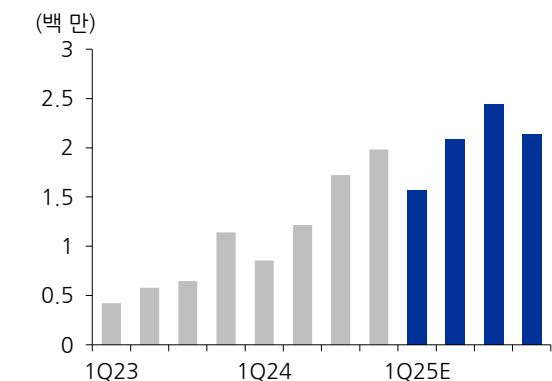
자료: 산업자료, 유진투자증권

도표 21. 엔비디아 연간 GPU 출하량 추이 및 전망



자료: 산업자료, 유진투자증권

도표 22. 엔비디아 분기별 GPU 출하량 추이 및 전망



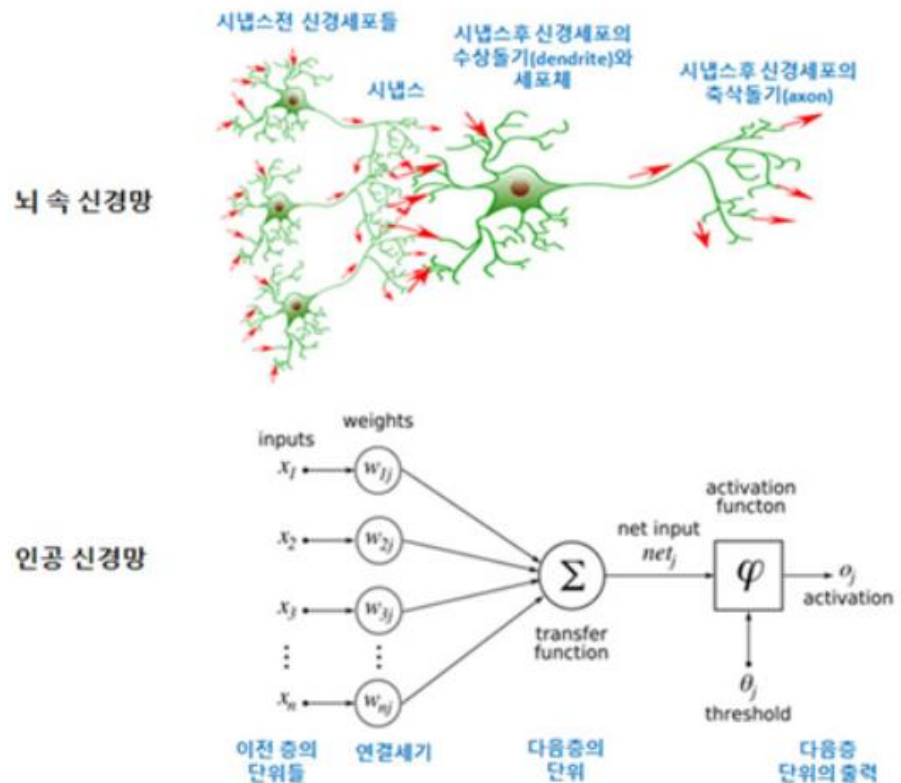
자료: 산업자료, 유진투자증권

구조 최적화와 파라미터 수 증가의 줄다리기

최적화된 구조 설계로 단위 비용이 줄어들고 있지만, 모델 크기와 파라미터 수가 급격하게 증가하며 비용이 상승하는 흐름이다. 이에 더해, 단순 텍스트 생성이 아닌 멀티모달 처리에 대한 필요성이 높아지며 모델 크기는 비용의 하락을 상쇄할 만큼 빠르게 커지고 있다.

결국 AGI 도달이 생성형 AI의 목표라고 본다면, AGI 도달에 필요한 파라미터가 얼마나 필요할지를 생각해 보면 된다. 생성형 AI 전문가들에 의하면 AGI 달성에 필요한 파라미터는 최소 1조 개 ~ 최대 100조 개의 파라미터가 필요하다. 사람의 뇌의 시냅스가 약 100조 개임을 참고하면 그럴싸한 수치다. 아키텍처의 혁신과 데이터 효율성이 AGI에 필요한 파라미터 수를 급격히 감소시킬 수는 있지만, 어쨌든, 최대 100T 수준의 파라미터가 필요하다는 사실을 생각하면 아직은 투자가 더 필요한 구간이라고 생각한다.

도표 23. 인간의 두뇌와 LLM 신경망 비교

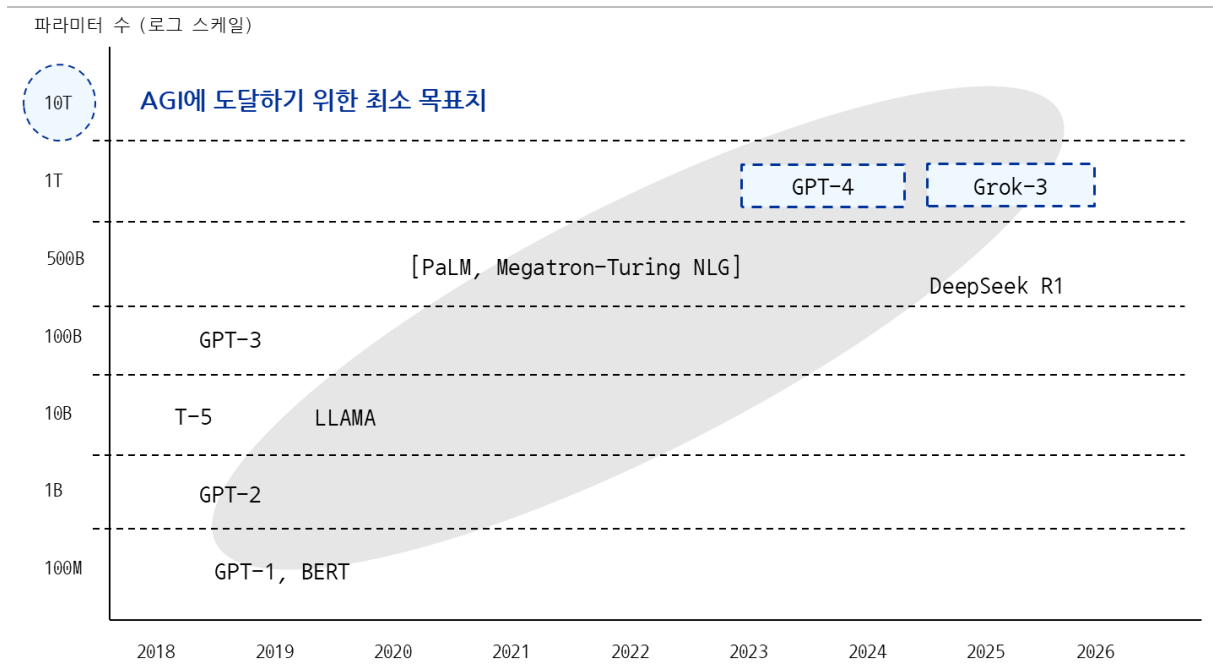


자료: SuperbAI, 유진투자증권 정리

AI에 대해 이해하기 위해서는 사람의 뇌를 비교 대상으로 생각해 보는 게 가장 쉽다. 사람의 뇌는 약 86억 개의 뉴런과 100조 개의 시냅스로 이루어져 있는데, 각 시냅스는 뉴런 간 연결 강도를 나타낸다. 이것을 LLM의 파라미터에 비유해 볼 수 있는데, AI가 뇌의 연결성을 모방하려면 100조 개 규모의 파라미터가 필요하다는 결론에 도달한다. 그러나 뇌는 sparse하고 효율적으로 작동하기 때문에 모든 시냅스가 동시에 활성화되지 않는다. 이를 고려하면 AGI는 10조 개 미만의 파라미터로도 달성 가능하긴 하다. (앞서 언급한 MoE나 MLA 기술을 참고하면 생성형 AI는 점차 뇌를 표방하는 방식으로 가고 있다)

현재 대규모 LLM의 파라미터 수는 이미 1T를 넘기 시작했다. 그러나 여전히 AGI 수준에는 한참 미치지 못한다. GPT-4나 Grok-3를 고려했을 때 가장 현실적인 목표는 10T~50T 수준이 될 것이다. 10T 수준의 파라미터를 가진 모델의 훈련 비용은 최소 100억 달러에서 1,000억 달러 수준이고, 100T 수준의 파라미터를 가진 모델 훈련에는 수천억 달러 이상의 자금이 필요하다. 따라서 10T~50T 수준의 파라미터를 목표로 한다면, 모델당 무난하게 수천억 달러의 자금이 필요하다. 스타게이트 프로젝트의 규모가 5,000억으로 정해진 데에는 이런 계산이 있었던 것 같다. 아키텍처 최적화로 비용이 급격하게 감소하고 있는 것은 맞지만, AGI를 목표로 할 때 아직 더 많은 투자가 필요한 것이 사실이다.

도표 24. 주요 LLM 모델의 파라미터 규모 증가 추이



자료: 각 사, 언론종합, 유진투자증권 정리

도표 25. 주요 LLM 모델의 파라미터 규모 증가 추이

모델명	기업	시기	파라미터	내용
GPT-1	OpenAI	2018	117M	최초의 GPT 모델, 트랜스포머 기반 소규모 파라미터
BERT	Google	2018	640M	양방향 학습 도입하며 당시 기준 큰 모델이었음
GPT-2	OpenAI	2019	1.5B	GPT-1 대비 10배 이상 증가, 텍스트 생성 능력 향상
T5	Google	2019	11B	text-to-text 프레임워크 도입, 파라미터 규모 확장
GPT-3	OpenAI	2020	175B	상업적 성공과 함께 파라미터 급격히 증가
Megatron	NVIDIA/MS	2021	530B	당시 최대 규모 모델, 대규모 훈련 인프라 필요
PaLM	Google	2022	540B	효율성과 성능을 동시에 추구한 대규모 모델
LLaMA	Meta AI	2023	65B	효율성 강조하며 상대적으로 작은 규모로 경쟁력 확보
Grok	xAI	2023	1T~10T	중대형 모델로 추정
GPT-4	Google	2023	1.76T	공식 수치는 없음. 멀티 모달 기능 추가
DeepSeek R1	DeepSeek	2025	671B	비용 효율성 강조하며 큰 규모 유지
Grok-3	xAI	2025	2.7T	공식 수치 없으나 초거대 Dense 모델일 가능성 높음

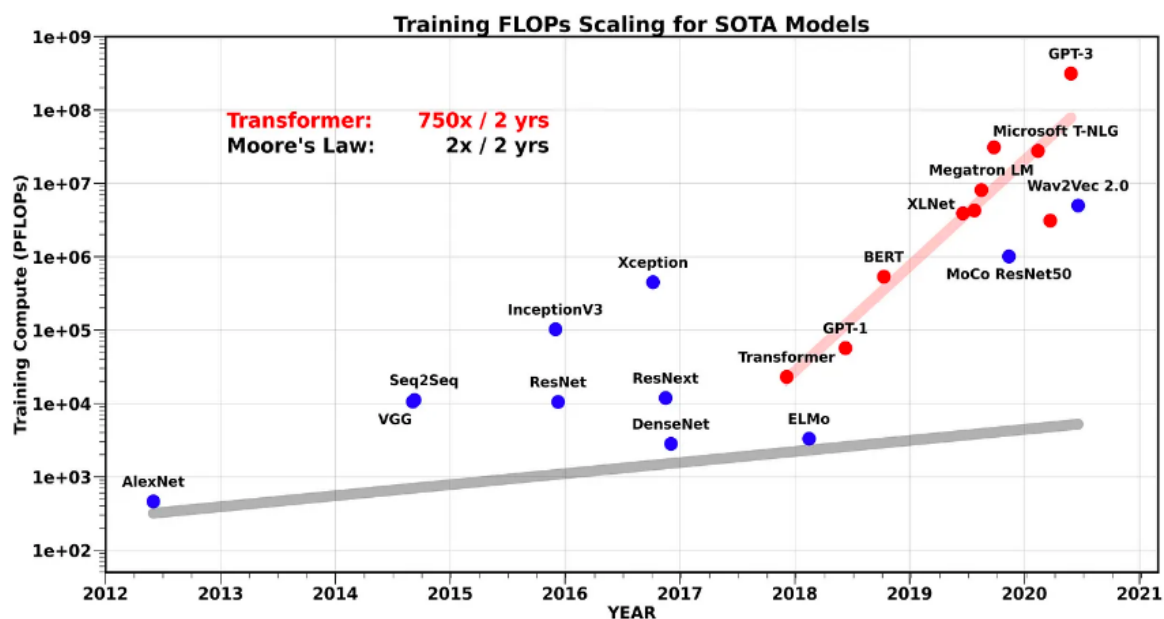
자료:각 사, 언론종합, 유진투자증권 정리

무어의 법칙을 능가한다, 압도적으로.

[도표 26]은 지난 10 년 간 딥러닝 모델 학습에 필요한 연산량이 어떻게 기하급수적으로 늘어났는지 보여준다. 특히 트랜스포머 모델 등장 이후 그 증가세가 무어의 법칙을 훨씬 상회한다는 부분이 눈에 띈다. X 축은 시간, Y 축은 모델 학습에 사용된 연산량으로, 그 단위를 PetaFLOPs(페타플롭스)로 표현했다. FLOPs(부동소수점 연산) 기준으로, 얼마나 많은 계산이 모델 학습에 투입되었는지 로그 스케일로 나타낸 그림이다. 쉽게 말하자면, 시기별로 주요 모델의 학습에 필요한 연산량이 어떻게 변화해 왔는지를 한눈에 보여준다. 또한 무어의 법칙과 트랜스포머 모델의 연산량 증가 추세를 비교하고 있다.

회색 실선으로 표시된 "Moore's Law: $2x / 2 \text{ yrs}$ "는 2 년에 2 배씩 증가하는 전통적인 무어의 법칙 추세를 나타낸다. 그러나, 트랜스포머 기반의 최신 모델들은 훨씬 가파른 증가율을 보이는데, 2 년 마다 2 배가 아니라 2 년마다 750 배 가까이 연산량이 증가하고 있다는 뜻이다. 이는 파라미터 수가 폭발적으로 증가함에 따라, 학습에 투입되는 컴퓨팅 자원도 기하급수적으로 증가한다는 것을 의미한다. 결론적으로 AGI 에 가까워지기 위해서는, 그리고 그 경쟁에서 뒤쳐지지 않기 위해서는 아직 투자가 필요한 상황이다.

도표 26. 기하급수적으로 증가하는 연산량 : 무어의 법칙과 비교



자료: IEEE Micro – AI and memory wall(2024.03) 유진투자증권

참고) 글로벌 주요 LLM 모델 비교

OpenAI GPT-4(o)

파라미터의 정확한 수치는 비공개이나 약 1.76 조 개로 추정된다. OpenAI 는 GPT-4 가 MoE 구조를 활용했는지 명확히 밝히지 않았지만, 일부 분석에서는 MoE 또는 하이브리드 구조일 가능성이 제기되고 있다. GPT-4o 는 기존 GPT-4 와 유사한 구조를 가지면서도 최적화된 모델로 보이며, 일부 파생 모델(GPT-4o mini 등)을 포함할 가능성이 있다. 학습 토큰의 경우 Common Crawl, 위키피디아, 서적, 코드 저장소 등에서 수조 개의 토큰을 학습했다.

학습 방법에 있어서는 지도 미세조정(SFT)과 강화학습(RLHF)을 적용했다. RLHF 는 AI 모델의 응답 품질을 인간 피드백을 통해 개선하는 방식으로, OpenAI는 이를 적극 활용하여 GPT-4o 의 응답 품질을 높였을 것으로 보인다.

Google Gemini 1.5

Gemini 역시 정확한 수치는 공개되지 않았으나, MoE 구조를 도입하여 활성화되는 파라미터는 수백억 개 수준으로 추정된다. 전체 파라미터 수는 수조 개에 이를 가능성이 있으며, MoE 의 특성상 한 번 추론 시 활성화되는 전문가(Experts) 네트워크만 연산에 참여하여 연산 효율성을 극대화한다. 이는 Google 이 기존 Switch Transformer 연구를 확장한 결과로 보이며, Gemini 1.5 Pro 는 성능과 효율성을 고려하여 설계되었다.

컨텍스트 윈도우는 기본 128K 토큰을 지원하며, 실험적으로 최대 100 만 토큰 까지 가능하다고 발표되었다. 이는 장문 처리와 다중 문서 분석에 강점을 가지며, Google 은 고성능 검색과 정밀한 문맥 이해를 목표로 설계한 것으로 보인다.

학습 방법은 MoE 아키텍처 활용했으며, SFT 및 RLHF 이 적용되었다. RLHF 를 통해 윤리적 판단과 안전성을 강화하는 방향으로 조정되었으며, 대량의 멀티모달 데이터 학습을 위해 MoE 를 적극적으로 활용하여 연산 부담을 최적화했다.

xAI Grok-3

파라미터 크기는 약 5,000 억 개~2.7 조 개로 추정된다. Grok-3 의 정확한 아키텍처는 공개되지 않았으나, 이전 모델인 Grok-1 이 MoE 기반이었기 때문에 Grok-3 역시 '대규모' MoE 모델일 가능성이 높다. 특히, MoE 구조를 기반으로 하면서도 대규모 컴퓨팅을 활용하여 성능을 극대화한 것으로 보인다. 컨텍스트 윈도우는 대 128K 토큰을 지원하며, 장문 문맥 유지 및 정확한 정보 검색을 위해 최적화된 것으로 보인다. 특히 실시간 검색 성능을 강화하는 방향으로 설계되었다.

학습 토큰의 경우 다른 모델들과 차이가 있다. X(트위터) 데이터를 포함하고 있으며, 학습방법은 RLHF 일부 적용했다. RLHF 를 통해 사용자 피드백을 반영하고, 코드 관련 학습에서는 RL 을 적용하여 실행 결과를 평가하는 방식으로 최적화되었을 가능성이 있다.

DeepSeek-V3

파라미터 크기는 총 6,710 억 개이며, 활성화되는 파라미터는 370 억 개 수준이다. 극단적인 스파스 모델로 설계되어 있으며, 필요할 때만 일부 파라미터를 활성화하는 방식으로 효율적인 연산이 가능하다. 기존 GPT-4 대비 연산량을 줄이면서도 성능을 유지하도록 설계되었고, 특히, MoE 아키텍처를 개선하여 효율성을 극대화했다. 학습 방법은 다른 모델과 마찬가지로 MoE 아키텍처 기반, SFT 및 RLHF 적용했다. 학습 과정에서 저품질 데이터를 필터링하고, 인간 피드백을 반영하여 정확성을 개선하는 방식으로 모델을 조정한다.

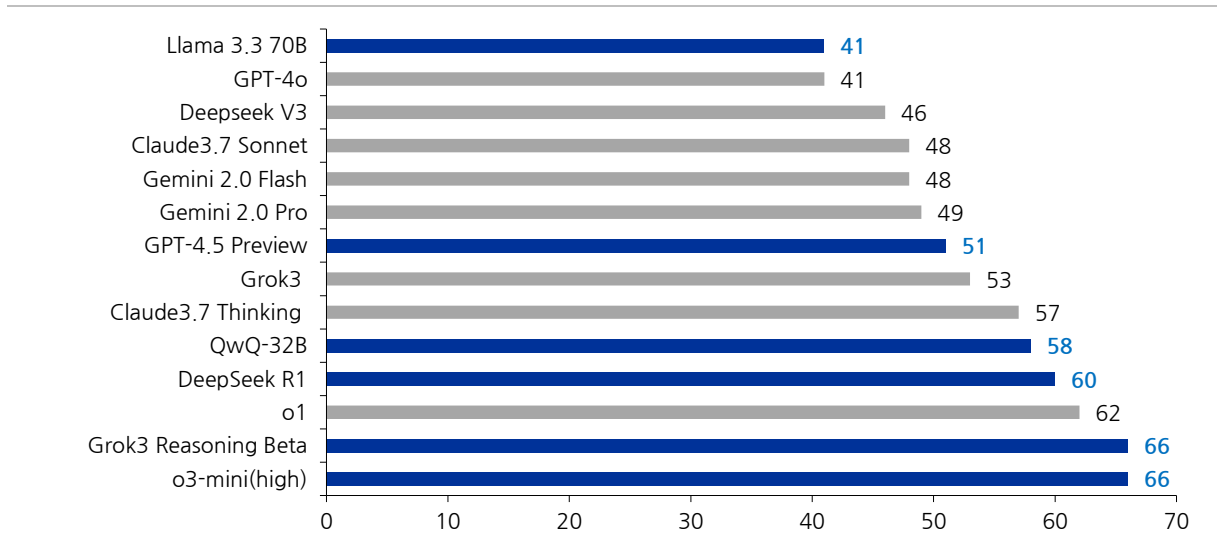
이들 모델은 Transformer 기반 아키텍처를 유지하면서도, MoE 기법을 적극 활용하여 성능과 효율성을 극대화하고 있다. MoE 구조를 채택한 모델들은 불필요한 연산을 최소화하여 대규모 파라미터를 유지하면서도 연산 비용을 절감하는 방식으로 최적화되고 있다. 컨텍스트 윈도우 또한 지속적으로 확장되면서 모델들이 장문의 정보를 더욱 효과적으로 처리할 수 있도록 개선되고 있다.

도표 27. 모델 별 비용 비교

	GPT-4o (OpenAI)	Gemini 1.5 (Google)	Grok-3 (xAI)	DeepSeek V3
파라미터 수	비공개 (1.76 조 개 추정)	비공개 (활성 파라미터 적음)	5 천 억~2.7 조 (Dense 가능성 有)	6,710 억 (활성 370 억)
컨텍스트 윈도우	128k 토큰 (일부 256k 지원 가능성)	128k (프리뷰 최대 1 백만 토큰 테스트)	128k (장기 기억 강화형)	128k (확장가능성有)
RLHF 여부	적용됨	적용됨 (AI 안전성 강화)	적용 여부 불명확 (실시간 피드백반영)	적용됨
기타 최신 기술 적용 여부 및 특징	MoE 가능성	MoE 기반	훈련비용 최적화 x	FP8 연산 적용 (훈련 비용 절감)
	Test-time Scaling 최적화	장문 컨텍스트 최적화 (100 만 토큰 실험)	실시간 검색 및 최신 정보 반영	오픈소스 공개, 파인튜닝 가능
	Multi-modal	Multi-modal	Think 모드 (사고과정 시각화)	저비용 고효율 학습 최적화 (MoE 활용)

자료: 각 사, 언론종합, 유진투자증권

도표 28. 모델 별 지능 지수 비교



자료: Artificial Analysis, 유진투자증권

주) 복합 지수, MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIIME, MATH-500 7 개 지표 점수 종합

[Part 2] 네카오의 AI 전략 분석

네이버

네이버 AI Phase1. AI 플랫폼 클로바(CLOVA)

2013년 네이버는 사내에 네이버랩스를 조직해 로보틱스, 자율주행, 인공지능 분야에 대한 연구를 시작했다. 또한 같은 해 국내 플랫폼 기업으로는 최초로 자체 데이터센터 '각(閣) 춘천'을 설립해 AI 기술 연구 기반을 다졌다. 이후 네이버랩스는 구글 딥마인드의 '알파고'가 산업계 전반에 AI 기술 기대감을 끌어올렸던 2016년부터 본격적인 AI 기술 개발에 들어갔다. 그리고 이듬해 선보인 것이 **네이버의 AI 플랫폼인 클로바(CLOVA)**이다. 당시 클로바는 음성 인식, 이미지 인식, 기계 번역 등 다양한 AI 기술을 통합한 플랫폼으로, 스마트 스피커 웨이브(WAVE)와 같은 디바이스에 적용됐다.

하지만 당시 네이버 클로바의 성능은 한계가 있다는 평가를 받았다. 클로바는 음성 인식, 번역, 챗봇 등의 기능을 제공했지만 자연어 이해(NLU)와 생성(NLG) 성능이 상대적으로 부족했으며, 한국어 특화 AI였지만 GPT-3 같은 대형 언어 모델과 비교하면 정교한 문맥 이해와 자연스러운 응답 생성에서 차이가 있었다.

네이버가 활용한 AI 모델의 매개변수(파라미터) 규모가 크지 않았고, 학습 데이터도 상대적으로 제한적이었다. 물론 네이버의 검색 데이터를 활용하긴 했지만, 지금과 같은 대규모 생성 AI 모델처럼 방대한 데이터 학습을 하지 못했기 때문에 답변 품질이 떨어졌다. 또한 텍스트 기반 AI 서비스도 상대적으로 단순한 기능(예: 뉴스 요약, 간단한 질의응답) 위주였기 때문에 활용성이 낮았다.

도표 29. 네이버랩스



자료: NAVER, 유진투자증권

도표 30. 네이버 클로바



자료: NAVER, 유진투자증권

도표 31. 데이터센터 '각(閣) 춘천'



자료: NAVER, 유진투자증권

도표 32. 네이버랩스 로봇



자료: NAVER, 유진투자증권

도표 33. 스마트 스피커 '웨이브(WAVE)'



자료: NAVER, 유진투자증권

도표 34. 클로바가 탑재된 스마트 디바이스들



자료: NAVER, 유진투자증권

네이버 AI Phase2. 첫 AI 모델 하이퍼클로바

2021 년 네이버는 현재의 자사 AI 언어 모델인 '하이퍼클로바 X'의 전신인 AI 모델 '하이퍼클로바(HyperCLOVA)'를 공개했다. 이 모델은 2,040억 개의 파라미터를 보유하고, 학습 데이터의 97%가 한국어로 구성되어 있어 한국어 처리에 특화됐다. 하지만 기술적으로 한국어 특화 대형 AI 모델이라는 점에서 의미가 있었음에도 불구하고 당시 시장에서 큰 주목을 받지 못했다. 이유로는 크게 경쟁모델 대비 차별성 부족, 국내에 국한된 시장 그리고 B2B·B2G 중심의 사업 전략을 꼽을 수 있다.

1) 경쟁모델 대비 차별성 부족

네이버는 하이퍼클로바를 한국어에 최적화된 초거대 AI 모델로 강조했지만, 이미 당시 경쟁 모델인 GPT-3가 한국어를 포함한 다국어를 지원했다. 또한 GPT-3는 2020 년부터 글로벌 시장에서 상용 서비스들이 빠르게 등장했는데, 하이퍼클로바는 네이버 자체 서비스(검색, 스마트스토어 등)에만 국한돼 사용되면서 외부 개발자들이 적극적으로 활용할 수 있는 환경이 부족했다는 지적을 받았다. 심지어 2022 년 11 월 ChatGPT 가 출시되면서 초거대 AI 시장의 중심이 OpenAI 로 완전히 넘어가게 된다.

2) 국내에 국한된 시장

GPT 시리즈는 글로벌 API 시장을 선점하고, 세계 여러 국가의 개발자들이 적극적으로 활용하면서 AI 생태계를 확장했다. 반면, 하이퍼클로바 X 는 네이버 자체 생태계 중심(한국 시장)에 머물러서 글로벌 AI 시장에서의 영향력이 거의 없었다. OpenAI 는 MS, 구글 등과 협업해 글로벌 시장을 공략했지만, 네이버는 주로 한국 기업과 협업하는 수준에 머물렀다.

3) B2B·B2G 중심 전략

OpenAI 는 GPT-3 를 API 형태로 공개하면서 스타트업과 기업들이 다양한 AI 서비스(예: Copy.ai, Jasper, GitHub Copilot 등)를 출시할 수 있도록 지원했지만, 네이버는 하이퍼클로바를 자사 서비스(검색, 광고, 쇼핑, 번역 등)에 적용하는 내부 중심 전략을 구사했다. 또한, 한국 정부 및 기업과 협업하는 B2B·B2G(정부용 AI) 모델에 집중하면서 일반 소비자가 체감할 기회가 적었다.

네이버 AI Phase3. 상용 서비스를 위한 하이퍼클로바 X

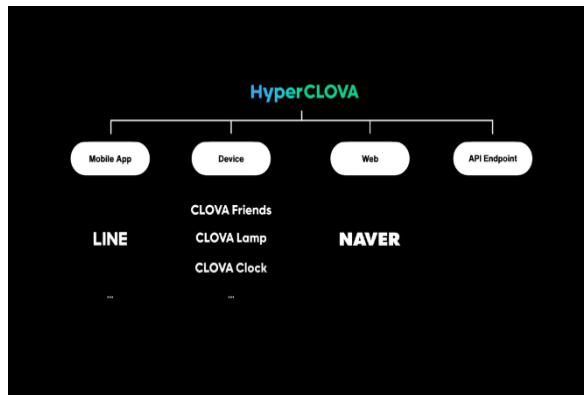
2023 년 네이버는 하이퍼클로바를 고도화한 하이퍼클로바 X(HyperCLOVA X)를 출시했다. 기본적으로 하이퍼클로바 X 는 전작인 하이퍼클로바 대비 성능이 크게 개선됐다. 또한 하이퍼클로바가 주로 기업 대상 B2B 서비스 및 네이버 내부 서비스용으로 활용된 기술 중심 모델이었다면, 하이퍼클로바 X 는 네이버 내부 서비스의 품질 향상뿐만 아니라 자체적인 상용화 서비스 출시도 가능케 했다.

하이퍼클로바 X에 기반한 대표적인 상용화 서비스로는 CLOVA X와 Cue:가 있다. CLOVA X는 네이버가 2023 년 8 월 하이퍼클로바 X 발표와 함께 선보인 대화형 생성형 AI 서비스다. Cue:는 검색 연동 AI 서비스로 네이버 검색 결과를 요약하고 정리해주는 기능을 제공한다. CLOVA X와 Cue:는 하이퍼클로바 X 기반의 서비스라는 측면에서는 유사하나, CLOVA X 는 독립적인 대화형 생성형 AI 로 ChatGPT 에, Cue:는 검색에 연동된 AI 서비스인 구글의 AI Overview 에 가깝다.

또한 하이퍼클로바 X는 기업 대상 API 제공 플랫폼인 CLOVA Studio 를 통해 다양한 비즈니스 서비스에 직접 활용될 수 있도록 설계되었으며, 모델 경량화 및 다국어 지원, 멀티모달 확장 가능성 등도 함께 고려되었다.

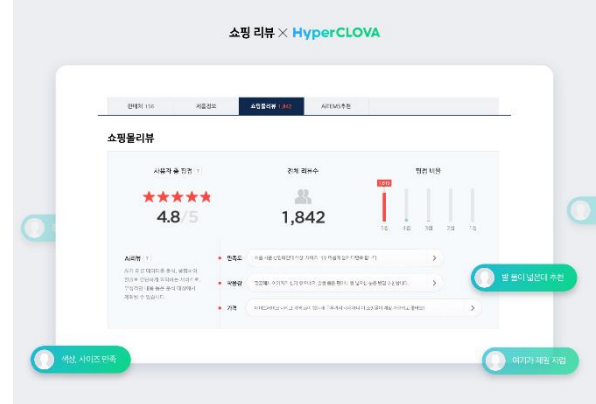
하지만 성능이 향상된 하이퍼클로바 X 기반의 상용 서비스도 시장의 반향을 일으키진 못했다. 앞서 언급한 하이퍼클로바의 한계점과 유사한데, 결국 글로벌 경쟁 모델과 비교해 성능이 뒤쳐졌으며 특히 다국어 지원이나 고난도 작업에서의 성능 차이가 두드러졌다. 차별화된 B2C 서비스도 부재했는데, 이미 시장에서는 여러 유사한 형태의 B2C 서비스들이 출시돼 있었기 때문이다. 이로 인해 국내 시장만을 타겟했음에도 불구하고 실사용자 입장에서는 사용 편의성, 기능, 응답 품질에서 큰 매력을 느끼지 못했던 것으로 보인다.

도표 35. 하이퍼클로바



자료: NAVER, 유진투자증권

도표 36. 쇼핑 리뷰에 적용된 하이퍼클로바



자료: NAVER, 유진투자증권

도표 37. 하이퍼클로바 X



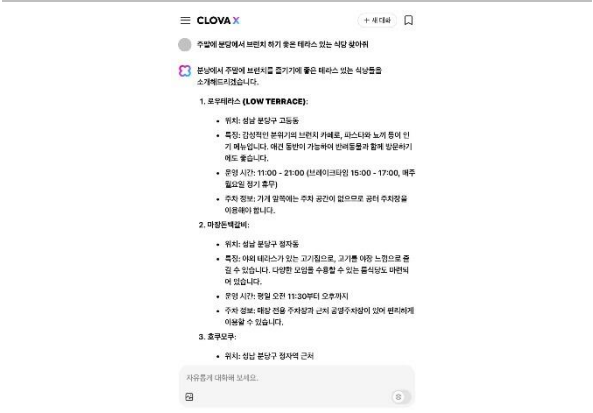
자료: NAVER, 유진투자증권

도표 38. 데이터센터 '각(閣) 세종'



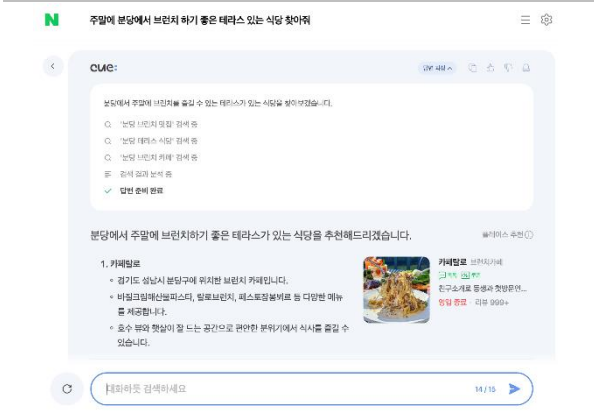
자료: NAVER, 유진투자증권

도표 39. 대화형 AI ‘CLOVA X’ 사용 화면



자료: 유진투자증권

도표 40. 검색형 AI ‘Cue.’ 사용 화면



자료: 유진투자증권

도표 41. 하이퍼클로바 X 기반 서비스 라인업

서비스	출시 시기	고객	내용
CLOVA X	2023 년 8 월	B2C	대화형 AI 서비스
Cue:	2023 년 9 월	B2C	네이버 검색 서비스에 생성형 AI를 결합
CLOVA for Writing	2023 년 10 월	B2C	AI 글쓰기 어시스턴트, 스마트에디터에 생성형 AI를 결합
Neurocloud for HyperCLOVA X	2023 년 10 월	B2B	AI 클라우드 인프라 서비스(보안 솔루션+전용 AI 플랫폼)
CLOVA Studio With HyperCLOVA X	2023 년 10 월	B2B	기업용 하이퍼스케일 AI 개발 도구
CLOVA for AD	2024 년 1 월	B2B	브랜드챗(AI 챗봇 형태, 브랜드 매니저 역할 수행) 서비스

자료: 유진투자증권

네이버 AI Phase4. 온서비스 AI(On-Service AI)

네이버의 최수연 대표는 연초 4 분기 실적 발표회에서 2025 년 네이버의 전략을 '온서비스 AI'라 요약했다. 온서비스 AI 란 AI 기술을 고도화하여 이를 네이버가 전개하는 모든 서비스(광고, 커머스, 콘텐츠 등)에 적용해 서비스 차별화에 나서겠다는 것이다. 네이버의 온서비스 AI 전략은 하이퍼클로바(2021 년)와 하이퍼클로바 X(2023 년)를 거쳐 확립된 것으로 파악된다. 하이퍼클로바는 출시 이후 경쟁 모델 대비 성능이 뒤처지며, 실용성이 떨어진다는 지적을 받았다. 이후 성능을 한 단계 올리고, 상용 서비스까지 이어진 하이퍼클로바 X 를 출시했지만 여전히 서비스 경쟁력이 떨어졌으며 자칫 비용 부담만 가중될 수 있었다. **네이버는 비용 부담은 줄이면서 AI 모델을 활용한 수익성 개선을 지향했으며, 실용성과 수익성을 모두 챙기기 위한 방안으로 온서비스 AI를 꺼내든 것이다.**

먼저 플랫폼 이용자는 온서비스 AI 를 통해 본인 맞춤형 광고 콘텐츠를 자동으로 추천받게 된다. 물론 온서비스 AI 가 도입되기 이전에도 사용자의 검색 이력이나 구매 기록을 바탕으로 광고를 띄우는 리타겟팅 광고(behavioral retargeting)는 존재했으나, 온서비스 AI는 단순 검색 기록을 넘어서 최근 이용자의 활동을 종합해 분석하고 이에 맞는 카테고리 상품을 추천하게 된다. 이 과정에서 광고를 콘텐츠처럼 자연스럽게 배치해 사용자의 거부감을 낮추고, 정보로 받아들이게 만든다. 또한 검색 서비스의 AI 브리핑 기능과 연계해 광고를 자연스럽게 노출시키기도 하는데, 이를 통해 광고 클릭률과 전환율을 높일 수 있다.

네이버는 지난해 10 월 '초개인화'를 내세운 네이버 플러스스토어를 선보인 바 있다. AI 를 통해 개인의 취향과 관심사를 분석해 상품, 혜택, 프로모션, 쇼핑 관련 콘텐츠 등을 추천하는 형태다. 또한 발견 탭에서는 개인화된 숏폼(쇼츠 형태) 콘텐츠를 통해 사용자가 새로운 상품을 탐색하고 바로 구매할 수 있도록 지원한다. 이 외에도 콘텐츠 플랫폼(블로그, 카페, 웹툰 등)에서는 음악 추천, 소음 제거 등 자동화 편집 지원 기능, 카페 글 추천 시스템 고도화 등의 AI 콘텐츠 보조 기능 등을 제공한다.

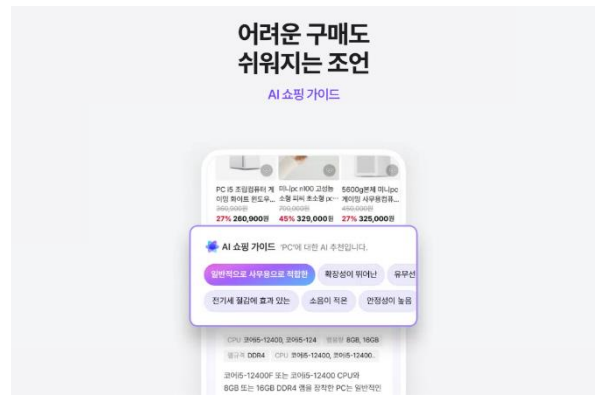
뛰어난 성능의 AI 모델을 보유한 전세계 주요 테크기업들도 수익화 측면에서 어려움을 겪고 있다. 네이버 또한 하이퍼클로바 X 를 개발하면서까지 상용화 서비스를 출시했으나 성능, 편의성, 범용성 등 여러 측면에서 아쉬운 모습을 보였다. **그보다 네이버라는 강력한 플랫폼을 보유하고 있다는 강점을 살린 온서비스 AI 전략이 현재 상황에서 타당한 전략이라 판단한다.**

도표 42. '단(DAN)24'에서 발표된 온서비스 AI



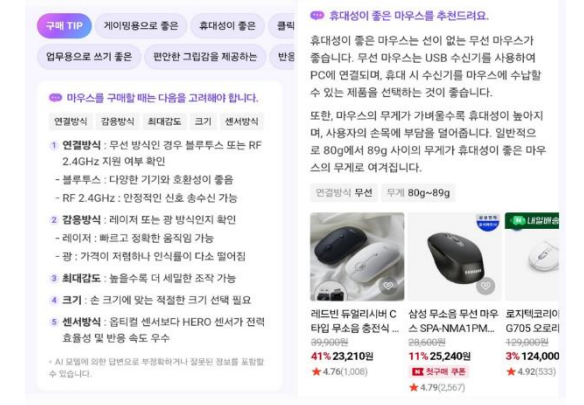
자료: NAVER, 유진투자증권

도표 43. 플러스스토어 'AI 쇼핑 가이드'



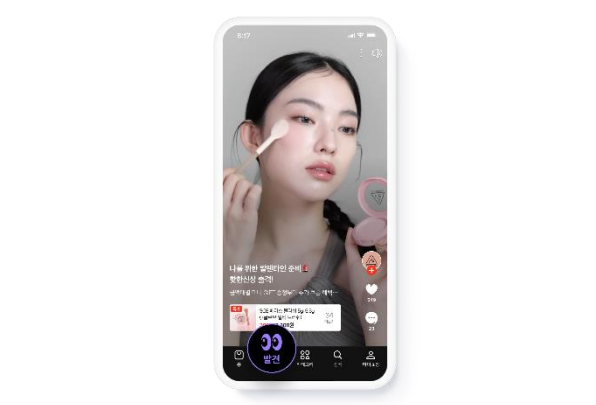
자료: NAVER, 유진투자증권

도표 44. 플러스스토어의 초개인화 추천 예시



자료: NAVER, 유진투자증권

도표 45. 플러스스토어 '발견' 탭



자료: NAVER, 유진투자증권

도표 46. 네이버검색 'AI 브리핑'



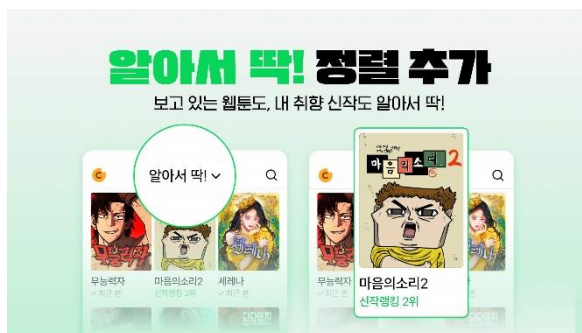
자료: NAVER, 유진투자증권

도표 47. 애드부스트 홈페이지 자동 타겟 광고



자료: NAVER, 유진투자증권

도표 48. 네이버웹툰 AI 추천 '알아서 딱!'



자료: NAVER, 유진투자증권

도표 49. 네이버웹툰 'AI 페인터'



자료: NAVER, 유진투자증권

카카오

카카오 AI Phase1. 헤이! 카카오

카카오는 2017 년 AI 개발 자회사 '카카오브레인'을 설립했다. 이어서 **자사 AI 플랫폼인 '카카오 i'**를 발표했는데, 이는 음성 인식, 이미지 분석, 자연어 처리 등 다양한 AI 기술을 기반으로 한 AI 플랫폼이다. 카카오 i 발표와 함께 '카카오 미니'라는 스마트 스피커를 출시했고, 여기에 음성비서 서비스인 '헤이 카카오' 서비스를 탑재했다. 하지만 당시의 카카오 i 플랫폼은 카카오 미니와 같은 스마트 스피커 중심의 서비스에 국한되었으며, 음성 인식, 자연어 처리 등에 있어서도 날씨 안내, 음악 재생, 일정 확인 등 주로 정해진 간단한 명령을 인식하고 수행하는 수준에 그쳐 상호작용에 있어 제약이 컸다.

카카오 AI Phase2. AI 모델 KoGPT1.0

2021 년 11 월, 카카오브레인은 한국어 특화 언어모델인 'KoGPT'를 발표했다. 이름에서도 알 수 있듯 오픈소스로 공개됐던 OpenAI의 GPT-2 아키텍처를 기반으로 만들었으며 한국어에 특화된 데이터셋을 학습했다. 다만 KoGPT 기반의 유의미한 상용 서비스는 부재했으며, 오픈소스를 활용해 국내 연구소, 대학 등에서 자연어 처리를 연구하는 수준에 그쳤다. 또한 KoGPT 가 발표된 시점에는 이미 GPT-3.0 이 공개된 상태였고, 이듬해에는 GPT-3.5 기반의 Chat GPT 가 시장에 등장했기 때문에 상대적으로 KoGPT 의 영향력은 미미한 수준이었다.

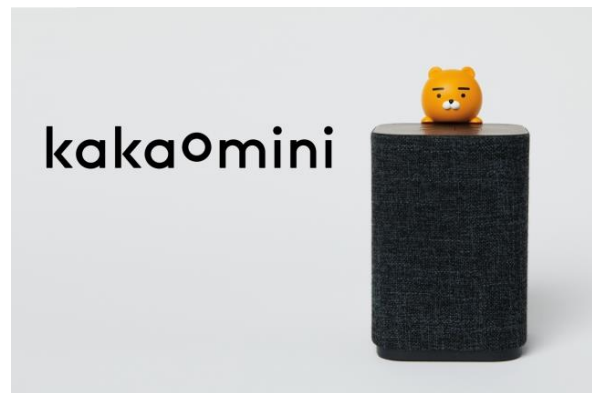
이후 카카오는 KoGPT 1.0 을 기반으로 성능과 비용 효율성을 균형적으로 갖춘 AI 파운데이션 모델인 KoGPT 2.0 의 개발에 몰두했다. 하지만 2023 년에 발표 예정이었던 KoGPT 2.0 은 계속해서 공개가 지연됐다. 이미 개발은 완료됐지만 이에 기반한 유의미한 서비스 출시가 부재한 상황에서 모델을 공개하는 것이 부담으로 작용했을 가능성도 적지 않다.

도표 50. 카카오 i



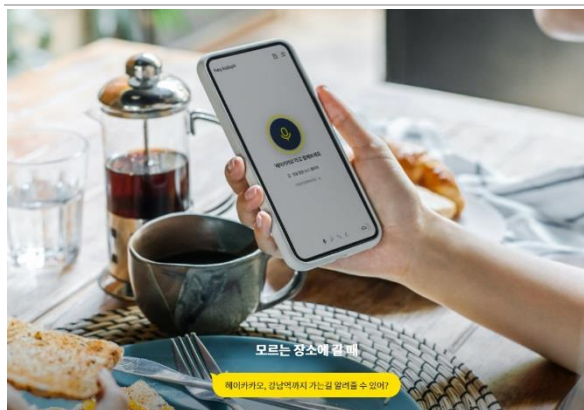
자료: 카카오, 유진투자증권

도표 51. 카카오 미니



자료: 카카오, 유진투자증권

도표 52. 헤이 카카오 서비스



자료: 카카오, 유진투자증권

도표 53. KoGPT



자료: 카카오, 유진투자증권

카카오 AI Phase3. (KoGPT2.0) Karlo 와 ddmm

이미지 생성 AI 서비스, Karlo

2022년 11월, 카카오브레인은 텍스트 기반 이미지 생성 모델인 'Karlo(칼로)'를 발표했다. Karlo는 기존 이미지 생성 AI 모델 Stable Diffusion의 오픈소스를 기반으로 개발됐으며, 이용자 얼굴 기반의 이미지 생성, AI 이미지 편집, 언어 모델 기반 프롬프트 자동 생성 등의 기능을 한국어에 맞춰 최적화한 이미지 생성 모델이다. 이후 2023년 5월에는 Karlo 2.0도 출시됐다. 약 3억 장 규모의 텍스트·이미지 데이터셋(Data Set)을 학습해 전작보다 성능, 해상도, 한국어 처리 능력 등 전반적인 기능이 크게 향상됐다.

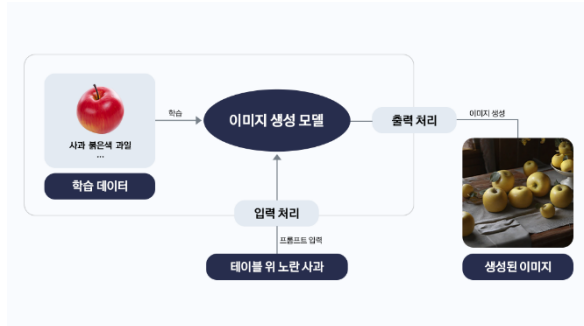
Karlo 2.0은 상업적 사용 허용, 고해상도 지원 등 강점을 갖췄지만 Midjourney, DALL-E와 비교해 퀄리티와 사용자 편의성 측면에서 경쟁력이 뒤떨어졌다. 또한 이미지 생성 AI는 GPU 등의 컴퓨팅 비용이 매우 크기 때문에 무료 또는 낮은 수익 모델로는 유지가 어려웠다. 이에 카카오는 2024년 7월부터 Karlo 서비스를 종료했다.

AI 기반 챗봇 서비스, ddmm(다다음)

2023년 3월 카카오브레인은 생성형 인공지능(AI) 기반 챗봇 서비스인 'ddmm(다다음)'의 오픈 베타 서비스를 시작했다. 카카오톡에서 친구 추가 방식으로 서비스를 이용할 수 있었는데, 서비스 공개 하루 만에 서비스 이용자가 1만명을 넘어서며 서비스가 중단됐다. 카카오는 '당초 내부 기대를 넘어서는 폭발적인 사용자 증가로 인해 20일 오후 4시를 기점으로 프로젝트 재정비에 들어가게 됐다'며 사용자 피드백을 바탕으로 서비스 재오픈 계획을 밝혔지만, 2년이 지난 현재까지도 ddmm 서비스는 재개되지 않고 있다.

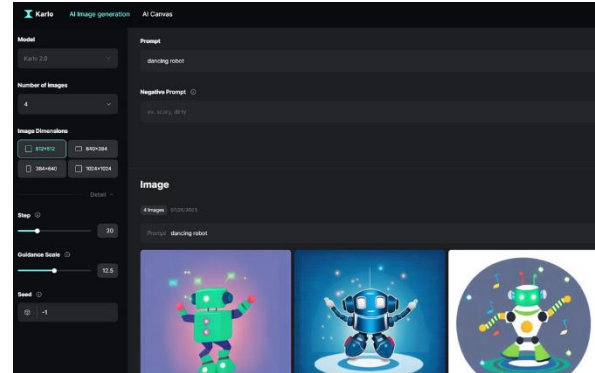
출시 초기 많은 이용자에게도 불구하고 ddmm 서비스가 재개되지 않는 이유를 몇 가지 추측해보면 첫째는 인프라 문제다. ddmm은 카카오톡 기반의 AI 챗봇으로 접근성이 뛰어났기 때문에 서비스 시작 하루 만에 서버 과부하로 다운됐다. 당시 카카오는 자체 데이터센터가 없었기 때문에 대규모 서비스 수요를 충족할 인프라가 부족했을 것이다. 두번째는 비용 문제다. ddmm의 경우 AI 챗봇의 기능에 이미지 생성 기능까지 보유하고 있어 클라우드 서버 비용이 크게 증가했을 것으로 추정된다(특히 외부 클라우드에 의존했다면 비용 부담은 더 커졌을 것). 이에 반해 당시 ddmm 서비스는 수익 구조가 부재했기 때문에 비용을 감당하기 힘들었을 것이다.

도표 54. 이미지 생성 모델 원리



자료: 카카오, 유진투자증권

도표 55. Karlo 웹 서비스



자료: 카카오, 유진투자증권

도표 56. Karlo 1.0 이 생성한 이미지 예시



자료: 카카오, 유진투자증권

도표 57. Karlo 2.0 이 생성한 이미지 예시



자료: 카카오, 유진투자증권

도표 58. 이미지 생성 AI 서비스 비교

	DALL-E	Stable Diffusion	Midjourney	Karlo
출시 시기	2021 년 1 월	2022 년 7 월	2022 년 8 월	2022 년 11 월
AI 모델	확산 모델	확산 모델	확산 모델	확산 모델
엑세스 방식	웹사이트	웹사이트 로컬 컴퓨터	웹사이트	웹사이트
소스	비공개	공개	비공개	공개
최고 해상도	1024x1024	1664x1664	1024x1024	1024x1024

자료: 유진투자증권

도표 59. ddmm(다다음)



자료: 카카오, 유진투자증권

도표 60. ddmm 서비스 예시



자료: 카카오, 유진투자증권

카카오 AI 개발 방향성은 어디로?

글로벌 테크 기업들의 AI 모델(GPT, Gemini, DeepSeek, Grok 등) 개발이 하루가 다르게 진전되는 가운데 상대적으로 국내 대표 소프트웨어 플랫폼 기업인 카카오의 AI 사업 성과는 아쉽다는 지적을 많이 받았다. 특히 KoGPT2.0의 개발이 거듭 지연되고 Karlo, ddmm 서비스도 중단된 와중 지난해 6월 AI 사업을 이끌었던 카카오브레인 김일두 대표도 회사를 떠나며, 향후 카카오의 AI 개발 방향성에 대해 물음표를 지울 수 없었다.

지난해 10월 열린 'ifkakaol 2024' 개발자 컨퍼런스에서 카카오의 정신아 대표는 자사의 새로운 AI 서비스인 '카나나(Kanana)'를 공개하며 카카오의 AI 비전과 전략을 발표했다. 카카오는 카나나를 그룹 단위 대화에서도 동작하는 관계 지향 서비스라고 소개했다. 기존 AI 서비스들이 일대일 대화를 통해 필요한 기능을 제공하는 방식이라면, 카나나는 그룹 대화에서도 맥락을 이해해 답변을 제시하는 등 관계의 형성과 강화를 돕는다는 점에서 차별화되며, 올해 상반기 중 별도 앱 출시도 예고했다. 다만 아직까지는 구체적인 서비스 정보가 부족하며, 별도앱 출시도 자칫 접근성이 떨어질 수 있어 시장에서도 이를 긍정적으로 평가하긴 무리가 있다는 반응이 주를 이뤘다.

도표 61. 카나나(Kanana)



자료: 카카오, 유진투자증권

도표 62. 카나나 모델 구조



자료: 카카오, 유진투자증권

카카오 AI Phase4, AI 오케스트레이션(with OpenAI)

이런 와중에 올해 2월 카카오는 오픈AI와 전략적 제휴 체결을 발표했다. 오픈AI와 제휴를 체결한 것은 국내 기업으로서는 최초이며, 양사는 AI 서비스를 쉽게 사용할 수 있도록 하는 데 초점을 맞춰 '**AI 서비스 고도화를 위한 기술 협력**'과 '**공동 상품 개발**'을 추진할 예정이라고 밝혔다.

기술 협력이나 공동 상품 개발과 관련해 구체적인 발표는 없었지만, 기본적으로 오픈AI의 최신 AI 기술을 국내 시장에 맞게 API 활용한 서비스가 출시될 것으로 보이며, 오픈AI의 ChatGPT, DALL-E 등의 기존 서비스와 유사한 형태일 것으로 추측된다. 물론 이 같은 서비스가 어떤 식으로 출시될 것인지(1) 별도앱 출시, 2) 카카오톡 내 친구추가, 3) 카카오톡 내 별도탭 신설)는 아직까지 미지수이다.

카카오는 그간 AI 개발에서 데이터센터 인프라 부족, 비용 부담 등의 문제를 해결하기 위해 AI 오케스트레이션 전략을 취했다. 오픈AI와의 협업으로 유의미한 성능을 지닌 AI 서비스가 카카오톡에 접목돼 출시된다면, 그간 카카오가 가지고 있던 카카오톡의 유의미한 유저 트래픽 향상이라는 숙제를 해결할 수 있다. 수익화 측면에서도 유료 구독 매출 혹은 트래픽 증가에 따른 광고 매출 성과가 드러날 것으로 기대된다.

도표 63. 카카오-오픈AI 전략적 제휴 체결



자료: 카카오, 유진투자증권

Game with AI

인간 챔피언을 이긴 최초의 AI, 딥블루

1997년 IBM이 만든 슈퍼컴퓨터 '딥 블루(Deep Blue)'가 당시 체스 챔피언인 그랜드마스터 가리 카스파로프에게 승리했다. 총 6번의 대국 끝에 3.5-2.5의 점수로 딥 블루가 승리한 것인데, 이로써 카스파로프는 세계 최초로 인공지능 컴퓨터에 패배한 챔피언이라는 불명예를 갖게 됐다.

하지만 당시에는 카스파로프를 이긴 딥 블루를 두고 지금과 같은 인공지능(AI)의 발전으로 여긴 사람은 많지 않았다. 그보다는 '슈퍼컴퓨터'의 계산 능력과 성능에 초점이 맞춰졌다. 특히 인간이 쉽게 따라올 수 없는 엄청난 연산 속도로 많은 수를 계산할 수 있다는 점이 강조된 것이다. 실제로 당시 언론 보도에서도 "AI가 인간을 이겼다"보다는 "IBM 슈퍼컴퓨터가 체스 챔피언을 이겼다"는 식으로 보도되는 경우가 많았다.

인공지능의 관점보다 슈퍼컴퓨터로 초점이 맞춰진 이유는 딥 블루가 오늘날의 AI처럼 딥러닝이나 신경망을 활용한 게 아니라, 철저히 '고성능 연산'을 바탕으로 한 '브루트 포스(Brute Force)' 방식에 기반했기 때문이다. 브루트 포스란 완전 탐색 알고리즘, 즉 가능한 모든 경우의 수를 모두 탐색하면서 요구조건에 충족되는 결과만을 가져오는 연산 방식이다. 따라서 당시 딥 블루는 AI라는 개념보다는 단순히 강력한 컴퓨팅 파워로 모든 경우의 수를 계산하는 기계로 여겨졌다.

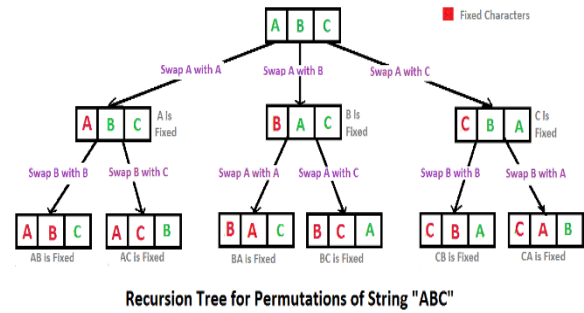
물론 딥 블루는 당시로서는 '슈퍼컴퓨터'였겠지만, 현재의 보급형 GPU들과 비교하면 성능이 지극히 떨어지는 수준이다. 오늘날의 일반적인 게이밍 PC에도 딥 블루보다 강력한 연산 성능을 가진 그래픽 카드가 들어가며, 심지어 스마트폰조차도 당시의 슈퍼컴퓨터의 연산 능력을 뛰어넘을 수 있는 수준에 이르렀다.

도표 64. 딥 블루 vs. 카스파로프



자료: 언론종합, 유진투자증권

도표 65. 브루트 포스 방식



자료: 유진투자증권

도표 66. 딥 블루 vs. 스마트폰 vs. GPU vs. AI GPU 비교

	딥 블루 (1997)	스마트폰 (iPhone 15 Pro)	GPU (RTX 4090)	AI GPU (NVIDIA H100)
연산 능력	약 11.38GFLOPS	약 10~15 TFLOPS (GPU)	약 82.6 TFLOPS (FP32)	약 700 TFLOPS (FP16, AI 연산)
메모리 용량	7.5GB RAM	8GB RAM	24GB GDDR6X	80GB HBM3
연산 구조	특수 목적 병렬연산	ARM 기반 SoC, GPU, NPU	대규모 병렬 GPU, CUDA 코어	텐서 코어 기반, AI 최적화 구조
전력 소비	약 20kW	약 5~10W	약 450W	약 700W
무게 / 크기	1.4 톤 / 대형	221g / 소형	2kg / 데스크탑용	약 1~4kg / 서버용
가격대	개발/구축 비용 약 수백만 달러	약 1300 달러	약 1600~2000 달러	약 3~4 만 달러
응용 범위	체스 AI 전용	통신, 게임, 가벼운 AI	게임, 영상, AI 추론	대규모 AI 학습/추론

자료: 유진투자증권

딥러닝의 알파고

2016년 3월 인공지능 기술은 다시 한 번 전세계를 놀라게 하는데, 구글 딥마인드의 인공지능 바둑 프로그램 '알파고(AlphaGo)'가 정상급 바둑 프로기사인 이세돌 9단과의 대결에서 4대1로 승리를 거둔 것이다. 당시 대국이 있기 전 대부분의 바둑 전문가들은 이세돌 9단의 승리를 점쳤다. 그 이유는 전년도에 치렀던 유럽 바둑 챔피언 판후이(2단)와의 대결에서 알파고가 승리를 거두긴 했으나, 대국만 놓고 봤을 때는 이세돌 9단의 실력에 한참 못 미친다는 평가가 지배적이었기 때문이다.

흥미로운 점은 당시 바둑 전문가들뿐만 아니라 AI 연구자들조차도 알파고보다는 이세돌 9단의 승리 가능성이 더 높다고 예상했다는 것이다. 이는 체스와 비교할 때 바둑에 압도적으로 많은 경우의 수가 존재하기 때문이다. 체스는 8x8(64 칸) 보드에서 진행되며, 말의 움직임도 제한적이기 때문에 이론상 체스 게임에서 나올 수 있는 경우의 수는 대략 10^{123} 가지이다. 반면 바둑의 경우 19x19(361 칸) 보드에서 진행되며 착점의 자유도도 높다. 이론상 가능한 경우의 수는 약 10^{700} 가지로 바둑의 경우의 수가 체스보다 압도적으로 많기 때문에, 기존 딥 블루처럼 경우의 수를 계산해 수읽기를 통한 승리 전략을 도출해 내는 방식(브루트 포스)은 바둑에서는 한계가 명확했다.

하지만 결과는 앞서 언급한 대로 4대1 알파고의 완승으로 끝났다. 오히려 현재 시점에서 보면 알파고를 이긴 이세돌 9단이 대단하다고 평가받는다. **알파고는 초일류 바둑 기사를 이기기 위해 브루트 포스 방식이 아닌 딥러닝 기반 지도학습을 활용했다.** 먼저 인간 프로 기사들의 기보(약 16만 판)를 학습해, 어떤 착수가 더 좋은지 기본적인 판단을 학습하고, 신경망을 사용해 패턴을 파악했다. 즉, 인간이 둔 기보를 보고 다음에 어디를 두면 승률이 높을 지를 예측하는 모델을 만든 것이다.

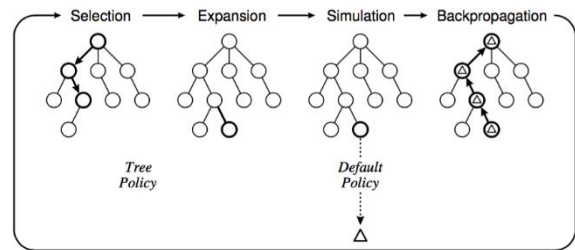
이후 알파고는 지도학습으로 익힌 기본 패턴을 바탕으로 수백만 판의 자기 대국(self-play)을 진행하며 스스로 더 나은 전략을 찾는 강화 학습에 들어갔다. 이 과정에서 AI는 이길 확률이 높은 수를 직접 찾아내고, 기존 인간 기보에서 보지 못한 새로운 전략도 발견하게 된다.

도표 67. 이세돌 vs. 알파고



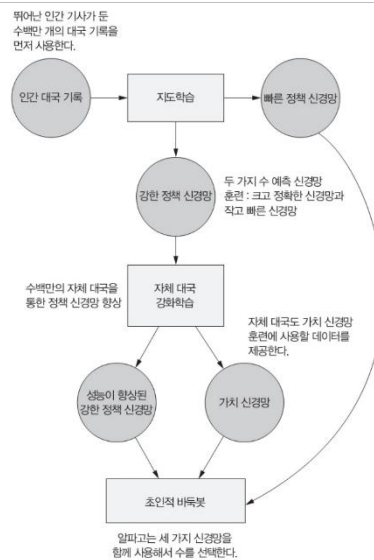
자료: 언론종합, 유진투자증권

도표 68. 바둑 AI의 몬테카를로 트리 방식



자료: 유진투자증권

도표 69. 알파고의 딥러닝 방식



자료: 딥러닝과 바둑(2020), 유진투자증권

도표 70. AI 바둑 프로그램 비교

	알파고 (2016)	알파고제로 (2017)	절예 (2017)	KataGo (2019)
개발사	구글 딥마인드	구글 딥마인드	중국 텐센트	오픈소스
훈련 방식	인간 기보+자가 대국	100% 자가 대국	인간 기보+자가 대국	자가 대국+실험적 강화
사용 자원	TPUv1 192 개	TPUv2 64 개	GPU 수십 개 추정	일반 GPU
대국 내역	이세돌 4:1 승, 초일류	알파고 압도, 초일류	커제 3:0 승, 초일류	프로 급 이상
AI 구조	CNN + MCTS	강화 CNN + MCTS	유사 구조로 추정	CNN + MCTS 개선
공개 여부	비공개	비공개	비공개	소스코드, 모델 일부공개

자료: 유진투자증권

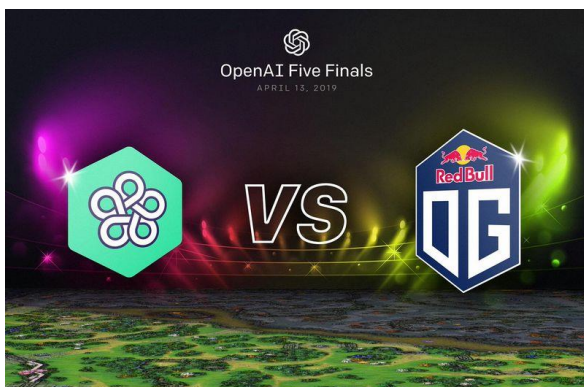
언어모델은 게임을 어떻게 바꿀까

2016년 알파고의 등장 이후 AI 연구가 게임 업계에서도 활발히 진행됐다. 2017년 오픈 AI 는 MOBA 장르 게임인 '도타 2'에서 'OpenAI Five'라는 AI 모델을 만들어 프로게이머들과의 경기에서 이긴 적이 있으며, 비슷하게 2018년 구글 딥마인드는 자체 개발한 AI모델 '알파스타(AlphaStar)'가 AOS장르 게임인 '스타크래프트 2'의 게임서버 배틀넷(Battle.net)에서 최고 레벨인 '그랜드 마스터'(게임 내 상위 0.2%)에 올랐다는 연구논문을 실었다.

하지만 알파고 이후 개발된 AI 모델들을 보면 게임 개발 측면에서 이점을 주는 AI 모델이라기 보다는 이미 개발된 게임(바둑, 도타2, 스타크래프트2 등) 내에서 사람과의 대결에서 승리하는 것에 초점이 맞춰졌다. 그러다 보니 AI 모델이 개발되는 장르도 보드게임, AOS, MOBA 등에 국한됐다.

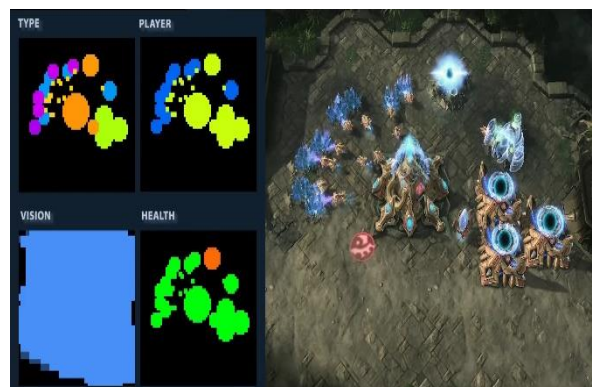
게임 개발에까지 AI가 영향을 미치기 시작한 시점은 2022년 ChatGPT의 등장 이후다. 이때부터 대형 언어 모델(LLM)이 게임 개발 과정에서 활용되기 시작했는데, 크게 보면 게임 개발 프로세스의 자동화 및 효율성 증가, 콘텐츠 생성 혁신, 그리고 플레이어 경험 향상이라는 세 가지 측면에서 변화를 가져왔다.

도표 71. 도타 2 OpenAI Five vs. OG



자료: OpenAI, 유진투자증권

도표 72. 알파스타의 스타크래프트 2 플레이 화면



자료: 구글 딥마인드, 유진투자증권

1) 게임 개발 프로세스의 자동화 및 효율성 증가

게임 개발 프로세스에서 LLM 은 특정 기능을 구현하는 스크립트를 자동으로 생성할 수 있다. 예를 들어, Unity에서 문이 플레이어가 가까이 가면 열리는 스크립트가 필요하다고 하면, LLM 은 이와 관련된 코드를 생성해 제공한다[도표 73]. 게임 내 스크립트 작성이 훨씬 간편해지면서 개발자들은 더 창의적인 부분에 집중할 수 있다. 또한 LLM 을 통해 버그가 발생하는 코드의 오류를 찾아 수정하는 것도 용이해지며, 성능 최적화를 위해 코드의 개선점을 제안하기도 한다.

LLM 은 NPC 의 행동 패턴을 만들거나, 이벤트 시스템을 구축하는 코드를 생성할 수도 있다. NPC 가 플레이어를 보면 대화창을 띄우고, 특정 선택지를 고르면 반응이 달라지는 시스템을 만들고 싶다면 LLM 이 필요한 코드를 제공해 줄 수 있다.

혹은 게임 기획 단계에서의 초기 스크립트를 디자인해주거나 아이디어를 제공해주기도 하며, 게임 스토리와 퀘스트의 스크립트를 직접 작성해주기도 한다. 예를 들어, "중세 판타지 RPG 에서 마법사 NPC 가 주는 퀘스트 스크립트를 만들어줘" 라고 요청하면 [도표 74]과 같이 대화 형식으로 생성하는 게 가능하다.

또한 스크립트의 자동 변환 및 최적화도 가능한데, 대표적으로 한 언어에서 다른 언어로 변환(예: Python → C#, Lua → C++)하는 것도 LLM 을 통해 손쉽게 가능하다[도표 75].

이처럼 LLM 을 활용해 코드 자동 완성, 버그 수정, 최적화 등의 작업이 더 빠르고 효율적으로 이루어지고 있다. 특히 반복적인 코드 작성이 필요한 부분(예: UI 스크립트, AI 행동 트리 생성 등)에서 많은 기여를 한다.

도표 73. 기능 스크립트 예시

```

csharp
using UnityEngine;

public class DoorTrigger : MonoBehaviour
{
    public Animator doorAnimator;
    private void OnTriggerEnter(Collider other)
    {
        if (other.CompareTag("Player"))
        {
            doorAnimator.SetTrigger("Open");
        }
    }
}

```

자료: 유진투자증권

도표 74. 퀘스트 스크립트 예시

```

rust
NPC (마법사) : "오, 여행자여! 나의 실험을 도와줄 수 있겠는가?"
플레이어 선택지:
1. "무슨 실험인가요?" -> [퀘스트 수락]
2. "죄송하지만 시간이 없습니다." -> [퀘스트 거절]
3. "보상은 뭔가요?" -> [보상 설명 후 선택 가능]

json
{
  "npc": "마법사",
  "dialogue": "오, 여행자여! 나의 실험을 도와줄 수 있겠는가?",
  "choices": [
    { "text": "무슨 실험인가요?", "result": "퀘스트 수락" },
    { "text": "죄송하지만 시간이 없습니다.", "result": "퀘스트 거절" },
    { "text": "보상은 뭔가요?", "result": "보상 설명 후 선택 가능" }
  ]
}

```

자료: 유진투자증권

도표 75. 코딩 언어 변환 예시

예제: Python에서 Lua로 변환

```

python
def move_character(x, y):
    print(f"캐릭터가 {x}, {y} 위치로 이동합니다.")

```

변환된 Lua 코드

```

lua
function move_character(x, y)
    print("캐릭터가 " .. x .. ", " .. y .. " 위치로 이동합니다.")
end

```

자료: 유진투자증권

2) 플레이어 경험 향상

AI 서비스를 통해 플레이어들에게 쾌적한 게임 환경을 제공할 수도 있다. 많은 게임들이 정식 출시 이후 게임 내에서 문제가 발생하거나 개발자들이 의도치 못한 방향으로 악용되는 경우가 많다. 대표적으로 게임 내 버그 악용, 작업장 운영, 욕설이나 비하 발언, 혹은 비인가 프로그램(헬퍼) 사용이나 승부조작(고의로 게임을 지게 만들어 팀원들에게 피해를 입히는 경우) 등이 있다.

위와 같은 문제들은 자칫 플레이어들로 하여금 게임의 재미를 크게 반감시켜 유저 이탈로까지 이어지게 하며, 실제로 피해를 입은 다수의 게임 사례가 존재한다. 이로 인해 많은 게임 운영사들은 이를 바로잡기 위해 전담 모니터링 인력들을 갖추는 등의 노력을 기울이고 있다. 하지만 게임 유저들이 늘어나고 비인가 프로그램이 고도화될수록 이를 모두 모니터링하고 제재를 가하는 데에는 한계가 있다.

AI 서비스는 이와 같은 문제 해결에 상당 부분 도움을 준다. LLM은 자연어 이해 능력이 뛰어나 지정된 욕설 필터링을 넘으려는 우회 표현까지도 그 문맥을 이해하고 의도를 파악해 잡아낼 수 있다. 비인가 프로그램 사용 혹은 승부조작의 경우에도 플레이어의 행동 데이터를 바탕으로 비정상 행동 시나리오를 생성형 AI가 학습해 예상할 수 있다.

도표 76. 작업장 모습



자료: 유진투자증권

도표 77. 롤 헬퍼프로그램 인게임 화면



자료: 유진투자증권

3) 콘텐츠 생성 혁신

LLM은 플레이어 선택에 따라 변화하는 스토리라인을 실시간으로 생성할 수 있다. 이는 기존의 고정된 스토리보다 더 몰입감 있는 게임 경험을 제공할 수 있는데, 게임 내 NPC의 기능 확대가 대표적인 예다.

NPC(Non-Player Character)란 비(非) 플레이어 캐릭터를 뜻하며, 게임 내에서 플레이어가 직접 조작할 수 없는 캐릭터를 의미한다. NPC가 처음 등장한 건 1975년 TRPG(Tabletop Roll Playing Game) '던전 앤 드래곤'에서다. 게임 마스터가 NPC를 조작하며 플레이어와 상호작용하게 했고, 이는 후대 디지털 게임에서도 그대로 적용되었다. NPC는 스토리를 전달하거나, 퀘스트를 제공하고, 적으로 등장하는 등 다양한 역할을 수행하며 RPG 게임에서 필수적인 요소로 자리잡았다.

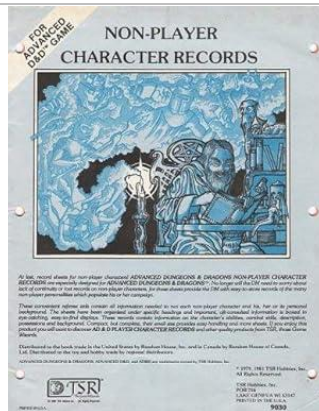
하지만 NPC와 플레이어 간의 상호 작용은 상당 부분 제한적이었다. NPC는 플레이어에게 기본적인 정보를 안내하거나, 아이템을 거래하고 스토리 전개를 돕는 등 주로 정적인 역할에 머물렀다. 게임 속 NPC는 간단한 역할만 수행하거나 미리 입력된 대사만을 반복한다. 이 같은 제한된 행동 패턴 혹은 반복되는 대사는 게임 플레이어의 게임 몰입도를 떨어트리게 된다.

NPC의 역할이 제한된 이유는 게임 개발자가 NPC를 설계할 때 트리 구조를 기반으로 플레이어의 말이나 행동에 따른 각각의 상호작용을 모두 개발해야 하기 때문이다. 물론 다소 복잡한 트리로 NPC를 개발할 경우 플레이어는 게임플레이의 자유도가 높다고 생각할 수 있지만, 이 또한 유한한 경우의 수가 존재할 뿐이며 그에 따른 비용과 개발시간이 늘어나게 된다.

최근 높은 퀄리티의 LLM이 연이어 등장하면서 AI NPC의 판도도 크게 달라질 전망이다. 물론 기존에도 NPC는 AI를 활용해 개발된 바 있지만, LLM 기반의 AI NPC는 기존과 궤를 달리한다. 기존 AI NPC는 미리 정해진 조건과 규칙에 따라 행동하는 NPC로 적 NPC가 플레이어를 보고 공격하거나 체력이 낮으면 도망가는 등의 FSM(유한 상태 머신, Finite State Machine) 기반으로 개발됐다면, LLM 기반 AI NPC는 자연어 처리(NLP)를 활용한 지능형 NPC라 할 수 있다.

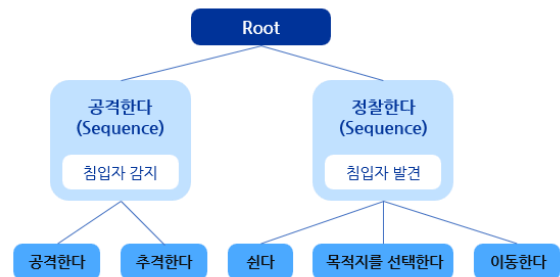
AI NPC는 기본적인 대화에 응답하는 것을 넘어, 플레이어의 행동에 실시간으로 반응해 생동감 있는 몸짓과 표정을 보이고, 플레이어와 협력해 미션을 수행한다. 결과적으로 플레이어는 NPC를 더 인간적으로 느낄 수 있어 몰입감을 높이고, 게임 설계 시 짜여진 각본에서 벗어나 자신만의 스토리를 만들어 나갈 수 있는 기회를 얻게 된다.

도표 78. AD&D(1978) NPC Records



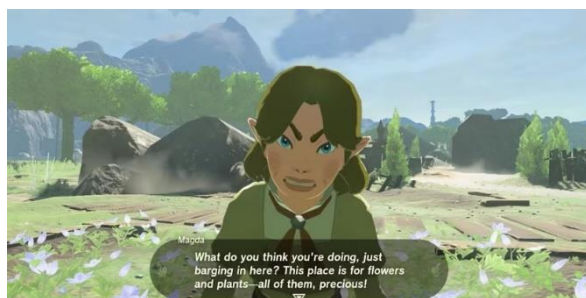
자료: 유진투자증권

도표 79. RPG 게임 NPC 의 행동트리 예시



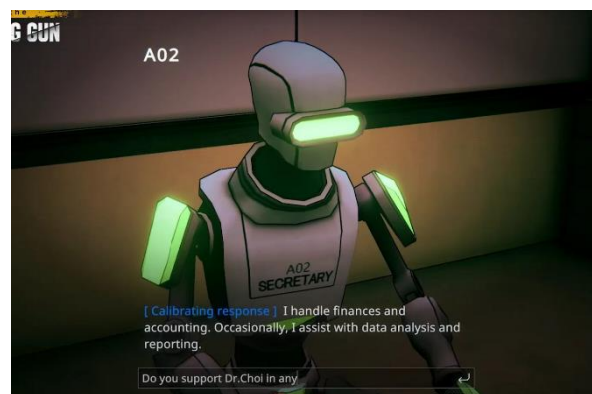
자료: 유진투자증권

도표 80. '젤다의 전설' NPC 오코바



자료: Nintendo, 유진투자증권

도표 81. '언커버 더 스모킹 건'의 대화형 AI NPC



자료: 크래프톤, 유진투자증권

AI NPC 의 한계

하지만 LLM 을 적용한 AI NPC 를 상용화함에 있어 큰 걸림돌이 생기는데, 바로 **비용**이다. LLM 은 수십~수백 기가바이트의 메모리와 고성능 GPU 를 필요로 하기 때문에, 대부분의 일반 게이머들의 PC 로는 로컬에서 LLM 을 구동하는 것이 어렵다. 특히 게임은 그래픽 처리에 리소스를 많이 쓰기 때문에, LLM 까지 돌리는 건 불가능에 가깝다.

만약 클라우드에서 LLM 을 호스팅할 경우, 사용량 기반으로 과금되는 구조이기 때문에 실시간 NPC 대화를 제공하게 되면 비용이 기하급수적으로 증가하게 된다. 예를 들어 수십만 명의 유저가 동시에 NPC 와 대화할 경우 LLM API 호출 수가 폭증하고, 특히 응답 속도(지연 시간)를 줄이기 위해 고성능 서버를 써야하기 때문에 단가가 더욱 올라간다. 결국 게임사의 수익 모델이 이 비용을 감당할 수 있어야 하는데, 현실적으로 어려움이 더 크다.

이에 대한 대안으로 1) 경량화 모델(SLM) 활용, 2) 프리셋+LLM 하이브리드, 3) 클라이언트-서버 협력 처리 등을 꼽을 수 있다. 경량화 모델이란 모델을 경량화하여 특정 대화 패턴이나 상황에 한정된 모델을 로컬에서 돌리는 방식을 의미한다. 게임 내 범용성이나 퀄리티는 LLM 과 비교해 떨어질 수 있지만 비용을 낮출 수 있다. 프리셋+LLM 하이브리드 방식이란 일반 대화는 기존 프리셋 대사로 처리하고, 특정 이벤트에 한해 LLM 을 활용하는 것을 말한다. 이를 통해 콘텐츠는 제한적이지만 퀄리티를 유지하고 트래픽과 비용을 통제하는 게 가능하다. 클라이언트-서버 협력이란 유저 클라이언트(PC/콘솔)와 개발사 서버가 역할을 나눠 LLM 관련 작업을 분담 처리하는 방식이다. 서버는 LLM 생성이라는 고비용 작업에 집중하고, 클라이언트는 상대적으로 가벼운 작업만 처리하는 것이다. 이를 통해 서버 부하와 클라우드 비용을 줄일 수 있으나, 이 또한 트래픽이 크게 늘어날 경우 감당하기 어렵다는 문제점은 지니고 있다.

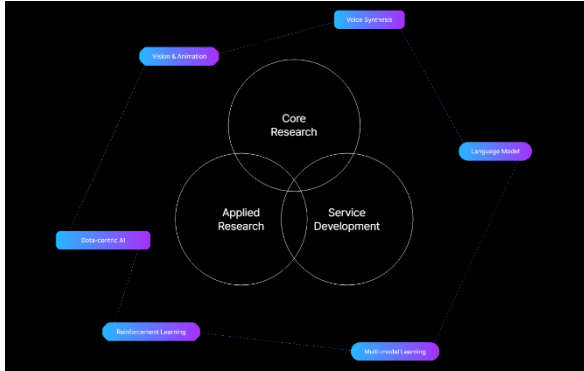
Game with AI에 진심인 크래프톤

크래프톤은 국내 게임 기업 중에서 AI기술 개발에 가장 적극적이다. 2022년부터 AI 전담 조직인 '크래프톤 AI'를 설립하고 이듬해부터 AI 사업 강화를 본격화하며 자사 게임 개발에 AI 기술을 적극적으로 활용하고 있다. 특히 생성형 AI를 중심으로 한 멀티모달 AI, 게임 개발 자동화, AI NPC 등에 집중하고 있는데, 올해 초 열린 CES2025에서 크래프톤은 게임 이용자와 상호작용할 수 있는 AI 협력 캐릭터 'CPC(Co-Playable Character)'를 발표했다. CPC는 엔비디아 에이스(ACE) 기술로 구축된 게임 특화 온디바이스 소형언어모델을 기반으로 이용자와 대화할 수 있으며, 마치 사람처럼 주변 상황을 인식해 대응할 수 있는 것이 특징이다.

3월 28일 얼리엑스로 출시 예정인 신작 인생 시뮬레이션 게임 '인조이(INZOI)'에 CPC가 적용될 예정인데, AI CPC로 헨리와 프시캣이 있다. 헨리는 전통적인 NPC의 역할을 하지만 이용자 반응에 따라 수많은 행동 선택지를 제공할 수 있다. 이는 헨리의 행동 트리가 정해져있는 것이 아니라 기본적인 방향성만 유지된 채 AI가 직접 행동 트리를 만들기 때문이다. 프시캣은 SLM을 기반으로 만들어진 챗봇으로, 기본적으로 게임을 안내해주는 역할을 하는 CPC이다. 게임 플레이 상 이해가 어려운 부분에 대해 자유롭게 질문하면 프시캣이 이를 인식하고 적절한 공략법을 제시해준다.

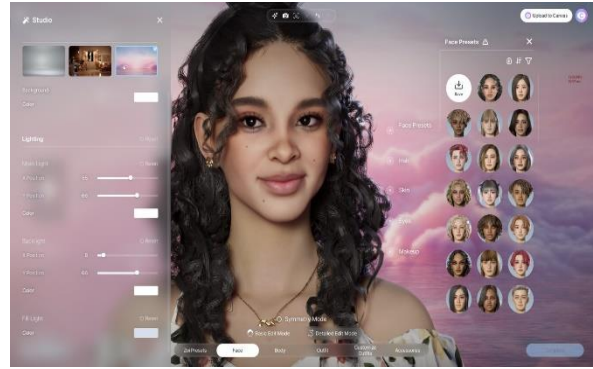
배틀그라운드에서는 자율적으로 학습하고 행동하며 플레이어를 돕는 '동료 CPC'를 도입할 예정이다. 플레이어가 음성으로 탄약과 차량을 요청하면 온디바이스 소형언어모델(SLM)로 구축된 CPC가 이를 수행한다. 요청에 반응하고 정해진 역할을 수행하는 것은 단순한 '봇' 캐릭터들도 수행할 수 있는 일이지만, CPC는 이보다 나아가 상호작용을 구현했다. 실제 사람과 마찬가지로 소통이 가능하고 게임에 대한 이해도가 높다. 적을 발견하면 경고하고, 상황을 판단해 공격 대신 은신을 선택하는 등 플레이어와의 경험을 학습해 나간다.

도표 82. '크래프톤 AI' 조직 구성



자료: 크래프톤, 유진투자증권

도표 83. 인조이 AI 캐릭터 스튜디오



자료: 엔비디아, 유진투자증권

도표 84. 인조이 AI CPC 'Smart Zoi'



자료: 엔비디아, 유진투자증권

도표 85. 배틀그라운드 AI 동료 CPC 'PUBG Ally'



자료: 엔비디아, 유진투자증권

도표 86. 엔비디아 에이스 AI 캐릭터



자료: 엔비디아, 유진투자증권

도표 87. 엔비디아 에이스 AI 캐릭터



자료: 엔비디아, 유진투자증권

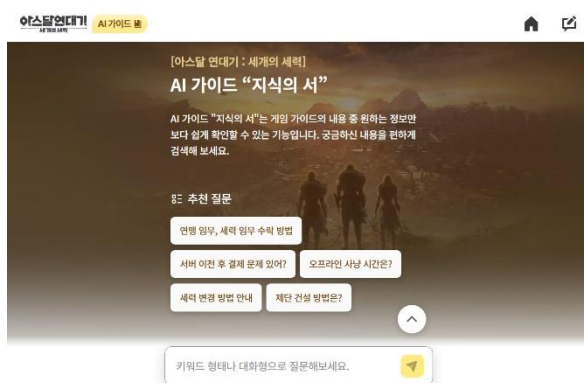
그 외 국내 게임사의 AI 적용 사례

엔씨소프트는 국내 게임사 중에서는 최초로 2011 년 AI 연구 조직을 설립하고, 지난해 11 월에는 해당 조직을 자회사 '엔씨 AI'로 분사하며 AI 전문성을 강화했다. 2023 년 8 월에는 국내 게임사 최초로 자체 거대언어모델(LLM) '바르코 LLM'을 공개했으며, 2024년 9월에는 한층 발전한 '라마 바르코 LLM'을 선보였다. 현재 주로 MMORPG 개발 과정에서 NPC의 텍스트 및 음성 생성, 게임 원화 제작 등에 AI를 활용 중이다.

넷마블 역시 2018 년 AI 연구개발 전담조직을 설립하여 AI 기술 개발에 집중하고 있다. 지난해 4 월 출시한 MMORPG '아스달 연대기: 세 개의 세력'에는 AI 가이드 시스템인 '지식의 서'를 도입하여, 이용자가 게임 플레이 관련 내용을 질문하면 공식 가이드 내용을 기반으로 생성형 AI를 통해 답변을 제공받을 수 있도록 했다.

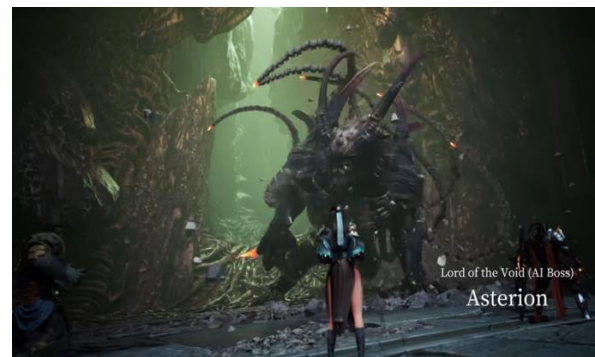
한편 위메이드는 엔비디아와 협력하여 출시 예정 신작에 엔비디아 에이스(ACE)를 활용한 AI 캐릭터, 인게임 전투 장비를 도입할 예정이다. 현재 위메이드넥스트는 엔비디아와 함께 MMORPG 신작 '미르 5'의 AI 보스 몬스터 '아스테리온'을 공동 개발 중이며, 디스이즈워는 FPS 신작 '블랙 벌처스: 프레이 오브 그리드'에서 AI 전투 분석 장비 '바이퍼'를 선보이는 등, 보다 정교한 AI 기반 전투 경험을 제공할 계획이다.

도표 88. 아스달 연대기 AI 가이드 '지식의 서'



자료: 넷마블, 유진투자증권

도표 89. 위메이드 AI 보스 '아스테리온'



자료: 엔비디아, 유진투자증권

기업분석

[NAVER\(035420\)](#)

BUY(유지) / TP 280,000 원(유지)

25년 키워드: 커머스 & 온서비스 AI

[카카오\(035720\)](#)

BUY(유지) / TP 48,000 원(유지)

AI 서비스로 쏠리는 시선

[크래프톤\(259960\)](#)

BUY(유지) / TP 470,000 원(상향)

트래픽이 어느 정도 높이까지 올라가는 거예요?

NAVER

(035420)

25년 키워드: 커머스 & 온서비스AI

투자의견

BUY(유지)

목표주가

280,000 원(유지)

현재주가

209,000 원(03/18)

시가총액

33,113(십억원)

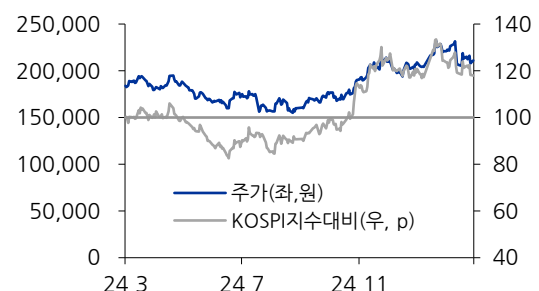
인터넷/게임/우주 정의훈_02)368-6170_uihoon0607@eugenefn.com

- 1분기 매출액 2.82조원(+12%yoy), 영업이익 5,222억원(+19%yoy)으로 컨센서스(5,232억원) 부합 전망
- 서치플랫폼(+12%yoy)은 확장검색 기능과 외부매체 광고 효과, 플레이스 광고 입찰 개선 효과가 1분기까지 이어질 것. DA매출은 피드 광고 지면 확장, 타겟팅 고도화를 통해 높아진 광고 효율 효과로 성장률 +11%yoy를 전망
- 커머스(+13%yoy)는 온플랫폼 거래액 성장이 매출 성장을 견인할 것. 3/12 별도앱 출시된 플러스스토어는 안드로이드 출시만으로 주간 앱 이용자수(WAU) 24만명을 기록, iOS 출시 이후 이용자 수 추가 확대 기대됨. 별도앱 출시와 병행된 수수료 개편의 매출 효과는 6월부터 반영될 예정, 커머스 매출 성장률 역시 상저하고 형태 예상함
- 콘텐츠(+6%yoy)는 전분기 국내 불매운동으로 인한 부정적 영향이 사그라들며 매출 성장률 소폭 회복 기대
- 비용에서는 플러스스토어 별도앱 출시에 따른 마케팅 비용의 유의미한 증가가 예상됨
- 지난해 네이버의 실적은 네이버앱 개편을 통한 트래픽 증가로 광고 매출 성장이 견인. 올해는 플러스스토어 별도앱 출시에 따른 커머스 매출 반등이 실적을 견인할 것. 플러스스토어앱 출시 효과는 커머스 내 중개, 판매 부문뿐만 아니라 커머스 광고 매출에도 긍정적으로 영향을 미쳐 실질적으로 광고와 커머스 매출 전반의 양호한 성장이 기대됨
- 2025년 네이버의 전략은 '온서비스 AI'로 요약됨, 네이버가 전개하는 모든 서비스(광고, 커머스, 콘텐츠 등)에 AI를 적용해 서비스 차별화에 나서겠다는 것. 비용 부담이 높은 별도의 상용 서비스 출시보다는 네이버라는 강력한 플랫폼 강점을 살린 온서비스 AI 전략이 현재 상황에서 타당한 전략이라고 판단

주가(원,03/18)	209,000		
시가총액(십억원)	33,113		
발행주식수	158,437천주		
52주 최고가	235,500원		
최저가	151,100원		
52주 일간 Beta	0.28		
60일 일평균거래대금	1,731억원		
외국인 지분율	48.5%		
배당수익률(2024F)	0.6%		
주주구성			
국민연금공단 (외 1인)	8.4%		
BlackRock Fund Advisors (외 14인)	6.1%		
주가상승	1M	6M	12M
상대기준	-5.5	29.1	16.3
절대기준	-6.1	30.5	13.5
(원, 십억원)	현재	직전	변동
투자의견	BUY	BUY	-
목표주가	280,000	280,000	-
영업이익(24)	1,919	1,979	-
영업이익(25)	2,304	2,277	▲

12월 결산(십억원)	2022A	2023A	2024P	2025E
매출액	8,220	9,671	10,738	11,915
영업이익	1,305	1,489	1,979	2,304
세전손익	1,084	1,481	2,246	2,569
당기순이익	673	985	1,862	1,863
EPS(원)	4,634	6,233	11,853	11,813
증감률(%)	-95.4	34.5	90.2	-0.3
PER(배)	38.3	35.9	17.6	17.7
ROE(%)	3.3	4.4	7.5	6.7
PBR(배)	1.3	1.6	1.2	1.1
EV/EBITDA(배)	16.1	17.8	11.2	7.3

자료: 유진투자증권



1Q25 Preview: 플러스스토어 별도앱 출시 효과 전임에도

1 분기 매출액 2.82 조원(+12%yoy), 영업이익 5,222 억원(+19%yoy)으로 컨센서스(5,232 억원)에 부합할 전망이다.

서치플랫폼(+12%yoy) 매출은 지난해 10 월 도입된 확장검색 기능 및 외부매체 광고 효과, 플레이스 광고 입찰 개선 효과가 1 분기까지 이어질 것이다. 특히 전년도 1 분기 부진했던 DA 매출은 피드 광고 지면 확장과 타겟팅 고도화를 통해 높아진 광고 효율로 인해 +11%yoy 성장률을 기록할 전망이다.

커머스(+13%yoy)는 시장 성장률을 상회하는 온플랫폼 거래액 성장이 매출 성장을 견인할 것이다. 3월 12일 별도앱으로 출시된 플러스스토어는 안드로이드만 다운로드 가능한 상황에서 주간 앱 이용자수(WAU)가 24 만명에 이른 것으로 조사됐다. 향후 iOS 출시까지 이어진다면 이용자 수는 더욱 늘어날 전망이다. 다만, 별도앱 출시와 함께 추진되는 수수료 개편으로 인한 매출 효과는 셀러 동의 과정을 거친 이후 6 월부터 반영될 전망이다. 따라서 커머스 매출 성장률 역시 상저하고 형태를 예상한다.

콘텐츠(+6%yoy)는 전분기 국내에서 불거진 불매운동으로 인한 부정적 영향이 사그라들면서 매출 성장률의 소폭 회복이 기대되며, 비용에서는 플러스스토어 별도앱 출시에 따른 마케팅 비용의 유의미한 증가가 예상된다.

지난해 네이버는 네이버앱 개편을 통한 트래픽 증가로 광고 매출 성장이 실적을 견인했다면, 올해는 플러스스토어 별도앱 출시에 따른 커머스 매출의 반등 효과가 실적을 견인할 것이다. 또한 플러스스토어앱 출시 효과는 비단 커머스 내 중개 및 판매 부문 뿐만 아니라 커머스 광고 매출에도 긍정적으로 영향을 미쳐 실질적으로 광고와 커머스 매출 전반의 양호한 성장이 기대된다.

2025 네이버 AI 전략: 온서비스 AI(On-Service AI)

네이버의 최수연 대표는 연초 4 분기 실적 발표회에서 2025 년 네이버의 전략을 '온서비스 AI'라 요약했다. 온서비스 AI 란 AI 기술을 고도화하여 이를 네이버가 전개하는 모든 서비스(광고, 커머스, 콘텐츠 등)에 적용해 서비스 차별화에 나서겠다는 것이다. 네이버의 온서비스 AI 전략은 하이퍼클로바(2021 년)와 하이퍼클로바 X(2023 년)를 거쳐 확립된 것으로 파악된다. 하이퍼클로바는 출시 이후 경쟁 모델 대비 성능이 뒤처지며, 실용성이 떨어진다는 지적을 받았다. 이후 성능을 한 단계 올리고, 상용 서비스까지 이어진 하이퍼클로바 X 를 출시했지만 여전히 서비스 경쟁력이 떨어졌으며 자칫 비용 부담만 가중될 수 있었다. 네이버는 비용 부담은 줄이면서 AI 모델을 활용한 수익성 개선을 지향했으며, 실용성과 수익성을 모두 챙기기 위한 방안으로 온서비스 AI를 꺼내든 것이다.

먼저 플랫폼 이용자는 온서비스 AI 를 통해 본인 맞춤형 광고 콘텐츠를 자동으로 추천받게 된다. 물론 온서비스 AI 가 도입되기 이전에도 사용자의 검색 이력이나 구매 기록을 바탕으로 광고를 띄우는 리타겟팅 광고(behavioral retargeting)는 존재했으나, 온서비스 AI는 단순 검색 기록을 넘어서 최근 이용자의 활동을 종합해 분석하고 이에 맞는 카테고리 상품을 추천하게 된다. 이 과정에서 광고를 콘텐츠처럼 자연스럽게 배치해 사용자의 거부감을 낮추고, 정보로 받아들이게 만든다. 또한 검색 서비스의 AI 브리핑 기능과 연계해 광고를 자연스럽게 노출시키기도 하는데, 이를 통해 광고 클릭률과 전환율을 높일 수 있다.

네이버는 지난해 10 월 '초개인화'를 내세운 네이버 플러스스토어를 선보인 바 있다. AI 를 통해 개인의 취향과 관심사를 분석해 상품, 혜택, 프로모션, 쇼핑 관련 콘텐츠 등을 추천하는 형태다. 또한 발견 탭에서는 개인화된 숏폼(쇼츠 형태) 콘텐츠를 통해 사용자가 새로운 상품을 탐색하고 바로 구매할 수 있도록 지원한다. 이 외에도 콘텐츠 플랫폼(블로그, 카페, 웹툰 등)에서는 음악 추천, 소음 제거 등 자동화 편집 지원 기능, 카페 글 추천 시스템 고도화 등의 AI 콘텐츠 보조 기능 등을 제공한다.

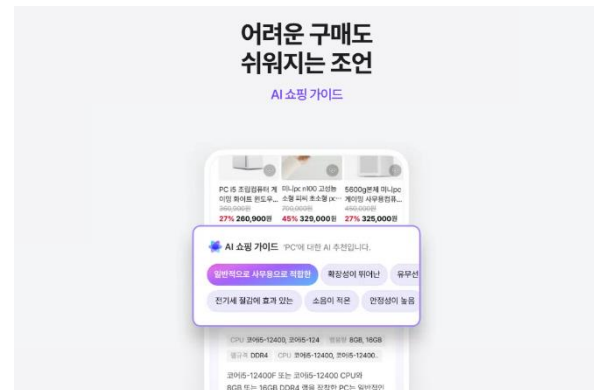
뛰어난 성능의 AI 모델을 보유한 전세계 주요 테크기업들도 수익화 측면에서 어려움을 겪고 있다. 네이버 또한 하이퍼클로바 X 를 개발하면서까지 상용화 서비스를 출시했으나 성능, 편의성, 범용성 등 여러 측면에서 아쉬운 모습을 보였다. 그보다 네이버라는 강력한 플랫폼을 보유하고 있다는 강점을 살린 온서비스 AI 전략이 현재 상황에서 타당한 전략이라 판단한다.

도표 90. '단(DAN)24'에서 발표된 온서비스 AI



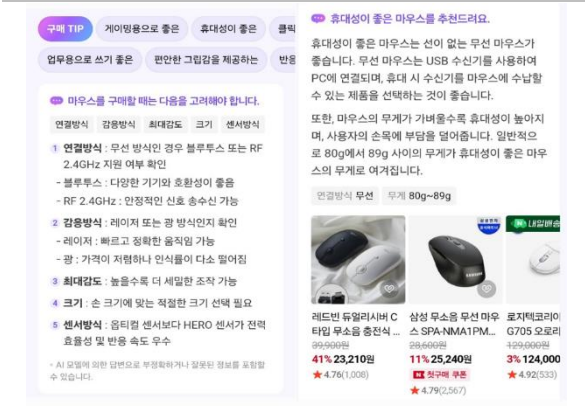
자료: NAVER, 유진투자증권

도표 91. 플러스스토어 'AI 쇼핑 가이드'



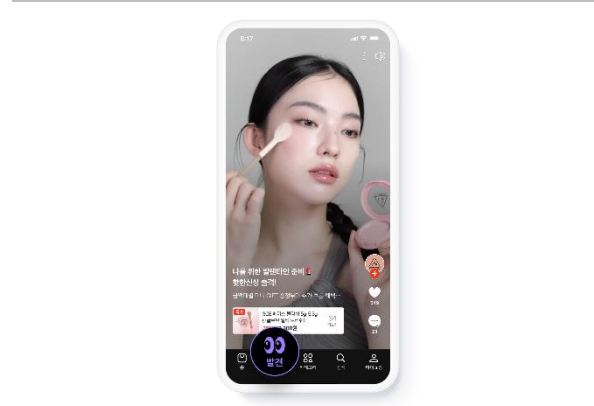
자료: NAVER, 유진투자증권

도표 92. 플러스스토어의 초개인화 추천 예시



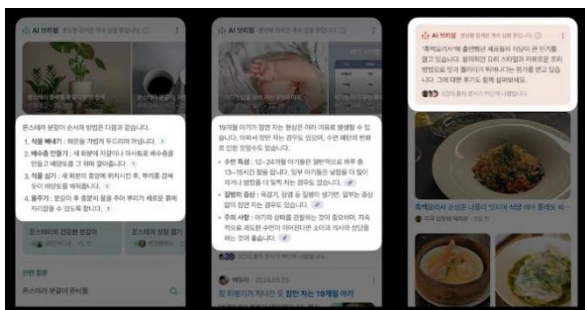
자료: NAVER, 유진투자증권

도표 93. 플러스스토어 '발견' 탭



자료: NAVER, 유진투자증권

도표 94. 네이버검색 'AI 브리핑'



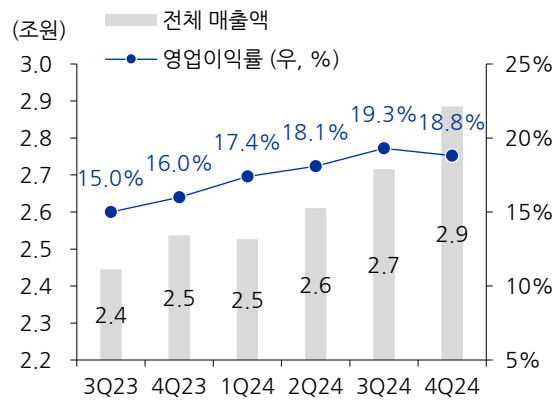
자료: NAVER, 유진투자증권

도표 95. 애드부스트 홈페이지 자동 타겟 광고



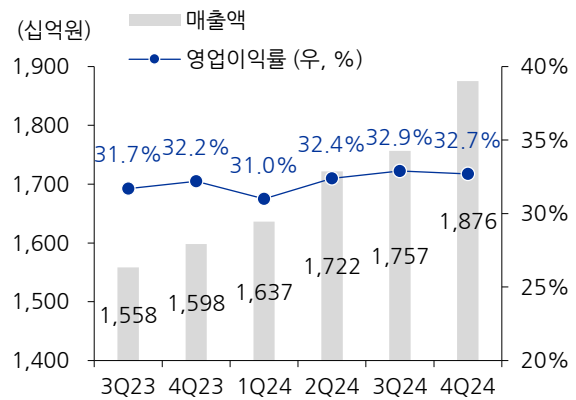
자료: NAVER, 유진투자증권

도표 96. 네이버 전체



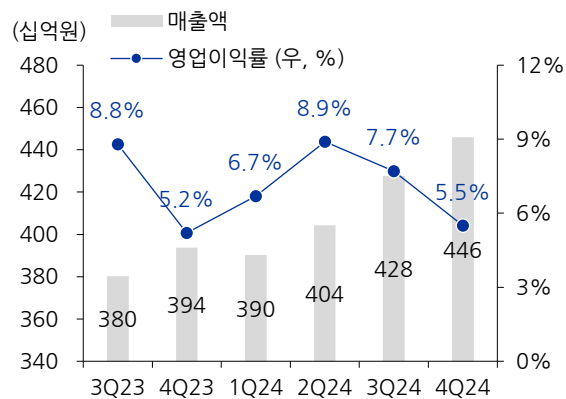
자료: NAVER, 유진투자증권

도표 97. 서치플랫폼 및 커머스



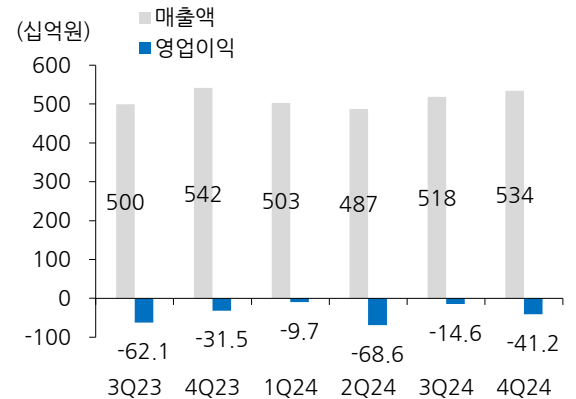
자료: NAVER, 유진투자증권

도표 98. 핀테크



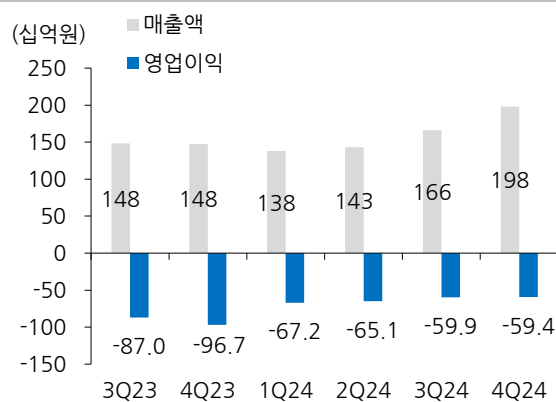
자료: NAVER, 유진투자증권

도표 99. 콘텐츠



자료: NAVER, 유진투자증권

도표 100. 클라우드



자료: NAVER, 유진투자증권

실적전망 및 밸류에이션

도표 101. 네이버 실적 추이

(십억원)	1Q24	2Q24	3Q24	4Q24	1Q25E	2Q25E	3Q25E	4Q25E	2024P	2025E
매출액	2,526.1	2,610.5	2,715.6	2,885.6	2,824.7	2,909.8	3,031.4	3,148.8	10,737.7	11,914.6
(%yoy)	10.8%	8.4%	11.1%	13.7%	11.8%	11.5%	11.6%	9.1%	11.0%	11.0%
서치플랫폼	905.4	978.4	997.7	1,064.7	1,010.9	1,064.2	1,074.2	1,127.4	3,946.2	4,276.8
커머스	703.4	719.0	725.4	775.1	793.7	815.5	872.4	900.4	2,923.0	3,382.0
핀테크	353.9	368.5	385.1	400.9	394.6	405.7	425.6	445.0	1,508.4	1,670.9
콘텐츠	446.3	420.0	462.8	467.3	471.9	460.2	495.2	508.2	1,796.4	1,935.4
클라우드	117.0	124.6	144.6	177.6	153.6	164.3	164.0	167.7	563.7	649.6
영업비용	2,086.8	2,137.8	2,190.3	2,343.6	2,302.5	2,345.2	2,430.5	2,532.1	8,758.5	9,610.3
(%yoy)	7.0%	5.0%	6.1%	10.0%	10.3%	9.7%	11.0%	8.0%	7.0%	9.7%
영업이익	439.3	472.7	525.3	542.0	522.2	564.6	600.9	616.7	1,979.3	2,304.3
(%yoy)	32.9%	26.8%	38.2%	33.7%	18.9%	19.4%	14.4%	13.8%	32.9%	16.4%
OPM(%)	17.4%	18.1%	19.3%	18.8%	18.5%	19.4%	19.8%	19.6%	18.4%	19.3%
세전이익	660.2	467.4	653.3	465.5	615.3	634.6	645.3	673.7	2,246.3	2,568.8
당기순이익	555.8	332.1	530.1	444.1	446.1	460.1	464.6	491.8	1,862.0	1,862.6
NPM(%)	22.0%	12.7%	19.5%	15.4%	15.8%	15.8%	15.3%	15.6%	17.3%	15.6%

자료: 유진투자증권

도표 102. 연간 실적 추정 및 밸류에이션

(십억원)	2022	2023	2024P	2025E
매출액	8,220.1	9,670.6	10,737.7	11,914.6
영업이익	1,304.7	1,488.8	1,979.3	2,304.3
세전이익	1,056.1	1,486.0	2,246.3	2,568.8
지배순이익	752.3	1,012.6	1,877.9	1,871.6
EPS(원)	4,634	6,233	11,853	11,813
PER(배)	38.3	35.9	17.6	17.7
OPM(%)	15.9%	15.4%	18.4%	19.3%
NPM(%)	9.2%	10.5%	17.5%	15.7%

자료: 유진투자증권

도표 103. 네이버 연간 실적 추정치 변경

(십억원)	수정후		수정전		변경률	
	2025E	2026E	2025E	2026E	2025E	2026E
매출액	11,915	12,749	11,792	12,618	1.0%	1.0%
영업이익	2,304	2,511	2,277	2,486	1.2%	1.0%
세전이익	2,569	2,716	2,542	2,691	1.1%	0.9%
지배순이익	1,872	2,010	1,852	1,991	1.1%	1.0%
% of Sales						
영업이익	19.3%	19.7%	19.3%	19.7%		
세전이익	21.6%	21.3%	21.6%	21.3%		
지배순이익	15.7%	15.8%	15.7%	15.8%		

자료: NAVER, 유진투자증권

NAVER(035420.KS) 재무제표

대차대조표

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
자산총계	33,899	35,738	37,697	39,615	41,540
유동자산	6,440	7,028	10,313	12,835	14,330
현금성자산	3,967	4,403	7,470	9,807	11,267
매출채권	1,515	1,706	1,904	2,077	2,103
재고자산	9	15	26	28	28
비유동자산	27,459	28,710	27,383	26,780	27,210
투자자산	23,520	22,523	21,749	22,632	23,551
유형자산	2,458	2,742	2,134	675	214
기타	1,481	3,446	3,501	3,473	3,446
부채총계	10,449	11,500	9,041	9,279	9,385
유동부채	5,481	6,306	3,780	3,947	3,980
매입채무	2,935	3,873	1,723	1,881	1,904
유동성이자부채	1,647	1,465	1,079	1,079	1,079
기타	899	967	977	987	997
비유동부채	4,968	5,194	5,261	5,332	5,406
비유동이자부채	3,306	3,417	3,445	3,445	3,445
기타	1,662	1,777	1,816	1,887	1,960
자본총계	23,450	24,238	28,655	30,336	32,155
지배지분	22,744	23,206	27,151	28,832	30,651
자본금	16	16	16	16	16
자본잉여금	1,556	1,243	1,416	1,416	1,416
이익잉여금	23,646	24,544	26,180	27,860	29,679
기타	(2,475)	(2,597)	(461)	(461)	(461)
비지배지분	706	1,032	1,504	1,504	1,504
자본총계	23,450	24,238	28,655	30,336	32,155
총차입금	4,953	4,883	4,524	4,524	4,524
순차입금	986	480	(2,946)	(5,283)	(6,743)

현금흐름표

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
영업현금	1,453	2,002	2,391	3,225	2,461
당기순이익	673	985	1,862	1,863	2,010
자산상각비	561	582	725	1,514	516
기타비현금성손익	757	745	2,041	(878)	(704)
운전자본증감	307	360	(2,424)	(18)	(2)
매출채권감소(증가)	(56)	43	(205)	(174)	(25)
재고자산감소(증가)	17	(2)	0	(2)	(0)
매입채무증가(감소)	133	347	158	157	23
기타	215	(28)	(2,378)	1	1
투자현금	(1,216)	(950)	(770)	(789)	(904)
단기투자자산감소	(920)	482	(1,145)	41	(37)
장기투자증권감소	445	534	793	(747)	(782)
설비투자	701	641	346	0	0
유형자산처분	12	7	8	0	0
무형자산처분	(55)	(51)	(22)	(27)	(27)
재무현금	(339)	(110)	(41)	(191)	(192)
차입금증가	(392)	(180)	(285)	0	0
자본증가	(152)	(34)	(174)	(191)	(192)
배당금지급	213	62	181	191	192
현금 증감	(57)	852	1,636	2,245	1,365
기초현금	2,781	2,724	3,576	5,213	7,458
기말현금	2,724	3,576	5,213	7,458	8,823
Gross Cash flow	1,991	2,312	5,219	3,243	2,463
Gross Investment	(11)	1,072	2,049	848	869
Free Cash Flow	2,003	1,241	3,170	2,395	1,594

자료: 유진투자증권

손익계산서

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
매출액	8,220	9,671	10,738	11,915	12,749
증가율(%)	20.6	17.6	11.0	11.0	7.0
매출원가	0	0	0	0	0
매출총이익	8,220	9,671	10,738	11,915	12,749
판매 및 일반관리비	6,915	8,182	8,758	9,610	10,237
기타영업손익	26	18	7	10	7
영업이익	1,305	1,489	1,979	2,304	2,511
증가율(%)	(1.6)	14.1	32.9	16.4	9.0
EBITDA	1,866	2,071	2,704	3,818	3,028
증가율(%)	6.1	11.0	30.6	41.2	(20.7)
영업외손익	(221)	(7)	267	265	205
이자수익	107	137	129	44	44
이자비용	72	127	13	52	52
지분법손익	10	244	145	81	79
기타영업외손익	(266)	(261)	6	192	134
세전순이익	1,084	1,481	2,246	2,569	2,716
증가율(%)	(49.0)	36.7	51.6	14.4	5.7
법인세비용	411	496	384	706	706
당기순이익	673	985	1,862	1,863	2,010
증가율(%)	(95.9)	46.3	89.0	0.0	7.9
지배주주지분	760	1,012	1,878	1,872	2,010
증가율(%)	(95.4)	33.2	85.5	(0.3)	7.4
비지배지분	(87)	(27)	(16)	(9)	0
EPS(원)	4,634	6,233	11,853	11,813	12,688
증가율(%)	(95.4)	34.5	90.2	(0.3)	7.4
수정EPS(원)	4,634	6,233	11,853	11,813	12,688
증가율(%)	(95.4)	34.5	90.2	(0.3)	7.4

주요투자지표

	2022A	2023A	2024P	2025F	2026F
주당지표(원)					
EPS	4,634	6,233	11,853	11,813	12,688
BPS	138,642	142,887	171,370	181,978	193,456
DPS	914	1,205	1,210	1,210	1,210
밸류에이션(배, %)					
PER	38.3	35.9	17.6	17.7	16.5
PBR	1.3	1.6	1.2	1.1	1.1
EV/ EBITDA	16.1	17.8	11.2	7.3	8.7
배당수익율	0.5	0.5	0.6	0.6	0.6
PCR	14.6	15.7	6.3	10.2	13.4
수익성(%)					
영업이익율	15.9	15.4	18.4	19.3	19.7
EBITDA이익율	22.7	21.4	25.2	32.0	23.7
순이익율	8.2	10.2	17.3	15.6	15.8
ROE	3.3	4.4	7.5	6.7	6.8
ROIC	66.7	35.1	37.1	36.4	51.7
안정성 (배, %)					
순차입금/자기자본	4.2	2.0	(10.3)	(17.4)	(21.0)
유동비율	117.5	111.5	272.8	325.2	360.1
이자보상배율	18.2	11.7	151.1	44.7	48.7
활동성 (회)					
총자산회전율	0.2	0.3	0.3	0.3	0.3
매출채권회전율	5.8	6.0	5.9	6.0	6.1
재고자산회전율	1,111.4	824.3	532.7	446.8	455.4
매입채무회전율	3.1	2.8	3.8	6.6	6.7

카카오

(035720)

AI 서비스로 쏠리는 시선

투자의견

BUY(유지)

목표주가

48,000 원(유지)

현재주가

44,000 원(03/18)

시가총액

19,524(십억원)

인터넷/게임/우주 정의훈_02)368-6170_uihoon0607@eugenefn.com

- 1분기 매출액 1.95조원(-2%yoy), 영업이익 1,070억원(-11%yoy)으로 컨센서스(1,100억원) 부합 전망
- 플랫폼 부문(+8%yoy)은 특비즈 매출이 전분기와 유사한 성장률을 이어갈 것으로 예상되는 가운데, 플랫폼기타에서 모빌리티와 카카오페이가 yoy 두자릿수 이상의 성장률을 이어갈 전망
- 콘텐츠 부문(-10%yoy) 부진은 지난해에 이어 1분기까지 이어질 전망. 게임은 기존 게임 매출 하향이 지속되는 가운데 신작 '발할라 서바이벌'(1/21출시)은 출시 직후 매출 순위가 크게 빠져 매출 기여도가 미미할 것으로 추정되며, 올해 상반기까지는 기대 신작 모멘텀 부재. 미디어는 방송광고 부진으로 편성 지연 이어지며 매출 역성장 지속
- 올해 카카오는 오픈AI와의 협업을 통한 유의미한 AI서비스(상반기 카나나, AI 메이트 등) 출시를 예고함. 이는 플랫폼 부문 내 다양한 사업부문과의 시너지로 이어질 수 있음. 다만 콘텐츠 부문의 부진이 추세적으로 이어지고 있고 올해도 어려운 시장환경은 여전하므로, AI 서비스의 유의미한 성과와 콘텐츠 부문의 수익성 방어 여부가 상반기 카카오의 주가에 최대 변수로 작용할 전망

주가(원,03/18)	44,000
시가총액(십억원)	19,524

발행주식수	443,721천주
52주 최고가	55,700원
최저가	32,550원
52주 일간 Beta	0.27
60일 일평균거래대금	1,478억원
외국인 지분율	27.7%
배당수익률(2024F)	0.1%

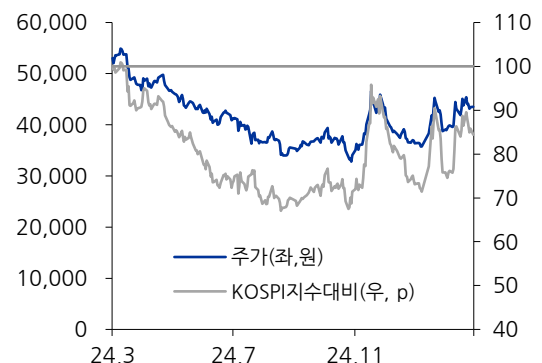
주주구성	
김범수 (외 93인)	24.2%
MAXIMO PTE (외 1인)	6.3%

주가상승	1M	6M	12M
상대기준	13.5	22.2	-14.2
절대기준	13.0	23.6	-17.0

(원, 십억원)	현재	직전	변동
투자의견	BUY	BUY	-
목표주가	48,000	48,000	-
영업이익(24)	460	491	▼
영업이익(25)	519	538	▼

12월 결산(십억원)	2022A	2023A	2024P	2025E
매출액	6,799	7,557	7,872	8,165
영업이익	569	461	460	519
세전손익	1,269	-1,648	-95	569
당기순이익	1,067	-1,817	-253	375
EPS(원)	3,049	-2,276	43	959
증감률(%)	-2.0	적전	흑전	2,126.1
PER(배)	17.4	na	1,021.5	45.9
ROE(%)	13.6	-10.3	0.2	3.9
PBR(배)	2.4	2.5	1.8	1.7
EV/EBITDA(배)	19.0	17.7	11.9	11.4

자료: 유진투자증권



오픈 AI와 손잡은 카카오

올해 2월 카카오는 오픈 AI와 전략적 제휴 체결을 발표했다. 오픈 AI와 제휴를 체결한 것은 국내 기업으로서는 최초이며, 양사는 AI 서비스를 쉽게 사용할 수 있도록 하는 데 초점을 맞춰 'AI 서비스 고도화를 위한 기술 협력'과 '공동 상품 개발'을 추진할 예정이라고 밝혔다.

기술 협력이나 공동 상품 개발과 관련해 구체적인 발표는 없었지만, 기본적으로 오픈 AI의 최신 AI 기술을 국내 시장에 맞게 API 활용한 서비스가 출시될 것으로 보이며, 오픈 AI의 ChatGPT, DALL-E 등의 기존 서비스와 유사한 형태일 것으로 추측된다. 물론 이 같은 서비스가 어떤 식으로 출시될 것인지(1) 별도앱 출시, 2) 카카오톡 내 친구추가, 3) 카카오톡 내 별도탭 신설)는 아직까지 미지수이다.

카카오는 그간 AI 개발에서 데이터센터 인프라 부족, 비용 부담 등의 문제를 해결하기 위해 AI 오케스트레이션 전략을 취했다. 오픈 AI와의 협업으로 유의미한 성능을 지닌 AI 서비스가 카카오톡에 접목돼 출시된다면, 그간 카카오가 가지고 있던 카카오톡의 유의미한 유저 트래픽 향상이라는 숙제를 해결할 수 있다. 수익화 측면에서도 유료 구독 매출 혹은 트래픽 증가에 따른 광고 매출 성과가 드러날 것으로 기대된다.

도표 104. 카카오-오픈 AI 전략적 제휴 체결



자료: 카카오, 유진투자증권

1Q25 Preview: AI 서비스로 쏘리는 시선

1 분기 매출액 1.95 조원(-2%yoy), 영업이익 1,070 억원(-11%yoy)으로 컨센서스(1,100 억원)에 부합할 전망이다. 플랫폼 부문(+8%yoy)은 톡비즈 매출이 전 분기와 유사한 성장률을 이어갈 것으로 예상되는 가운데, 플랫폼기타에서 모빌리티와 카카오페이가 yoy 두자릿수 이상의 성장률을 이어갈 전망이다.

콘텐츠 부문(-10%yoy)의 부진은 지난해에 이어 1 분기까지 지속될 전망이다. 특히 게임과 미디어의 부진이 심한데, 게임은 기존 게임 매출 하향이 지속되는 가운데 신작 '발할라 서바이벌'(1/21 출시)은 출시 직후 매출 순위가 크게 빠져 매출 기여도가 미미할 것으로 추정된다. 올해 상반기까지는 기대 신작 모멘텀이 부재한 상황이다. 미디어의 경우에도 방송광고 부진으로 잇따른 편성 지연이 겹쳐 매출 역성장이 지속되고 있다.

올해 카카오는 오픈 AI 와의 협업을 통한 유의미한 AI 서비스(상반기 카나나, AI 메이트 등) 출시를 예고했다. 이는 플랫폼 부문 내 다양한 사업부문과의 시너지로 이어질 수 있다. 다만 콘텐츠 부문의 부진이 추세적으로 이어지고 있으며, 올해도 어려운 시장환경은 지속될 전망이다. 따라서 AI 서비스의 유의미한 성과와 콘텐츠 부문 수익성 방어 여부가 상반기 카카오의 주가에 최대 변수로 작용할 전망이다.

실적전망 및 밸류에이션

도표 105. 카카오 실적 추이

(십억원)	1Q24	2Q24	3Q24	4Q24P	1Q25E	2Q25E	3Q25E	4Q25E	2024P	2025E
매출액	1,988.4	2,004.9	1,921.4	1,957.0	1,957.5	2,029.8	2,045.3	2,132.7	7,871.7	8,165.2
(%yoy)	22.5%	4.2%	-4.5%	-2.1%	-1.6%	1.2%	6.4%	9.0%	4.2%	3.7%
플랫폼	954.8	955.3	943.5	1,047.2	1,031.3	1,038.8	1,024.7	1,130.7	3,900.7	4,225.6
특비즈	522.1	513.9	507.3	562.7	548.3	546.9	536.9	599.2	2,106.1	2,231.4
포털비즈	84.7	87.9	76.4	83.2	81.5	85.0	75.0	82.5	332.2	324.0
플랫폼 기타	348.0	353.5	359.8	401.2	401.5	406.9	412.8	449.0	1,462.5	1,670.2
콘텐츠	1,033.6	1,049.6	977.9	909.9	926.1	990.9	1,020.5	1,002.0	3,971.0	3,939.6
게임	243.1	233.5	233.6	163.0	160.2	157.6	192.2	195.6	873.2	705.6
뮤직	468.2	510.9	470.9	470.0	468.6	521.6	516.9	500.1	1,920.0	2,007.2
스토리	227.0	215.7	218.7	203.0	218.8	227.1	231.5	214.7	864.3	892.2
미디어	95.2	89.6	54.7	73.9	78.5	84.6	79.9	91.6	313.4	334.6
영업비용	1,868.1	1,870.9	1,790.9	1,881.6	1,850.5	1,897.5	1,905.4	1,992.7	7,411.5	7,645.9
(%yoy)	19.7%	3.4%	-5.1%	2.4%	-0.9%	1.4%	6.4%	5.9%	4.4%	3.2%
인건비	479.0	480.6	459.0	498.0	478.0	472.6	474.0	482.6	1,916.6	1,907.2
매출연동비	819.9	796.8	747.7	653.9	655.8	692.2	685.2	746.5	3,018.4	2,779.5
외주/인프라비	202.4	195.6	194.1	287.6	312.5	317.0	326.3	338.9	879.6	1,294.7
마케팅비	89.3	108.2	105.5	103.5	106.7	106.0	111.0	114.0	406.4	437.7
상각비	197.8	207.3	210.7	246.0	215.0	225.6	225.0	225.7	861.8	891.3
기타	79.7	82.5	73.9	92.7	82.5	84.1	83.9	85.0	328.7	335.5
영업이익	120.3	134.0	130.5	75.4	107.0	132.3	139.9	140.1	460.2	519.3
(%yoy)	92.2%	18.5%	5.0%	-53.1%	-11.0%	-1.2%	7.2%	85.7%	-0.1%	12.8%
OPM(%)	6.0%	6.7%	6.8%	3.9%	5.5%	6.5%	6.8%	6.6%	5.8%	6.4%
세전이익	140.6	164.9	135.7	-536.7	130.7	142.7	139.4	156.1	-95.5	568.9
당기순이익	67.7	87.1	78.5	-485.9	86.3	94.2	92.0	103.0	-252.6	375.5
(%yoy)	19.2%	58.8%	112.0%	적자지속	27.5%	8.1%	17.2%	흑자전환	적자지속	흑자전환
NPM(%)	3.4%	4.3%	4.1%	-24.8%	4.4%	4.6%	4.5%	4.8%	-3.2%	4.6%

자료: 유진투자증권

도표 106. 연간 실적 추정 및 밸류에이션

(십억원)	2022	2023	2024P	2025E
매출액	6,799.0	7,557.0	7,871.7	8,165.2
영업이익	569.4	460.9	460.2	519.3
세전이익	1,268.9	-1,648.2	-95.5	568.9
지배순이익	1,352.9	-649.4	19.1	425.3
EPS(원)	3,049	-2,276	43	959
PER(배)	17.4	na	1,021.5	45.9
OPM(%)	8.4%	6.1%	5.8%	6.4%
NPM(%)	19.9%	-8.6%	0.2%	5.2%

자료: 유진투자증권

도표 107. 카카오 연간 실적 추정치 변경

(십억원)	수정후		수정전		변경률	
	2025E	2026E	2025E	2026E	2025E	2026E
매출액	8,165	8,737	8,329	8,912	-2.0%	-2.0%
영업이익	519	629	538	642	-3.5%	-2.0%
세전이익	569	677	587	690	-3.1%	-1.9%
지배순이익	425	453	438	462	-3.0%	-1.9%
% of Sales						
영업이익	6.4%	7.2%	6.5%	7.2%		
세전이익	7.0%	7.7%	7.0%	7.7%		
지배순이익	5.2%	5.2%	5.3%	5.2%		

자료: 카카오, 유진투자증권

카카오(035720.KS) 재무제표

대차대조표

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
자산총계	22,964	25,180	26,166	26,838	27,428
유동자산	7,756	8,943	9,973	11,101	11,651
현금성자산	6,022	6,998	8,033	9,041	9,551
매출채권	772	1,134	1,148	1,251	1,281
재고자산	160	139	113	123	126
비유동자산	14,451	15,090	14,866	14,356	14,339
투자자산	7,464	7,992	8,385	8,725	9,079
유형자산	885	1,322	1,011	320	101
기타	6,102	5,776	5,470	5,311	5,159
부채총계	9,448	11,321	11,319	11,593	11,756
유동부채	4,768	6,245	6,366	6,535	6,590
매입채무	1,642	3,623	1,806	1,968	2,015
유동성이자부채	1,376	1,925	3,856	3,856	3,856
기타	1,750	697	704	711	718
비유동부채	3,798	3,737	3,333	3,372	3,412
비유동이자부채	2,825	2,577	2,378	2,378	2,378
기타	973	1,160	955	993	1,033
자본총계	13,516	13,859	14,847	15,245	15,672
지배지분	9,984	9,764	10,845	11,243	11,670
자본금	45	45	44	44	44
자본잉여금	8,116	8,840	8,927	8,927	8,927
이익잉여금	3,031	1,922	1,905	2,303	2,730
기타	(1,209)	(1,043)	(31)	(31)	(31)
비지배지분	3,532	4,095	4,002	4,002	4,002
자본총계	13,516	13,859	14,847	15,245	15,672
총차입금	4,200	4,503	6,235	6,235	6,235
순차입금	(1,821)	(2,496)	(1,798)	(2,806)	(3,317)

현금흐름표

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
영업현금	678	1,341	1,847	1,430	986
당기순이익	1,067	(1,817)	(253)	375	460
자산상각비	581	761	1,027	950	470
기타비현금성손익	(455)	2,240	1,043	(36)	29
운전자본증감	(460)	248	(306)	49	15
매출채권감소(증가)	(69)	(108)	(100)	(103)	(30)
재고자산감소(증가)	(73)	43	(28)	(10)	(3)
매입채무증가(감소)	64	364	(67)	162	47
기타	(382)	(51)	(111)	0	0
투자현금	(1,574)	(1,780)	(157)	(457)	(513)
단기투자자산감소	(75)	227	239	(62)	(65)
장기투자증권감소	21	(9)	(44)	(175)	(224)
설비투자	453	575	309	0	0
유형자산처분	14	14	28	0	0
무형자산처분	(91)	(135)	(77)	(100)	(100)
재무현금	412	1,020	(442)	(27)	(27)
차입금증가	265	(75)	(325)	0	0
자본증가	50	(20)	(8)	(27)	(27)
배당금지급	23	55	27	27	27
현금 증감	(490)	553	1,236	946	446
기초현금	5,326	4,836	5,269	6,505	7,451
기말현금	4,836	5,389	6,505	7,451	7,897
Gross Cash flow	1,193	1,184	2,121	1,381	971
Gross Investment	1,960	1,759	702	345	434
Free Cash Flow	(766)	(575)	1,419	1,035	537

자료: 유진투자증권

손익계산서

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
매출액	6,799	7,557	7,872	8,165	8,737
증가율(%)	15.0	11.2	4.2	3.7	7.0
매출원가	0	0	0	0	0
매출총이익	6,799	7,557	7,872	8,165	8,737
판매 및 일반관리비	6,229	7,096	7,411	7,646	8,108
기타영업손익	17	14	4	3	6
영업이익	569	461	460	519	629
증가율(%)	(3.1)	(19.1)	(0.1)	12.8	21.1
EBITDA	1,151	1,222	1,487	1,470	1,100
증가율(%)	19.8	6.1	21.8	(1.2)	(25.2)
영업외손익	700	(2,109)	(556)	50	48
이자수익	179	220	176	93	93
이자비용	110	163	27	109	109
지분법손익	1,603	(135)	54	46	6
기타영업손익	(973)	(2,031)	(758)	20	59
세전순이익	1,269	(1,648)	(95)	569	677
증가율(%)	(44.5)	적전	적지	흑전	19.0
법인세비용	202	168	157	193	217
당기순이익	1,067	(1,817)	(253)	375	460
증가율(%)	(35.0)	적전	적지	흑전	22.6
지배주주지분	1,358	(1,013)	19	425	453
증가율(%)	(2.1)	적전	흑전	2,126.1	6.6
비지배지분	(291)	(804)	(272)	(50)	7
EPS(원)	3,049	(2,276)	43	959	1,022
증가율(%)	(2.0)	적전	흑전	2,126.1	6.6
수정EPS(원)	3,049	(2,276)	43	959	1,022
증가율(%)	(2.0)	적전	흑전	2,126.1	6.6

주요투자지표

	2022A	2023A	2024P	2025F	2026F
주당지표(원)					
EPS	3,049	(2,276)	43	959	1,022
BPS	22,414	21,948	24,452	25,349	26,312
DPS	60	61	60	60	60
밸류에이션(배, %)					
PER	17.4	n/a	1,021.5	45.9	43.0
PBR	2.4	2.5	1.8	1.7	1.7
EV/EBITDA	19.0	17.7	11.9	11.4	14.7
배당수익율	0.1	0.1	0.1	0.1	0.1
PCR	19.8	20.4	9.2	14.1	20.1
수익성(%)					
영업이익율	8.4	6.1	5.8	6.4	7.2
EBITDA이익율	16.9	16.2	18.9	18.0	12.6
순이익율	15.7	(24.0)	(3.2)	4.6	5.3
ROE	13.6	(10.3)	0.2	3.9	4.0
ROIC	7.4	4.9	4.7	4.6	6.2
안정성(배, %)					
순차입금/자기자본	(13.5)	(18.0)	(12.1)	(18.4)	(21.2)
유동비율	162.7	143.2	156.7	169.9	176.8
이자보상배율	5.2	2.8	17.0	4.8	5.8
활동성(회)					
총자산회전율	0.3	0.3	0.3	0.3	0.3
매출채권회전율	9.5	7.9	6.9	6.8	6.9
재고자산회전율	53.9	50.6	62.6	69.2	70.1
매입채무회전율	4.3	2.9	2.9	4.3	4.4

크래프톤 (259960)

트래픽이 어느 정도 높이까지 올라가는 거예요?

투자의견

BUY(유지)

목표주가

470,000 원(상향)

현재주가

351,000 원(03/18)

시가총액

16,816(십억원)

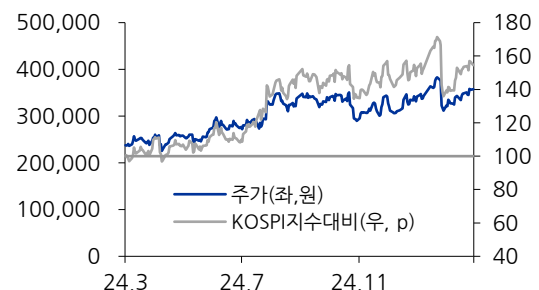
인터넷/게임/우주 정의훈_02)368-6170_uihoon0607@eugenefn.com

- 1분기 매출액 7,799억원(+17%yoy), 영업이익 3,774억원(+22%yoy)으로 컨센서스(3,620억원) 소폭 상회 전망
- PC펍지의 트래픽 상승세는 말 그대로 파죽지세. 4분기에도 높은 트래픽(77~80만명)을 기록했지만, 1월 79만명, 2월 83만명에 이어 3월 현재 90만명을 돌파하며 역대급 분기 트래픽을 달성 중. 이미 작년 한 해 유저 트래픽이 크게 상승했던 터라 연말까지만 해도 높아진 기저의 부담도 없지 않았으나, 여전히 견조한 성장세를 보여주고 있음. 이와 함께 연초 설맞이 성장형 무기스킨 판매로 PC매출은 전년 대비 20% 성장이 예상됨
- 모바일 매출은 중국 지역 매출이 1분기 성수기 효과와 맞물려 매출 반등에 성공한 것으로 파악되며, 기존 펍지모바일도 론도맵 업데이트와 다양한 콜라보 상품 출시로 매출 고성장이 예상됨
- 전분기 실적 발표에서 신규 대형 IP 확보를 위한 개발 투자를 발표해 비용 증가에 대한 우려가 있었으나, 개발 인력 충원의 시차를 감안하면 1분기 인건비 부담도 크지 않을 전망
- 전일(3/18) 크래프톤은 배틀그라운드2의 2025년 개발 계획을 발표, 주요 개발 방향으로 1) 신규 콘텐츠 개발, 2) 라이브 서비스 고도화, 3) 장기적인 서비스 준비 등을 꼽음. 연말 대규모 월드 업데이트 등을 준비하고 있으며, 올해 두 번의 프리미엄 차량 브랜드 콜라보, K팝 아티스트 콜라보 등도 예정되어 있어 트래픽 확대에 따른 수익화 기대
- 올해 출시 예정인 신작인 인조이(3/28 얼리엑세스), 프로젝트 아크(하반기), 서브노티카2(하반기) 또한 기대감을 갖기 충분하다고 판단, 목표주가를 기존 43만원에서 47만원으로 상향하고 투자의견 BUY, 업종 내 Top-pick 유지

주가(원,03/18)	351,000		
시가총액(십억원)	16,816		
발행주식수	47,908천주		
52주 최고가	390,000원		
최저가	225,000원		
52주 일간 Beta	0.58		
60일 일평균거래대금	579억원		
외국인 지분율	41.8%		
배당수익률(2024F)	0.0%		
주주구성			
장병규 (외 33인)	36.6%		
국민연금공단 (외 1인)	6.6%		
주가상승	1M	6M	12M
상대기준	9.7	2.4	50.2
절대기준	9.2	3.8	47.5
(원, 십억원)	현재	직전	변동
투자의견	BUY	BUY	-
목표주가	470,000	430,000	▲
영업이익(24)	1,182	1,182	-
영업이익(25)	1,309	1,226	▲

12월 결산(십억원)	2022A	2023A	2024P	2025E
매출액	1,854	1,911	2,710	3,011
영업이익	752	768	1,182	1,309
세전손익	684	829	1,723	1,505
당기순이익	500	594	1,303	1,129
EPS(원)	10,190	12,310	27,269	23,561
증감률(%)	-4.0	20.8	121.5	-13.6
PER(배)	16.5	15.7	12.9	14.9
ROE(%)	10.3	11.2	21.4	15.6
PBR(배)	1.6	1.7	2.5	2.2
EV/EBITDA(배)	9.1	10.0	12.2	10.4

자료: 유진투자증권



트래픽이 어느정도 높이까지 올라가는 거예요?

1 분기 매출액 7,799 억원(+17%yoy), 영업이익 3,774 억원(+22%yoy)으로 컨센서스(3,620 억원) 소폭 상회를 전망한다.

PC 펌지의 트래픽 상승세는 말 그대로 파죽지세다. 4분기에도 높은 트래픽(10 월 77 만명, 11 월 80 만명, 12 월 80 만명)을 기록했지만, 1 월 79 만명에 이어 2 월에는 83 만명을 기록했고 3 월 현재는 90 만명을 돌파하며 역대급 분기 트래픽을 달성하고 있다. 이미 작년 한 해 동안 크게 상승했던 유저 트래픽이었기 때문에 연말까지만 해도 높아진 기저에 대한 부담도 없지 않았으나, 성장세가 꺾이지 않았음을 보여주고 있다. 이와 함께 연초 설맞이 성장형 무기스킨 판매로 PC 매출은 전년 대비 20% 성장이 예상된다.

모바일 매출은 전분기 기대에 미치지 못했던 중국 지역 매출이 1 분기 성수기 효과와 맞물려 매출 반등에 성공한 것으로 파악된다. 또한 기존 펌지모바일도 론도 맵 업데이트와 다양한 콜라보 상품 출시로 매출 고성장이 전망된다.

전분기 실적 발표에서 신규 대형 IP 확보를 위한 개발 투자를 발표해 비용 증가에 대한 우려가 있었으나, 개발 인력 충원에 시차가 있음을 감안하면 1 분기 인건비 부담도 크지 않을 전망이다.

전일(3/18) 크래프톤은 배틀그라운드 2025 년 개발 계획을 발표했다. 주요 개발 방향으로 1) 신규 콘텐츠 개발, 2) 라이브 서비스 고도화, 3) 장기적인 서비스 준비 등을 꼽았다. 연말 대규모 월드 업데이트, 지형 파괴 시스템 도입 등을 준비하고 있으며 추가적으로 올해 두 번의 프리미엄 차량 브랜드 콜라보, K팝 아티스트 콜라보, 패션 브랜드 콜라보 등도 예정되어 있어 높아진 트래픽에 따른 수익화도 기대할 수 있다.

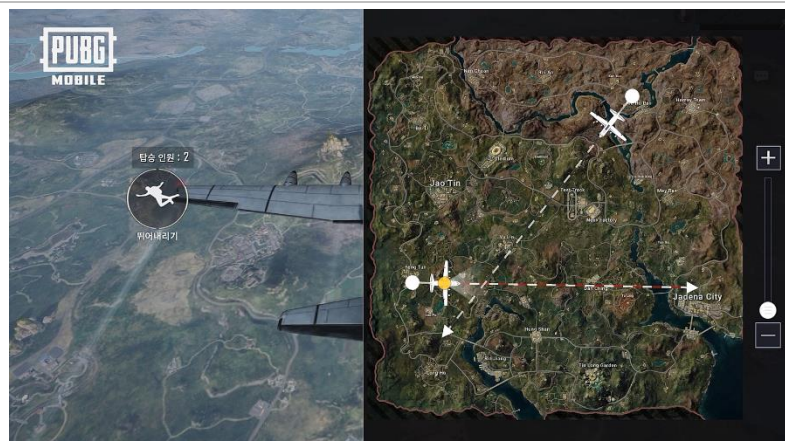
올해 출시 예정인 신작인 인조이(3/28 얼리엑세스), 프로젝트 아크(하반기), 서브노티카 2(하반기) 또한 기대감을 갖기 충분하다고 판단해 목표주가를 기존 43 만원에서 47 만원으로 상향(2025 년 추정 EPS 23,650 원에 타겟 멀티플 20 배 적용)하고 투자의견 BUY와 업종 내 Top-pick을 유지한다.

도표 108. 펍지 PC 설맞이 대축제 성장형 무기 스킨



자료: 크래프톤, 유진투자증권

도표 109. 펍지 모바일 신규 맵 '론도'



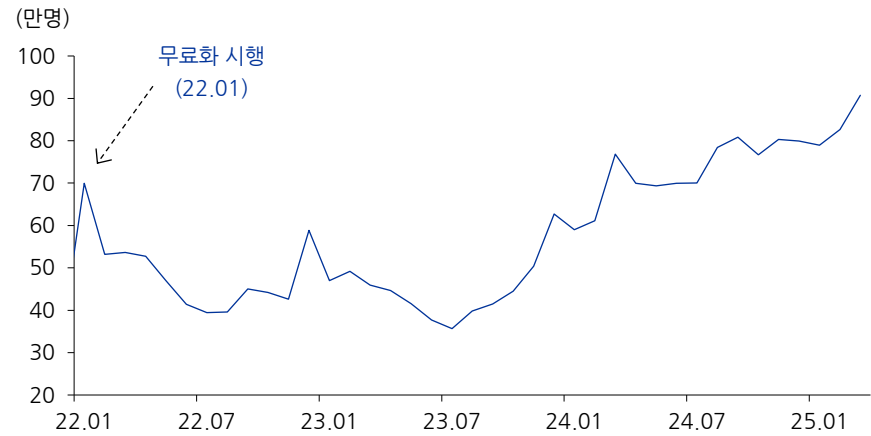
자료: 크래프톤, 유진투자증권

도표 110. 배틀그라운드 2025 로드맵 - 신규 콘텐츠



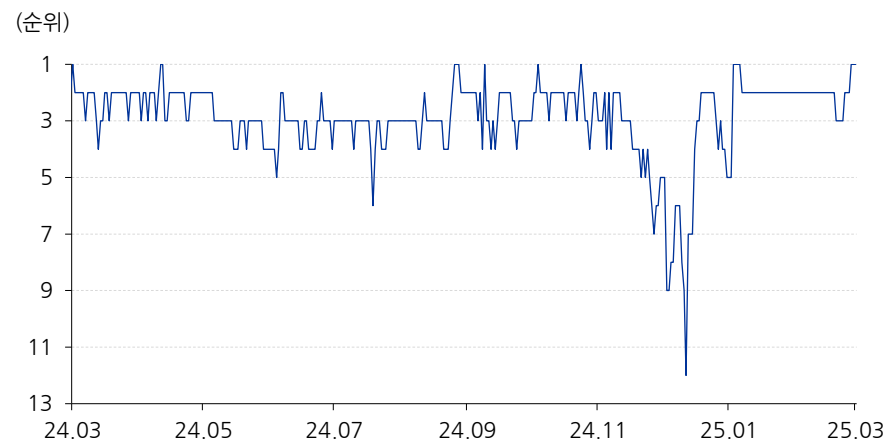
자료: 크래프톤, 유진투자증권

도표 111. 펍지 배틀그라운드 스팀 동시 접속자 수 추이



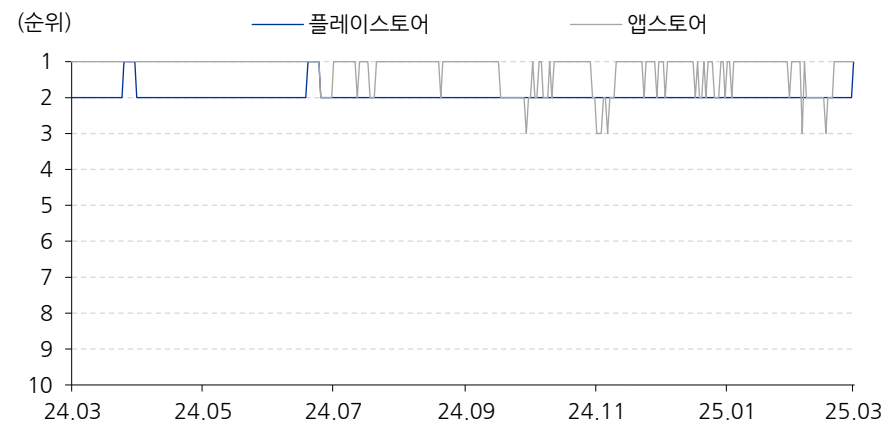
자료: SteamDB, 유진투자증권

도표 112. 중국 화평정영 앱스토어 매출 순위



자료: 모바일인덱스, 유진투자증권

도표 113. 인도 BGMI 매출 순위



자료: 모바일인덱스, 유진투자증권

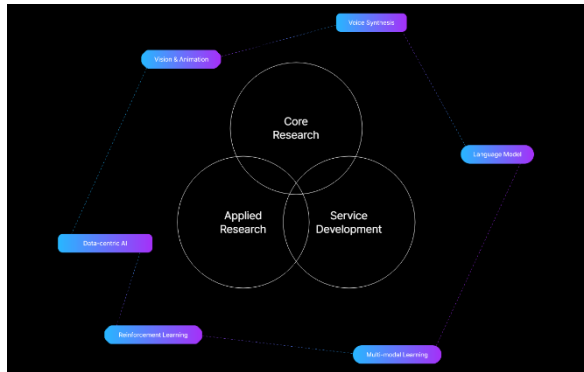
Game with AI에 진심인 크래프톤

크래프톤은 국내 게임 기업 중에서 AI기술 개발에 가장 적극적이다. 2022년부터 AI 전담 조직인 '크래프톤 AI'를 설립하고 이듬해부터 AI 사업 강화를 본격화하며 자사 게임 개발에 AI 기술을 적극적으로 활용하고 있다. 특히 생성형 AI를 중심으로 한 멀티모달 AI, 게임 개발 자동화, AI NPC 등에 집중하고 있는데, 올해 초 열린 CES2025에서 크래프톤은 게임 이용자와 상호작용할 수 있는 AI 협력 캐릭터 'CPC(Co-Playable Character)'를 발표했다. CPC는 엔비디아 에이스(ACE) 기술로 구축된 게임 특화 온디바이스 소형언어모델을 기반으로 이용자와 대화할 수 있으며, 마치 사람처럼 주변 상황을 인식해 대응할 수 있는 것이 특징이다.

3월 28일 얼리엑세스로 출시 예정인 신작 인생 시뮬레이션 게임 '인조이(INZOI)'에 CPC가 적용될 예정인데, AI CPC로 헨리와 프시캣이 있다. 헨리는 전통적인 NPC의 역할을 하지만 이용자 반응에 따라 수많은 행동 선택지를 제공할 수 있다. 이는 헨리의 행동 트리가 정해져있는 것이 아니라 기본적인 방향성만 유지된 채 AI가 직접 행동 트리를 만들기 때문이다. 프시캣은 SLM을 기반으로 만들어진 챗봇으로, 기본적으로 게임을 안내해주는 역할을 하는 CPC이다. 게임 플레이 상 이해가 어려운 부분에 대해 자유롭게 질문하면 프시캣이 이를 인식하고 적절한 공략법을 제시해준다.

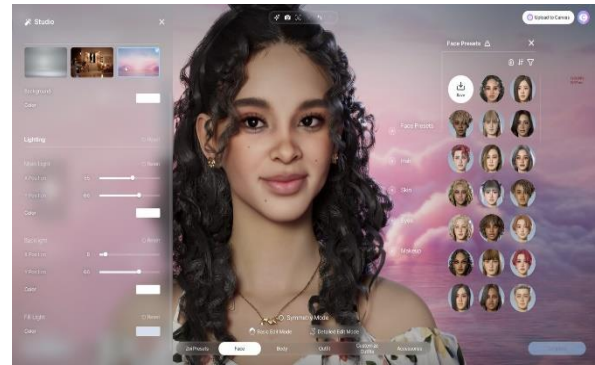
배틀그라운드에서는 자율적으로 학습하고 행동하며 플레이어를 돕는 '동료 CPC'를 도입할 예정이다. 플레이어가 음성으로 탄약과 차량을 요청하면 온디바이스 소형언어모델(SLM)로 구축된 CPC가 이를 수행한다. 요청에 반응하고 정해진 역할을 수행하는 것은 단순한 '봇' 캐릭터들도 수행할 수 있는 일이지만, CPC는 이보다 나아가 상호작용을 구현했다. 실제 사람과 마찬가지로 소통이 가능하고 게임에 대한 이해도가 높다. 적을 발견하면 경고하고, 상황을 판단해 공격 대신 은신을 선택하는 등 플레이어와의 경험을 학습해 나간다.

도표 114. '크래프톤 AI' 조직 구성



자료: 크래프톤, 유진투자증권

도표 115. 인조이 AI 캐릭터 스튜디오



자료: 엔비디아, 유진투자증권

도표 116. 인조이 AI CPC 'Smart Zoi'



자료: 엔비디아, 유진투자증권

도표 117. 배틀그라운드 AI 동료 CPC 'PUBG Ally'



자료: 엔비디아, 유진투자증권

실적전망 및 밸류에이션

도표 118. 크래프톤 실적 추이

(십억원)	1Q24	2Q24	3Q24	4Q24P	1Q25E	2Q25E	3Q25E	4Q25E	2024P	2025E
매출액	665.9	707.0	719.3	617.6	789.9	765.1	782.6	673.6	2,709.8	3,011.2
(%yoy)	23.6%	82.7%	59.7%	15.5%	18.6%	8.2%	8.8%	9.1%	41.8%	11.1%
온라인	243.7	191.3	274.3	232.6	291.1	288.1	318.7	252.9	941.9	1,150.8
모바일	402.3	499.9	425.4	362.2	479.7	457.6	442.6	398.0	1,689.8	1,777.9
콘솔	11.5	8.8	11.8	12.3	13.0	13.0	13.5	13.7	44.3	53.2
기타	8.5	7.0	7.8	10.5	6.1	6.4	7.8	9.0	33.7	29.3
영업비용	355.4	374.9	394.9	402.1	412.5	417.9	422.5	449.6	1,527.3	1,702.5
(%yoy)	39.0%	46.7%	51.3%	8.6%	16.1%	11.5%	7.0%	11.8%	33.7%	11.5%
인건비	121.4	129.6	133.1	132.7	138.6	149.0	156.6	160.1	516.8	604.3
앱수수료/매출원가	85.9	77.5	101.1	84.8	102.7	103.3	105.7	90.9	349.3	402.6
지급수수료	61.3	76.4	80.4	97.9	83.2	91.0	85.7	105.0	316.1	364.9
광고선전비	12.4	21.2	29.9	37.9	45.0	29.1	28.5	40.0	101.4	142.6
주식보상비용	42.1	35.2	14.5	10.5	10.5	12.0	12.0	15.0	102.4	49.5
기타	32.3	34.9	35.9	38.3	32.5	33.5	34.0	38.6	141.3	138.6
영업이익	310.5	332.1	324.4	215.5	377.4	347.2	360.1	224.0	1,182.5	1,308.7
(%yoy)	9.7%	152.6%	71.4%	31.1%	21.5%	4.5%	11.0%	3.9%	54.0%	10.7%
OPM(%)	46.6%	47.0%	45.1%	34.9%	47.8%	45.4%	46.0%	33.2%	43.6%	43.5%
영업외손익	154.8	126.0	-127.7	387.1	87.9	49.2	33.2	25.8	540.2	196.0
세전이익	465.3	458.1	196.7	602.6	465.3	396.4	393.3	249.8	1,722.7	1,504.7
당기순이익	348.6	341.4	121.4	491.2	348.9	297.3	295.0	187.3	1,302.6	1,128.5
(%yoy)	30.5%	165.7%	-42.6%	흑자전환	0.1%	-12.9%	143.1%	-61.9%	119.3%	-13.4%

자료: 유진투자증권

도표 119. 연간 실적 추정 및 밸류에이션

(십억원)	2022	2023	2024P	2025E
매출액	1,854.0	1,910.6	2,709.8	3,011.2
영업이익	751.6	768.0	1,182.5	1,308.7
세전이익	683.9	828.6	1,722.7	1,504.7
지배순이익	500.2	595.4	1,306.1	1,128.5
EPS(원)	10,190	12,310	27,269	23,561
PER(배)	16.5	15.7	12.9	14.9
OPM(%)	40.5%	40.2%	43.6%	43.5%
NPM(%)	27.0%	31.2%	48.2%	37.5%

자료: 유진투자증권

도표 120. 크래프톤 연간 실적 추정치 변경

(십억원)	수정후		수정전		변경률	
	2025E	2026E	2025E	2026E	2025E	2026E
매출액	3,011	3,252	2,924	3,157	3.0%	3.0%
영업이익	1,309	1,382	1,226	1,342	6.8%	3.0%
세전이익	1,505	1,485	1,376	1,444	9.4%	2.8%
지배순이익	1,129	1,121	1,032	1,091	9.4%	2.7%
% of Sales						
영업이익	43.5%	42.5%	41.9%	42.5%		
세전이익	50.0%	45.7%	47.1%	45.7%		
지배순이익	37.5%	34.5%	35.3%	34.6%		

자료: 크래프톤, 유진투자증권

도표 121. 크래프톤 목표주가 변경

적용 EPS(원)	23,650	2025 년 추정 EPS
적용 PER(배)	20.0	
목표주가(원)	470,000	기존 43 만원 대비 9.3% 상향

자료: 유진투자증권

크래프톤(259960.KS) 재무제표

대차대조표

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
자산총계	6,038	6,440	7,567	8,703	9,832
유동자산	3,893	3,964	4,928	6,116	7,279
현금성자산	829	851	1,727	2,819	3,783
매출채권	558	717	781	852	1,028
재고자산	0	0	0	0	0
비유동자산	2,145	2,476	2,639	2,587	2,552
투자자산	1,054	1,611	1,807	1,838	1,871
유형자산	223	257	230	177	137
기타	868	608	601	572	544
부채총계	921	882	903	911	919
유동부채	411	521	534	536	538
매입채무	223	224	0	0	0
유동성이자부채	64	73	307	307	307
기타	124	224	227	229	231
비유동부채	510	361	369	375	380
비유동이자부채	375	215	235	235	235
기타	135	146	134	139	145
자본총계	5,116	5,559	6,664	7,792	8,913
지배지분	5,111	5,554	6,662	7,790	8,911
자본금	5	5	5	5	5
자본잉여금	1,449	1,472	1,478	1,478	1,478
이익잉여금	3,468	3,895	5,081	6,210	7,331
기타	190	183	97	97	97
비지배지분	5	4	2	2	2
자본총계	5,116	5,559	6,664	7,792	8,913
총차입금	439	288	543	543	543
순차입금	(390)	(563)	(1,184)	(2,277)	(3,240)

현금흐름표

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
영업현금	513	662	1,658	1,124	996
당기순이익	500	594	1,303	1,129	1,121
자산상각비	107	108	103	83	68
기타비현금성손익	253	216	297	(68)	(69)
운전자본증감	(134)	(139)	163	(93)	(198)
매출채권감소(증가)	(42)	(148)	(89)	(71)	(176)
재고자산감소(증가)	0	0	0	0	0
매입채무증가(감소)	8	(62)	(15)	0	0
기타	(100)	70	267	(22)	(22)
투자현금	(2,863)	(394)	(532)	(37)	(38)
단기투자자산감소	(2,581)	166	(271)	(5)	(5)
장기투자증권감소	(55)	(188)	(178)	0	0
설비투자	27	34	18	0	0
유형자산처분	0	1	0	0	0
무형자산처분	5	25	(3)	(1)	(1)
재무현금	(56)	(225)	(244)	0	0
차입금증가	(46)	(57)	(47)	0	0
자본증가	0	(168)	(197)	0	0
배당금지급	0	0	0	0	0
현금 증감	(2,345)	46	892	1,088	958
기초현금	3,019	675	721	1,613	2,701
기말현금	675	721	1,613	2,701	3,660
Gross Cash flow	861	919	1,722	1,217	1,195
Gross Investment	417	700	98	125	232
Free Cash Flow	444	219	1,623	1,093	963

자료: 유진투자증권

손익계산서

(단위:십억원)	2022A	2023A	2024P	2025F	2026F
매출액	1,854	1,911	2,710	3,011	3,252
증가율(%)	(1.7)	3.1	41.8	11.1	8.0
매출원가	0	0	0	0	0
매출총이익	1,854	1,911	2,710	3,011	3,252
판매 및 일반관리비	1,102	1,143	1,527	1,702	1,870
기타영업손익	(11)	4	34	11	10
영업이익	752	768	1,182	1,309	1,382
증가율(%)	15.5	2.2	54.0	10.7	5.6
EBITDA	859	876	1,285	1,392	1,450
증가율(%)	19.6	2.0	46.7	8.3	4.2
영업외손익	(68)	61	540	196	103
이자수익	34	38	0	1	1
이자비용	8	9	1	6	6
지분법손익	(35)	(40)	(30)	0	0
기타영업손익	(59)	71	572	201	108
세전순이익	684	829	1,723	1,505	1,485
증가율(%)	(10.2)	21.2	107.9	(12.7)	(1.3)
법인세비용	169	235	420	376	364
당기순이익	500	594	1,303	1,129	1,121
증가율(%)	(3.8)	18.8	119.3	(13.4)	(0.7)
지배주주지분	500	595	1,306	1,129	1,121
증가율(%)	(3.8)	19.0	119.4	(13.6)	(0.7)
비지배지분	(0)	(1)	(3)	0	0
EPS(원)	10,190	12,310	27,269	23,561	23,404
증가율(%)	(4.0)	20.8	121.5	(13.6)	(0.7)
수정EPS(원)	10,190	12,310	27,269	23,561	23,404
증가율(%)	(4.0)	20.8	121.5	(13.6)	(0.7)

주요투자지표

	2022A	2023A	2024P	2025F	2026F
주당지표(원)					
EPS	10,190	12,310	27,269	23,561	23,404
BPS	104,137	114,849	139,086	162,647	186,051
DPS	0	0	0	0	0
밸류에이션(배, %)					
PER	16.5	15.7	12.9	14.9	15.0
PBR	1.6	1.7	2.5	2.2	1.9
EV/ EBITDA	9.1	10.0	12.2	10.4	9.4
배당수익율	0.0	0.0	n/a	n/a	n/a
PCR	9.6	10.2	9.8	13.8	14.1
수익성(%)					
영업이익율	40.5	40.2	43.6	43.5	42.5
EBITDA이익율	46.3	45.9	47.4	46.2	44.6
순이익율	27.0	31.1	48.1	37.5	34.5
ROE	10.3	11.2	21.4	15.6	13.4
ROIC	20.9	13.4	20.9	22.0	22.9
안정성 (배, %)					
순차입금/자기자본	(7.6)	(10.1)	(17.8)	(29.2)	(36.3)
유동비율	946.0	761.4	923.1	1,140.8	1,352.1
이자보상배율	99.2	86.2	909.6	233.7	246.8
활동성 (회)					
총자산회전율	0.3	0.3	0.4	0.4	0.4
매출채권회전율	3.4	3.0	3.6	3.7	3.5
재고자산회전율	n/a	n/a	n/a	n/a	n/a
매입채무회전율	6.6	8.6	n/a	n/a	n/a

Compliance Notice

당사는 자료 작성일 기준으로 지난 3개월 간 해당종목에 대해서 유가증권 발행에 참여한 적이 없습니다

당사는 본 자료 발간일을 기준으로 해당종목의 주식을 1% 이상 보유하고 있지 않습니다

당사는 동 자료를 기관투자자 또는 제 3자에게 사전 제공한 사실이 없습니다

조사분석담당자는 자료작성일 현재 동 종목과 관련하여 재산적 이해관계가 없습니다

동 자료에 게재된 내용들은 조사분석담당자 본인의 의견을 정확하게 반영하고 있으며, 외부의 부당한 압력이나 간섭 없이 작성되었음을 확인합니다

동 자료는 당사의 제작물로서 모든 저작권은 당사에게 있습니다

동 자료는 당사의 동의 없이 어떠한 경우에도 어떠한 형태로든 복제, 배포, 전송, 변형, 대여할 수 없습니다

동 자료에 수록된 내용은 당사 리서치센터가 신뢰할 만한 자료 및 정보로부터 얻어진 것이나, 당사는 그 정확성이나 완전성을 보장할 수 없습니다. 따라서 어떠한 경우에도 동 자료는 고객의 주식투자의 결과에 대한 법적 책임소재에 대한 증빙자료로 사용될 수 없습니다

투자기간 및 투자등급/투자의견 비율

종목추천 및 업종추천 투자기간: 12 개월 (추천기준일 종가대비 추천종목의 예상 목표수익률을 의미함)

당사 투자의견 비율(%)

· STRONG BUY(매수)	추천기준일 종가대비 +50%이상	0%
· BUY(매수)	추천기준일 종가대비 +15%이상 ~ +50%미만	96%
· HOLD(중립)	추천기준일 종가대비 -10%이상 ~ +15%미만	4%
· REDUCE(매도)	추천기준일 종가대비 -10%미만	0%

(2024.12.31 기준)

