

생성 AI 심화편 ②

The Great Monetization: 그 많은 파이는 누가 다 먹을까

(Analyst) 박연주 yeonju.park@miraeasset.com (Analyst) 김수진 soojin.kim@miraeasset.com (Analyst) 임희석 heeseok.lim@miraeasset.com
(Analyst) 김영건 younggun.kim.a@miraeasset.com (Analyst) 박용대 yongdai.park@miraeasset.com (RA) 박주연 park.zuyeon@miraeasset.com



CONTENTS

I. Investment Summary	3
II. 주요 산업 트렌드와 시사점	6
III. 반도체	16
1. AI 추론 시장의 성장성	16
2. 주요 기업들의 AI 대응 전략	19
3. AI 반도체 투자포인트	39
IV. 소프트웨어/통신	41
1. 빅테크 생성AI 비즈니스 현황 - 승자는?	41
2. 수익화의 기로 - 1) LLM 효율화 진행 중	56
3. 수익화의 기로 - 2) 결국 Vertical 접근 중요	64
4. 통신 서비스	66
V. 인터넷/게임	67
1. 광고: 타겟팅 고도화 및 수수료 BM 등장 예상	67
2. 커머스: 구매 전환율 상승, 라이브커머스 활성화	72
3. 콘텐츠: 생성 AI 도입을 통한 공급난 해소 기대	75
4. 게임: 생산성 증대와 비용 감소에 따른 공급 확대 전망	79
VI. 금융	92
1. 금융 AI 시장은 26년 562억달러까지 성장할 전망	92
2. 금융권에서의 생성형 AI의 역할	93
3. 해외 금융업계의 생성형 AI 도입 사례	94
4. 국내 금융업계의 생성형 AI 도입 현황	99
5. 금융기관의 생성형 AI 기술 채택 증가 전망	103
6. 생성형 AI 도입을 위해 풀어나가야 할 과제	104
Global Company Analysis	105
엔비디아	106
마이크로소프트	110
어도비	116
유니티	120
테슬라	130
NAVER	136
SK텔레콤	150
삼성SDS	156
Global X Artificial Intelligence & Technology	160
Global X Cyber Security ETF	162

I. Investment Summary

본격화되는 생성 AI의 수익화: 그 많은 파이는 누가 다 먹을까

생성 AI가 본격적인 상용화 단계에 진입하면서 그 막대한 부가가치를 누가 가져갈 것인지 관심이 커지고 있다. 인터넷/모바일 혁명을 되돌아 보면 향후 2~3년 내에 승자가 결정될 가능성이 높다. 아직 초기지만, 주목할 트렌드는 다음과 같다.

첫째, **기반 기술의 확산 속도가 예상보다 빨라 기업의 비즈니스 전략이 더 중요해질** 전망이다. 특히 메타의 오픈 소스 전략으로 인해 다양한 크기와 성능을 가진 오픈 소스 모델이 대거 출시되었다. 이에 방대한 고객 기반과 질 좋은 데이터, 민첩한 시장 대응력과 저작권 등 법적 이슈 대응, 초기 시장 선점 등 비즈니스 전략의 중요성이 커졌다.

둘째, **응용 서비스의 성장 속도는 시장 기대를 상회할 가능성이 높고, 이에 따라 추론 영역의 중요성이 높아질** 전망이다. 상용화 가능한 수준의 기술을 확보하기가 용이해졌고 주요 기업들이 기업 가치를 확대하기 위해 경쟁적으로 생성 AI 제품을 출시하면서 어플리케이션 시장은 급격히 성장할 전망이다. 기술 특성상 재생산의 비용이 제로에 가깝게 하락하기 때문에 콘텐츠 등 다양한 분야에서 생산량이 급격히 증가할 전망이다.

셋째, **기반 기술이 빠르게 발전 중이기 때문에 잠재 시장 규모와 경쟁 구도가 급변할** 가능성이 높다. 생성 AI 기술은 답변의 정확도/신뢰도 개선, 멀티 모달로의 발전, 도구 사용, Agent간 상호 작용을 통한 발전 등의 방향으로 이루어지고 있으며 그 진전 정도에 따라 AI가 할 수 있는 일의 범위와 신뢰도, 선발 업체의 기술적 경쟁 우위 등이 달라질 수 있다. 특히 도구를 사용하는 톨포머, AI 에이전트끼리의 협업을 통해 복잡한 업무를 수행하는 생성적 에이전트 기술은 잠재력이 높다.

넷째, **어떤 서비스가 킬러 어플리케이션이 되고 누가 그 시장을 장악할 지는 향후 2~3년 내에 결정될** 전망이다. 몇 가지 시그널은 확인되고 있는데 완전히 개인화된 AI 비서가 새로운 플랫폼으로 부상할 가능성이 있고, 단순한 생산성 개선 서비스는 차별성을 잃고 차별화된 서비스로 고객군을 빠르게 확보하는 것이 중요할 것이라는 점이다. 이는 향후 사용자 트래픽 추세를 보면서 승자를 확인해 나갈 필요가 있다. 나아가 생성 AI 기술을 기반으로 메타버스, 로봇틱스 산업의 성장 잠재력도 높아 보인다.

산업별 영향 및 Top Picks, 추천 ETF(AIQ, BUG ETF)

반도체 산업은 추론용 수요가 급증하는 가운데 엔비디아가 추론에서도 주도권을 유지할 것으로 예상된다. 빅테크 업체들이 전용 반도체를 개발하고 있지만 개발 생태계를 오픈하기는 어려워 빅테크 외 대부분 업체들은 엔비디아 CUDA 생태계를 벗어나기 어려울 전망이다. 한편 추론용 반도체 수요가 증가하면서 IP(반도체 설계 자산) 업체들의 수혜가 기대된다. 빅테크들의 전용 반도체 개발 시도가 지속될 것이고 옛지 디바이스 시장에서는 칩 메이커보다는 IP 업체(나스닥의 ARM, 국내의 칩스앤미디어, 오픈엠티테크놀로지 등)에 더 큰 기회가 될 수 있기 때문이다.

Top Pick: 엔비디아(NVDA US Equity)

소프트웨어/통신서비스 어플리케이션 개발 플랫폼에서 구글과 아마존, 어플리케이션 중에서는 이미지 생성AI 시장을 주도하는 어도비가 성과를 내기 시작할 것으로 전망한다. 사이버보안 산업도 구조적 성장이 예상된다. 스스로 진화하는 생성AI의 특성 때문에 보안 방법도 늘어나고, 해킹 공격도 늘어나고 있기 때문이다. 통신 업체들도 새로운 성장 동력으로 AI에 눈을 돌리고 있다. 특히 SK텔레콤은 에이닷 뿐만 아니라 엔트로픽 투자 등을 통해 기술력을 확보했을 뿐만 아니라, 글로벌 얼라이언스를 통해 해외 판매 전략도 구축하고 있다.

Top Pick: 글로벌 X Cybersecurity ETF(BUG US Equity), 어도비(ADBE US Equity), SK텔레콤(017670 KS Equity)

인터넷 산업에서 광고 플랫폼 업체들의 타케팅 능력 고도화와 그에 따른 광고 단가 상승, 멀티 모달 서비스 도입에 따른 커머스 분야의 고도화, 웹툰, 웹소설 등 콘텐츠 공급난을 완화가 기대된다.

Top Pick: NAVER(035420 KS Equity)

게임 산업에서는 1) 프로그래밍 생산성 향상, 2) 소규모 개발 활성화, 3) 노코드 툴 도입에 따른 개발자 저변 확대, 4) 아트 리소스 재분배 효과 등으로 공급이 확대되어 게임 엔진사의 수혜가 기대된다.

Top Pick: 유니티(U US Equity)

금융 산업에서는 생성형 AI를 통해 1) 생산성 및 효율성을 향상시키고, 2) 고객경험을 개선시키고, 3) 보안 및 리스크 관리 강화를 모색할 수 있다. 남은 과제는 1) 정보 신뢰성 한계 2) 고객 데이터 보호 우려 3) 도입비용 부담 4) 규제 도입 등이다.

한편 **테슬라**는 생성 AI에 따른 GPU 공급 부족 속에서 자체 개발한 슈퍼 컴퓨터 도조의 경쟁 우위가 강화될 전망이고 언어 모델의 발전으로 로봇의 인지 및 판단 능력이 개선되면서 중장기적으로 오피머스의 기업 가치도 커질 전망이다.

Top Pick: 테슬라(TSLA US Equity)

생성 AI에 따른 응용 서비스 시장 규모가 급증하는 가운데 막대한 부가가치가 창출되고 이를 소수의 승자가 가져갈 것이라는 점은 분명하지만, 그 승자가 구체적으로 누가 될지는 향후 2~3년 내에 결정될 가능성이 높다. 인터넷 시대의 구글과 아마존, 모바일 시대의 메타와 트위터 등과 같이 생성 AI 시대의 주인공이 누가 될지 계속 찾아 나갈 필요가 있지만 그와 함께 글로벌 AI 업체에 꾸준히 분산 투자할 수 있는 ETF에 투자하는 것도 매력적인 투자 대안이다.

Top Pick: Global X Artificial Intelligence & Technology (AIQ US)

표 1. Top Picks 및 관심 종목

(십억원, 배, %)

종목명	티커	시가총액	매출액	영업이익	당기순이익	PER	PBR	ROE
엔비디아	NVDA US	1,463,392	43,184	41,677	14,155	101.3	39.1	25.2
마이크로소프트	MSFT US	3,193,471	288,474	120,504	98,503	35.0	11.5	47.1
테슬라	TSLA US	1,064,803	133	15	14	83.7	15.5	18.5
유니티	U US	15,120	2,940	414	-1,108	-	3.46	-
어도비	ADBE US	310,450	26,380	11,931	7,330	32.2	13.9	41.7
NAVER	035420 KS	33,056	9,764	1,513	955	34.6	1.3	4.1
SK텔레콤	017670 KS	11,379	17,588	1,749	1,090	10.4	0.9	9.3
삼성SDS	018260 KS	10,461	13,527	756	731	14.3	1.2	8.6

주. 밸류에이션은 23년 실적 기준. 당사 커버리지 종목은 당사 추정치, 나머지는 블룸버그 컨센서스
 자료: 블룸버그, 미래에셋증권 리서치센터

표 2. 주요 AI ETF 현황

(%, 백만달러)

	Global X Artificial Intelligence & Technology		Global X Cybersecurity ETF	
수수료율	0.68		0.5	
개시일	2018-05-11		2019-10-25	
순자산	570.37		604.73	
주요 보유 주식	종목명	비중	종목명	비중
	알파벳	3.5	지스케일러	8.9
	인텔	3.3	클라우드 스트라이크	7.0
	메타	3.3	팔로알토 네트워크	6.5
	시스코	3.3	체크포인트 소프트웨어	6.0
	IBM	3.2	옥타	5.7
	엔비디아	3.2	퀄리스	5.4
	어도비	3.2	다크트레이스	5.4
	아마존	3.2	바로니스 시스템즈	5.2
	액센추어	3.1	사이버 아크 소프트웨어	5.1
테슬라	3.1	테너블	5.0	

자료: 글로벌 X, 미래에셋증권 리서치센터

II. 주요 산업 트렌드와 시사점

1. 빠른 기술 확산의 속도: 비즈니스 전략의 중요성 확대

대형언어모델은 대량의 데이터를 대규모 모델로 학습하기 때문에 학습과 파인튜닝(특정 업무를 수행하기 위해 관련 데이터를 추가 학습하는 과정), 추론(학습된 모델을 사용하는 과정)에 많은 시간과 비용이 든다. 그래서 챗GPT 같은 모델을 구축하기도 어렵고 이를 기반으로 다양한 서비스를 하려고 하면 서버 부족 등 한계가 있었다.

그러나 최근 빅테크와 스타트업을 중심으로 다수의 대형언어모델이 개발되고 있다. 유사한 데이터셋과 아키텍처를 사용하는 이들 모델들은 정확도 측면에서 'GPT 4'만큼은 아니지만 상용화 가능한 수준이고 안전성, 속도 등에서는 더 뛰어난 측면도 있다. 마이크로소프트 등 클라우드 업체들도 다양한 모델을 지원해 사용자가 원하는 모델을 선택해서 사용할 수 있게 제공하고 있다.

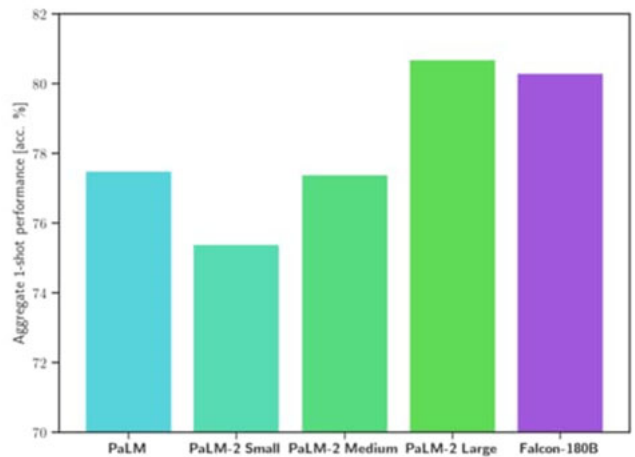
또한 다양한 오픈소스 모델이 등장하고 있으며, 특히 메타의 오픈소스 전략에 따라 'GPT 3.5' 급의 성능을 가진 모델을 구하기 쉬워지고 있다. 메타에서 발표한 'LLaMA 2'는 GPT 3.5 급의 정확도를 가진 것으로 평가되는데, 이를 기반으로 다양한 주체들이 개발에 참여하여 성능 및 보안 등의 기능을 발전시키고 있다. 최근에는 매개변수가 1800억개나 되는 Falcon-180B (UAE 기술혁신협회에서 개발)이 오픈 소스로 공개되어 커뮤니티에서 가장 높은 점수를 받기도 했다.

그림 1. 다양한 언어 모델의 등장

Model	Provider	Open-Source	Speed	Quality	Params	FINE-TUNEABILITY
gpt-4	OpenAI	No	☆☆☆	★★★★	-	No
gpt-3.5-turbo	OpenAI	No	☆☆☆	★★★★	175B	No
gpt-3	OpenAI	No	☆☆☆	★★★★	175B	No
ada, babbage, curie	OpenAI	No	☆☆☆	☆☆☆☆	350M - 7B	Yes
claude	Anthropic	Yes	☆☆☆	★★★★	52B	NO
claude-instant	Anthropic	Yes	☆☆☆	☆☆☆☆	52B	No
command-xlarge	Cohere	No	☆☆☆	☆☆☆☆	50B	Yes
command-medium	Cohere	No	☆☆☆	☆☆☆☆	6B	Yes
BERT	Google	Yes	☆☆☆	☆☆☆☆	345M	Yes
T5	Google	Yes	☆☆☆	☆☆☆☆	11B	Yes
PaLM	Google	Yes	☆☆☆	☆☆☆☆	540B	Yes
LLaMA	Meta AI	Yes	☆☆☆	☆☆☆☆	65B	Yes
CTRL	Salesforce	Yes	☆☆☆	☆☆☆☆	1.6B	Yes
Dolly 2.0	Databricks	Yes	☆☆☆	☆☆☆☆	12B	Yes

자료: dev.to, 미래에셋증권 리서치센터

그림 2. 높은 점수를 받은 오픈 소스 모델 Falcon-180B



자료: 허깅 페이스, 미래에셋증권 리서치센터

그림 3. 언어 모델의 진화

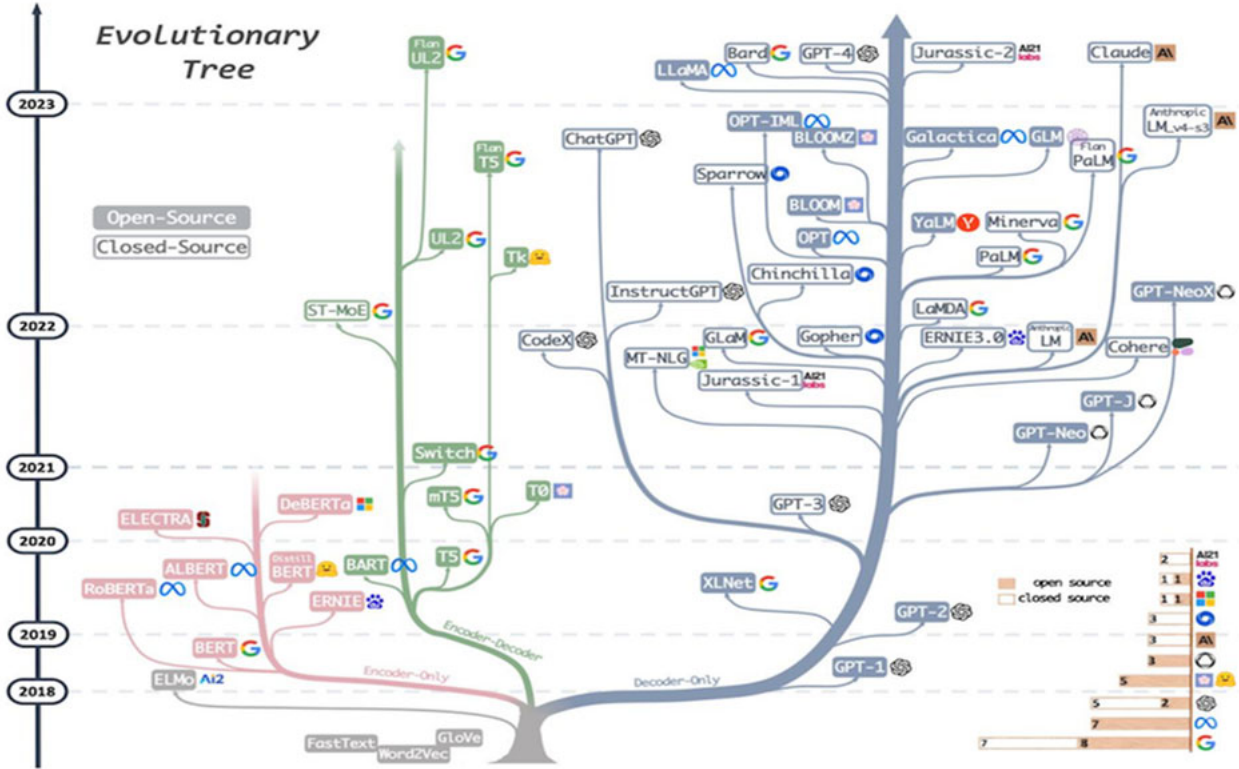


Image from Mooler0410/LLMsPracticalGuide

Yang, J., Jin, H., Tang, R., Han, X., Feng, Q., Jiang, H., ... Hu, X. (2023). Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. arXiv [Cs.CL]. Retrieved from <http://arxiv.org/abs/2304.13712>

자료: A Survey on Chat GPT and beyond, 미래에셋증권 리서치센터

그림 4. 메타의 오픈 소스 모델 LLaMA와 그 파생 모델

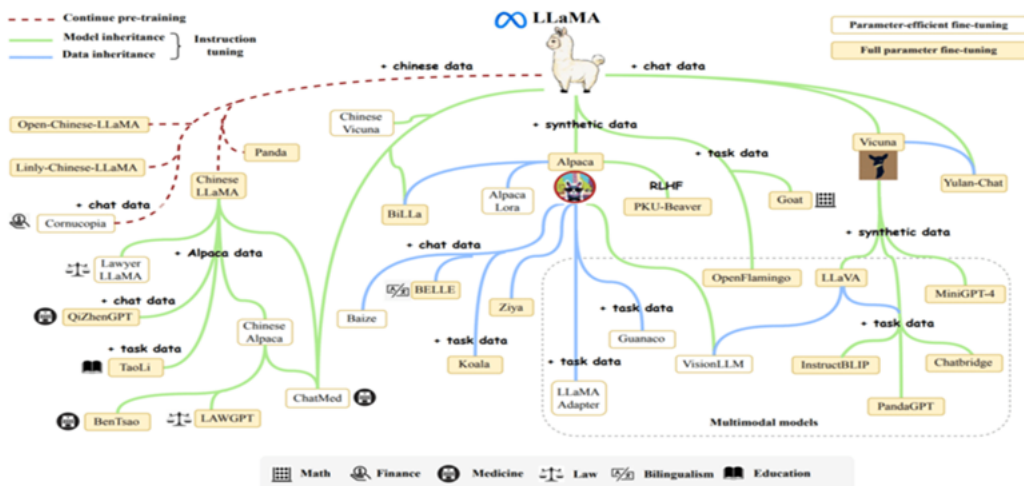


Fig. 4: An evolutionary graph of the research work conducted on LLaMA. Due to the huge number, we cannot include all the LLaMA variants in this figure, even much excellent work. To support incremental update, we share the source file of this figure, and welcome the readers to include the desired models by submitting the pull requests on our GitHub page.

자료: 깃허브, deci, 미래에셋증권 리서치센터

특히 데이터 중심의 접근법이 확산되면서 추론에서 연산 부담이 줄어들고 있다. 메타의 'LLaMA 2'는 'GPT 3.5' 대비 매개 변수가 작지만 학습 데이터의 양을 크게 늘려 유사한 정확도를 구현했다. 오픈 AI 진영에서도 'GPT 3.5' 파인튜닝을 통해 특정 영역에서 매개 변수가 10배 이상인 'GPT 4' 급 성능을 구현하는 기법을 공개했다. 매개 변수가 적으면 추론 연산 부담이 줄어들어 상용화에 유리하다.

이와 함께 특정 분야에 적용 가능한 소형 언어모델도 확산되고 있다. 소형 언어모델은 매개 변수를 줄이는 대신 질 좋은 데이터를 학습시키고 파인 튜닝하는 방식으로 성능을 개선시키는데 특정 분야에 한정되어 있지만 훈련 비용과 시간을 줄이고 정확도를 높일 수 있다. 신속하고 저렴하게 만들 수 있는 데다 보유한 데이터를 활용해 맞춤형으로 구축할 수 있기 때문에 기업에서 도입하기에도 유리하다. 다양한 소형 언어모델의 등장으로 AI 기술을 상용화하기가 좀더 용이해졌다.

표 3. 오픈 AI의 GPT와 메타의 LLaMA 비교

	매개변수(십억개)	학습 데이터	컨텍스트 길이	정확도(%)	모달리티
GPT 3.5 turbo	154-175	570GB	4096-16384	70	텍스트
GPT 4	1000-1760	NA	8192-32768	86	텍스트, 이미지
LLaMA 2	7-70	2조 토큰	4096	69	텍스트

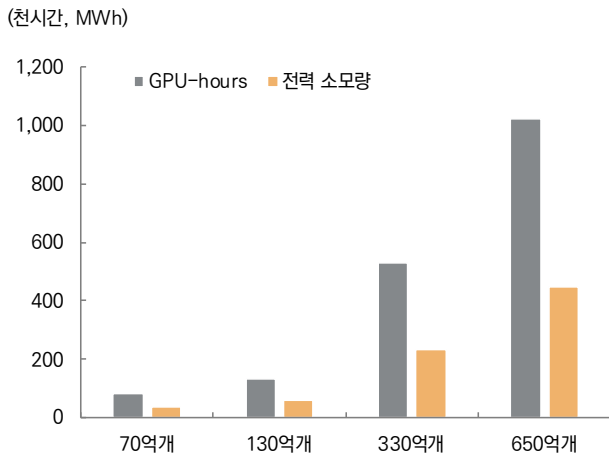
주: 정확도는 Massive Multitask Language Understanding (MMLU), GPT 4 매개변수는 추정
 자료: neoteric, 미래에셋증권 리서치센터

표 4. 소형언어모델 사례

개발자 및 모델	특징
스탠퍼드대학교의 알파카	매개변수 70억 개, 52,000개 데이터를 기반으로 클라우드 컴퓨터 8대로 이용해 3시간 만에 훈련. 비용은 600달러
마이크로소프트의 파이-1	매개변수 13억 개를 가진 새로운 코딩 특화 언어모델, 100배 크기인 GPT-3.5를 일부 능가했다는 연구 결과 발표
데이터브릭스의 돌리(Dolly) 2.0	100달러 비용으로 1대 서버에서 3시간 훈련해 구축한 매개변수 60억 개의 소형 언어모델
세레브라스	매개 변수 1억개~30억개의 소형 언어 모델 7개 오픈 소스로 공개
갯잇 AI의 엘마	클라우드를 이용하지 않는 온프레미스(사내구축)형 소형언어모델.
스테빌리티 AI의 스테이블LM	텍스트와 이미지, 코드를 생성할 수 있는 대화형 오픈소스 언어모델. 30억개/70억개 파라미터

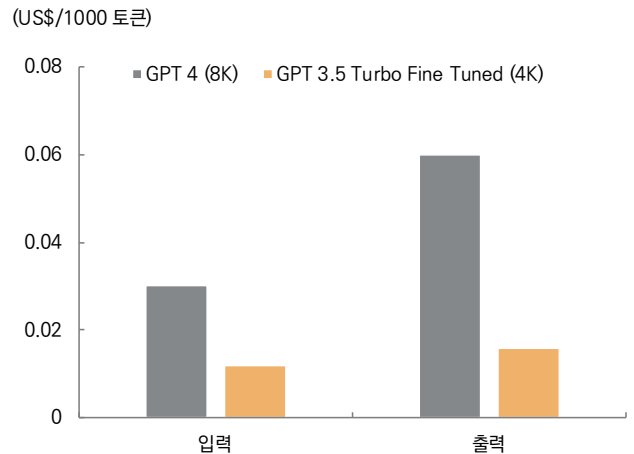
자료: 언론 자료, ICT 브리프에서 재인용, 미래에셋증권 리서치센터

그림 5. LLaMA 매개 변수에 따른 GPU 시간과 전력 소모량



자료: shaped.ai, 미래에셋증권 리서치센터

그림 6. GPT 3.5의 파인 튜닝 모델과 GPT 4 추론 비용 비교



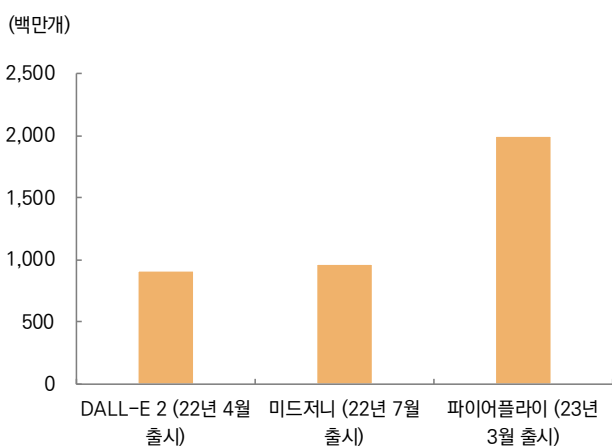
자료: 오픈 AI, 미래에셋증권 리서치센터

상용화 가능한 수준의 기술을 확보하기가 용이해지고 추론 비용이 하락함에 따라 기술 자체보다는 이를 적용해 제품화하는 기업의 비즈니스 전략이 점차 중요해질 전망이다. 얼마나 많은 고객군과 질 좋은 데이터를 확보하고 있는지, 어떻게 사용자 니즈를 만족시키고 빠른 제품화로 시장을 선점할 수 있는지, 바뀌는 기술과 시장 상황에 얼마나 민첩하게 대응할 수 있는지, 저작권 등 법적 이슈를 얼마나 잘 대응할 수 있는지 등이 관건이 될 전망이다.

예를 들어 성공적인 상용화 사례로 꼽히는 어도비의 파이어플라이는 출시 이후 20억장의 이미지를 생성해 당초 예상을 크게 뛰어넘었다. 텍스트를 이미지로 바꾸어주는 기술은 이미 상용화되어 있었다. 그러나 파이어플라이는 기존 포토샵 등 이미 방대한 고객군을 가진 제품에 통합했고 사용자가 텍스트 프롬프트만 넣으면 편리하게 사용할 수 있도록 했으며 라이선스를 확보한 이미지로만 학습(2.7억개 이상의 이미지를 보유하고 있는 Adobe Stock 등을 이용)하고 저작권 문제가 발생할 경우 보상해 준다는 정책을 제시해 사용자가 저작권 문제로부터 벗어날 수 있게 하였다.

24년에 주요 기업들의 생성 AI 제품이 쏟아질 것으로 예상된다. 마이크로소프트 MS 코파일럿처럼 기존 제품에 생성 AI 기능을 통합해 고객을 확보하거나 판매 가격을 인상시키는 전략부터 개인, 기업용 AI 어시스턴트 도입을 통한 생산성 개선 및 소비자 효용 증대, 원가 절감, 콘텐츠 생성 플랫폼 제공 등 다양한 전략이 준비되고 있다. 이들 전략의 설득력에 따라 기업 가치의 차별화가 예상된다.

그림 7. AI 이미지 생성 개수(23년 8월 기준)



자료: 언론 자료, 어도비, 미래에셋증권 리서치센터

그림 8. 스테이블 디퓨전과 차별화되는 파이어플라이의 이점

Adobe Firefly만의 차별화된 이점

- Firefly는 브라우저 기반의 인터페이스에 입력된 간단한 텍스트 프롬프트를 바탕으로 예술적이고 사실적인 이미지를 생성하며, 크리에이티브 워크플로우를 가속화합니다.
- Firefly가 이미지를 생성하면 프롬프트에 설명을 더 추가하거나 스타일, 테마, 조명, 구도 옵션을 사용하여 결과를 조정할 수 있습니다.
- Firefly는 Adobe Stock, 공개 라이선스 콘텐츠, 저작권이 만료된 공개 도메인 콘텐츠를 학습하므로 상업적으로 안전합니다.

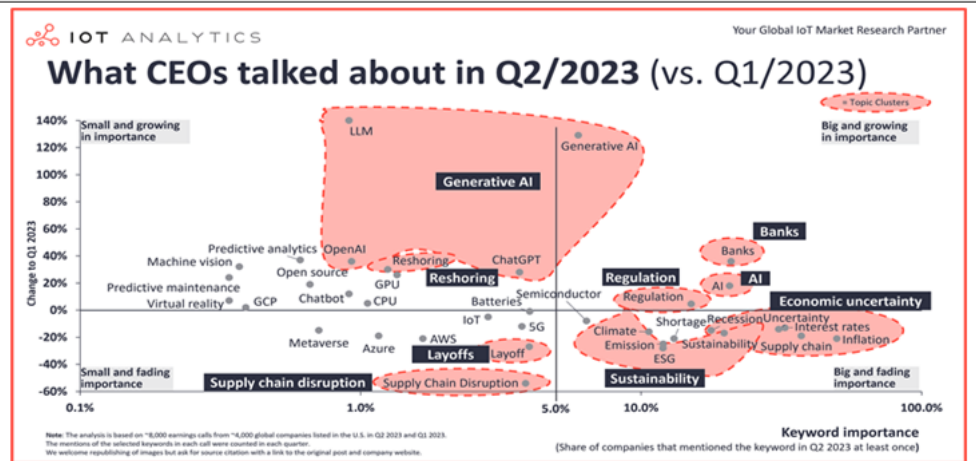
자료: 어도비, 미래에셋증권 리서치센터

2. 예상을 상회할 응용 서비스의 성장: 추론 중요성 확대 예상

상용화 가능한 수준의 기술을 확보하기가 용이해진 가운데 주요 기업들의 경쟁적인 제품화로 생성 AI 어플리케이션은 급격히 성장할 것으로 예상된다. 관련해서 추론 비용을 줄일 수 있는 소프트웨어와 전용 반도체의 중요성이 부각될 전망이다.

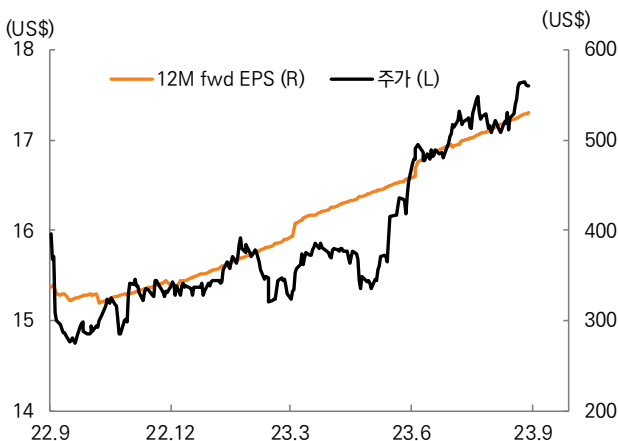
챗GPT 이후 글로벌 주요 기업들은 앞다퉀 생성 AI 전략을 발표하고 있다. 이는 생성 AI 기술을 도입해 판매 가격을 인상하거나 점유율을 확대할 수 있기 때문이기도 하지만, 도입하지 않으면 경쟁에서 뒤쳐질 수 있기 때문이다. 주가 흐름도 이를 반영하고 있다. 포토샵에 생성 AI를 도입한 신제품 파이어 플라이가 좋은 반응을 얻자 어도비의 실적 추정치와 주가는 상승했다. 언어학습 플랫폼 업체인 듀오링고는 역할극 등의 기능을 추가하면서 기존 제품 가격을 최대 50% 인상하겠다고 밝히면서 수혜주로 부각되었지만 과제 도움 서비스 업체인 체그는 챗GPT로 인해 사용자들의 관심이 줄었다는 언급에 주가가 폭락하였다.

그림 9. 23년 2분기에 CEO의 언급(전분기대비)



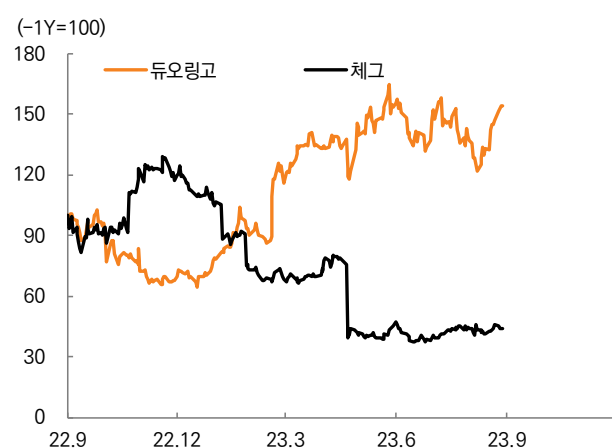
자료: IoT, 미래에셋증권 리서치센터

그림 10. 어도비 주가와 실적 추정치 추이



자료: 블룸버그, 미래에셋증권 리서치센터

그림 11. 미국 교육 업체 듀오링고 Vs 체그 주가 추이



자료: 블룸버그, 미래에셋증권 리서치센터

표 5. 주요 기업의 생성 AI 기술 상용화 계획

회사명	상용화 현황 및 계획
어도비	포토샵 등 기존 제품에 생성 AI 기능을 통합한 'Firefly' 출시 방대한 고객 기반, 저작권 이슈 해결 등으로 Firefly는 3월 출시되어 9월까지 베타 서비스만으로 20억장 이상의 이미지 생성 가격은 기존 어도비 크리에이티브 클라우드(CC) 구독자들은 일정량을 무료로 생성할 수 있고 이후에는 건당 5센트 25년 말 크리에이티브 클라우드/익스프레스 구독자 수는 4500만명에 달할 것으로 추정. 이중 일부가 Firefly 유료 구독자로 전환 시 본격적 수익화 가능
듀오링고	언어학습 플랫폼을 제공하는 듀오링고는 기존 제품에 GPT-4를 결합한 '듀오링고 맥스' 출시 시가 답변을 설명해주는 기능과 역할극을 할 수 있는 기능 추가한 듀오링고 맥스 가격은 기존 대비 최대 50% 인상한 월 30달러로 책정 듀오링고는 7,400만명의 월간활성사용자수를 보유하고 있고 이중 7.9%인 520만명이 유료 구독자
인튜이트	세무신고 소프트웨어, 중소기업을 위한 회계 소프트웨어 등을 제공하는 인튜이트는 23년 9월 생성 AI 비서인 '인튜이트 어시스트' 발표 인튜이트 어시스트는 기존 프로그램에 통합되어 고객들에게 현금 부족 여부, 이메일 마케팅 등 개인화된 서비스를 제공할 계획 가격 정책 등은 아직 미정. 인튜이트의 고객은 전세계적으로 1억명 이상으로 추산
줌	24년 초 AI 비서 '줌 컴패니언' 출시 예정. 미팅 전 프로젝트 현황을 문의할 수 있고 미팅 중에도 답변 초안 작성, 회의 후 미팅 요약 등 기능 제공 줌 컴패니언은 기존 유료 사용자에게 추가로 제공. 줌의 일 미팅 참여자 수는 23년 기준 3억명이 넘는 것으로 추산
워크데이	인사관리 소프트웨어를 제공하는 워크데이는 당사 인사관리 플랫폼에 문서 이해, 언어 생성, 요약 등 다양한 기능을 향후 6~12개월 내 추가할 계획 워크데이는 6000만명 이상의 사용자들이 매년 4420억건의 트랜잭션을 실행하고 있음
세일즈포스	오픈 AI와 제휴해 기존 서비스에 생성 AI 통합한 '아인슈타인 AI' 발표. 영업, 서비스, 마케팅 등 전 업무 시스템에 생성 AI 기술을 활용 9월 대화형 어시스턴트인 아인슈타인 코파일럿과 코파일럿 빌더 공개. 사용자가 자연어로 질문하면 업무에 필요한 정보와 콘텐츠를 자동으로 생성 세일즈포스 고객 업체는 23년 기준 15만개 이상으로 추산
서비스나우	기존 제품에 생성 AI 기술을 도입한 'Pro+' 출시 계획. Text to Code(자연어로 말하면 소프트웨어 코드를 제공)와 사례 요약(즉 IT/HR, 고객 서비스 등에서 고객 또는 사건의 정보, 이전 기록, 해결 방안 등을 읽고 요약) 등의 기능 제공 Pro+는 기존 대비 최대 60% 높은 가격으로 출시할 계획
유니티	크리에이터의 생산성 향상과 실시간 3D 경험의 몰입도를 높여줄 신규 AI 플랫폼 '유니티 센티스'(Unity Sentsis)와 '유니티 뮤즈'(Unity Muse)를 출시 유니티 센티스는 AI 모델을 유니티 런타임에 임베드해 지연 시간 없이 여러 기기에서 프로그램을 실행할 수 있게 해줌. 유니티 뮤즈는 크리에이터가 게임, 애플리케이션, 디지털 트윈을 위한 실시간 3D 콘텐츠를 제작하는 방식을 가속화하는 AI 솔루션
스냅	3월 챗GPT로 구동되는 챗봇 마이 AI 출시. 초반에는 유료였으나 4월부터 무료로 공개. 광고 등에 활용 부절적인 콘텐츠 생성 등으로 일부 논란이 있으나 2월 말 출시된 이후 6월까지 1.5억명(월간 사용자의 20%)이 100억건의 메시지를 챗봇에게 보냄
마이크로소프트	파워 포인트 등이 포함된 MS 365에 생성 AI 기능을 도입한 MS 코파일럿 출시 계획(24년 예상). 저작권에 대해 법적 책임 제공 월 사용료는 30달러로 기존 월 9~16달러 대비 약 50~80% 상승. MS 365 기업용 버전 사용자수는 약 3.8억명으로 추산 기존 검색엔진에 챗GPT를 도입한 채팅형 검색 Bing 출시. 7월에는 보안을 강화해 기업용으로 사용할 수 있는 Bing 엔터프라이즈도 출시
구글	MS 코파일럿에 대응해 구글 워크스페이스에 생성 AI 기능을 도입한 듀엣 AI 출시. 24년 초 일반 이용자로 출시 예정 월 사용료는 30달러로 코파일럿과 동일하게 책정. 구글 워크 스페이스의 유료 가입자는 600만명으로 추산 PaLM2 기반으로 만들고 구글 검색과 연동한 챗봇 서비스 바드 출시 기업이 대규모 언어 모델(LLM)을 활용할 수 있는 플랫폼 '버텍스 AI(Vertex AI)' 강화. 100개 이상의 모델을 구동 가능
메타	9월 Connect 행사에서 제품화 관련 발표 예상 주커버그의 인터뷰에 따르면 AI '페르소나' 형태의 챗봇을 페이스북, 인스타그램 등 자사 서비스에 선보여 앱 이용률과 참여도를 높일 것으로 예상 페르소나는 다양한 캐릭터를 구축해 원하는 콘텐츠를 일대일로 추천해줄 예정. 또한 크리에이터의 작업 효율성을 높여주고 기업체와의 협상을 용이하게 해 줌. 광고 관련 생성 AI 기능도 도입 예상

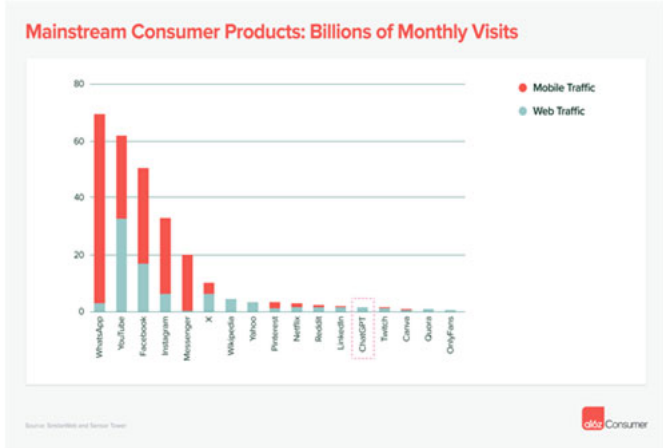
자료: 미래에셋증권 리서치센터

실제 사용자가 자연어로 물어보면 답변해주는 AI 어시스턴트와 같은 기능은 대부분의 업체들이 출시할 계획이다. 특히 주요 기업의 출현으로 생성 AI 적용 범위가 획기적으로 넓어질 전망이다. 주요 기업은 스타트업과 달리 1) 고객 기반이 방대하며, 2) 편리한 이용 환경을 제공할 수 있고, 3) 저작권 등의 법적 이슈를 보완할 수 있다는 점에 기인한다.

또한 생성 AI 기술 특성상 재생산의 비용이 제로에 가깝게 하락하기 때문에 급격한 물량 증가가 예상된다. 과거 인터넷/모바일 혁명으로 비디오 등 콘텐츠 제조 비용이 급감했듯 생성 AI 기술의 도입으로 콘텐츠 제조 비용은 급감하고 이를 생성해 내는 플랫폼의 가치가 확대될 전망이다.

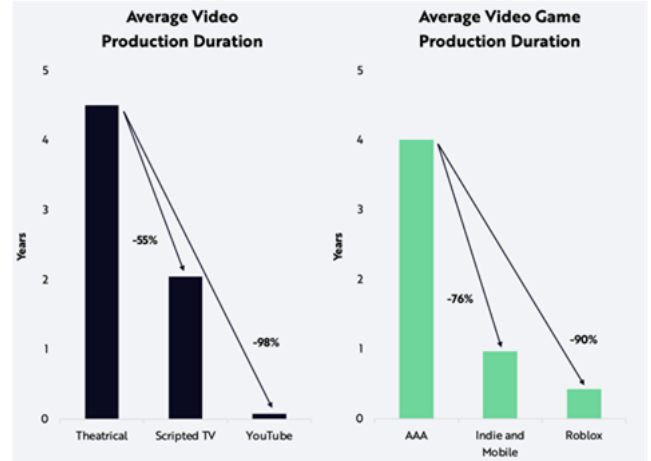
어플리케이션의 확대와 함께 추론의 효율성이 점차 중요해질 전망이다. 현재 학습용 반도체 시장을 장악하고 있는 엔비디아가 추론 시장에서도 경쟁 우위를 지켜낼지 주목할 필요가 있다.

그림 12. 주요 어플리케이션 월 방문자수



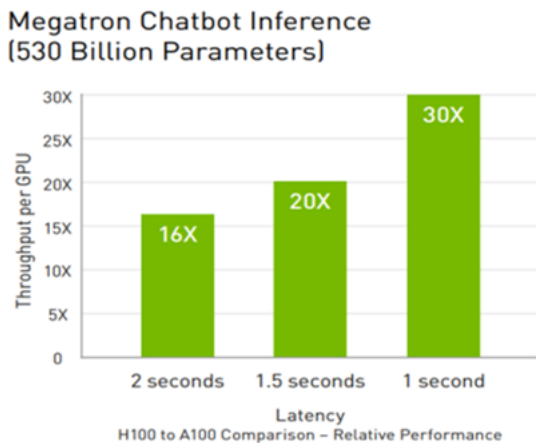
자료: 미래에셋증권 리서치센터

그림 13. 평균 비디오/비디오 게임 제조 기간



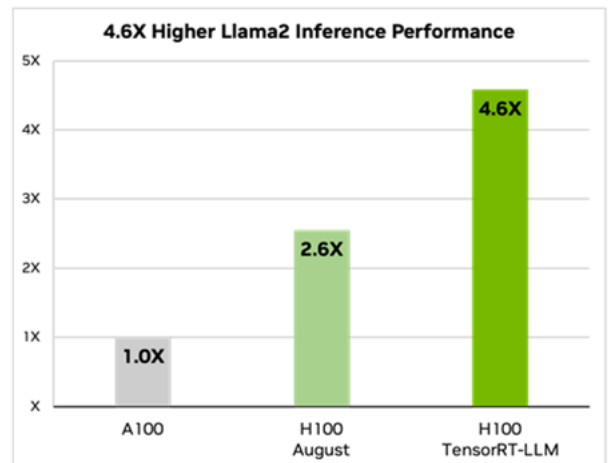
자료: 아크 인베스트, 미래에셋증권 리서치센터

그림 14. 챗봇 추론에서 A100 대비 H100 성능 개선



자료: 엔비디아, 미래에셋증권 리서치센터

그림 15. 엔비디아, 추론 성능 개선 소프트웨어 개발



자료: 엔비디아, 미래에셋증권 리서치센터

3. 빠르게 발전하는 기술: 시장 규모와 경쟁 구도의 급변 가능성

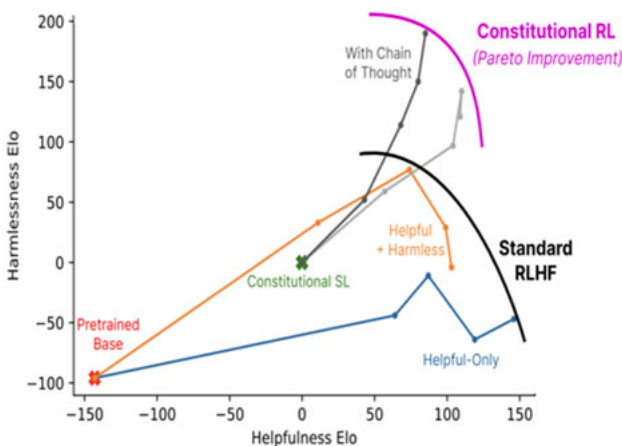
인터넷/모바일 기술과 차별화되는 생성 AI 기술의 특징은 기반 기술 자체가 빠르게 발전 중이라는 점이다. 생성 AI 기술은 답변의 정확도/신뢰도 개선, 멀티 모달 모델로의 발전, 도구 사용, Agent간 상호 작용을 통한 발전 등의 방향으로 이루어지고 있으며 그 진전 정도에 따라 AI가 할 수 있는 일의 범위와 신뢰도, 선발 업체의 경쟁 우위 정도 등이 달라질 수 있다.

생성 AI는 기본적으로 다음에 나올 확률이 가장 높은 단어를 내 놓기 때문에 답변의 정확도가 떨어질 수 있다. 이에 현재는 상대적으로 정확도의 요구가 낮은 콘텐츠 생성, 흥미 위주의 대화, 브레인 스토밍 등의 용도에 사용되고 있지만, 정확도가 일정 수준을 넘어서면 부가가치가 높은 전문적인 분야에서도 사용할 수 있게 된다.

이를 위해 오픈 AI에서는 사람이 직접 피드백을 주고 이를 통해 모델을 강화학습시키는 RLHF를 통해 정확도를 개선시켰고, 최근 오픈 AI 출신이 설립한 스타트업인 엔트로픽은 사람의 피드백을 AI의 피드백으로 대체하는 RLAIIF이라는 기술로 모델의 안전성을 획기적으로 개선시킬 계획이다.

학습하는 데이터의 양을 늘리거나 원하는 영역에 대해 추가 학습하는 것도 정확도를 개선시킬 수 있는 방법이다. 입력/출력하는 토큰(단어)수를 늘림으로써 AI가 맥락을 좀 더 잘 이해하고 사용자가 원하는 답변을 내 놓을 수 있도록 모델을 조정할 수도 있다. 모델이 사용자에게 대한 데이터를 축적할 수 있도록 벡터 데이터베이스와 같은 기능을 통해 메모리 기능을 개선시키는 것도 답변의 완성도를 높이는 방법이다.

그림 16. 엔트로픽, 사람이 아닌 AI 피드백에 기반한 강화학습



자료: 엔트로픽, 미래에셋증권 리서치센터

그림 17. 차세대 언어 모델의 발전 방향

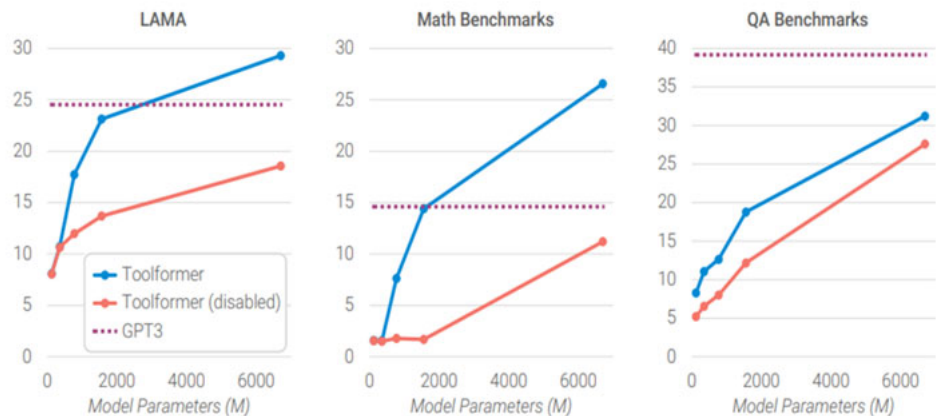


자료: Appypie, 미래에셋증권 리서치센터

AI가 도구를 사용하는 기술(툴포머, Toolformer)도 발전하고 있는데, 외부 데이터 베이스를 활용해 성능을 개선시키고 필요한 업무를 대신 수행할 수 있다. 오픈 AI의 플러그인 기능이 대표적인데, 챗GPT를 주요 어플리케이션에 연결해 해당 영역의 데이터를 활용하고 업무를 수행한다. 플러그인 중 하나인 코드 인터프리터는 AI가 직접 코딩을 하고 해당 코드를 실행까지 할 수 있어 데이터 분석 등 다양한 기능을 수행할 수 있다.

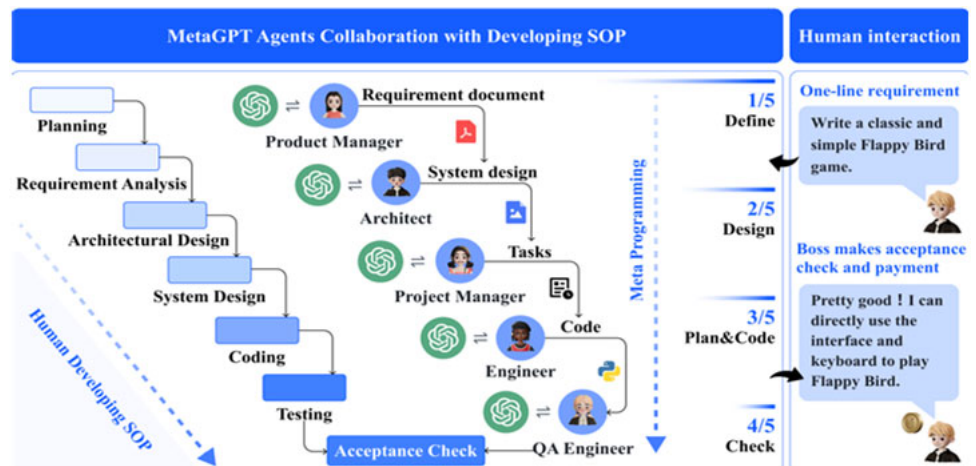
생성적 에이전트(Generative Agents) 기술도 주목할 필요가 있다. 언어 모델로 다양한 AI Agent를 만든 뒤 이들끼리 서로 상호 작용하면서 복잡한 업무를 수행하는 기술이다. 23년 8월 공개된 ‘메타 GPT’는 제품 매니저, 개발자, 리뷰어 등 7개의 에이전트를 만들어 게임 소프트웨어를 개발하도록 했는데 이들 Agent들은 각자의 역할을 기반으로 서로에서 조언도 하고 피드백도 주면서 업무를 진행시켜 나갔다. 복잡한 업무는 사람이 단계별로 지시하거나 직접 가이드를 줘야 하는데 메타 GPT에서는 AI Agent가 이를 대신하기 때문에 개발 속도가 훨씬 빨라질 수 있을 뿐 아니라 그 성과도 크게 개선될 잠재력이 있다.

그림 18. 툴포머: 도구를 이용한 성능 개선



자료: arxiv.org/pdf/2302.04761, 미래에셋증권 리서치센터

그림 19. 생성 에이전트를 이용한 메타 GPT



자료: <https://arxiv.org/abs/2308.00352>, 미래에셋증권 리서치센터

4. 어플리케이션: AI 비서, 사용자 트래픽에 주목

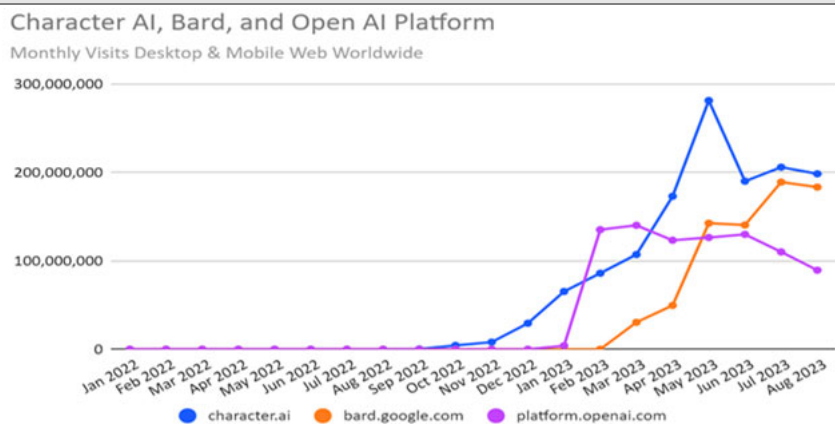
어떤 서비스가 킬러 어플리케이션이 되고 누가 그 시장을 장악할 지는 향후 2~3년 내에 결정될 것으로 예상된다. 현재 몇 가지 시그널은 나타나고 있는데 완전히 개인화된 AI 비서가 새로운 플랫폼으로 부상할 가능성이 있고, 단순한 생산성 개선 서비스는 빠르게 커머디티화되고 ‘재미’ 등 차별적인 효용으로 고객을 확보하고 플랫폼을 구축하는 것이 중요할 것으로 예상된다. 이는 향후 사용자 트래픽을 보면서 확인해 나갈 필요가 있다. 나아가 메타버스, 로보틱스 등 유관 산업의 잠재력도 높아 보인다.

빌 게이츠는 최근 컨퍼런스에서 AI 비서 시장을 장악하는 업체가 산업의 승자가 될 것으로 전망한 바 있다. 모델의 메모리 기능이 강화되어 사용자에게 대한 정보를 축적하게 되면 완전히 개인화된 검색, 쇼핑 등 최적화된 서비스를 제공할 수 있기 때문이다. 도구를 사용하는 툴포머 기능도 발전하면 원하는 업무를 수행해 줄 수도 있다. 나아가 최근에는 AI가 단순히 정확한 답변을 주는 것뿐 아니라 친구처럼 공감하고 조언을 해줄 수 있는 기능도 발전하고 있다.

한편 단순한 생산성 개선 서비스는 차별화가 쉽지 않을 전망이다. 글쓰기 서비스를 제공하는 스타트업 재스퍼 AI나 웹사이트 개인화를 지원하는 Mutiny AI은 최근 구조조정을 시작했다. 반면 개인화된 챗봇 서비스를 제공하는 캐릭터 AI는 최근 월간 방문자 수가 오픈 AI를 넘어설 정도로 인기를 끌고 있다. 챗GPT와 차별화되는 캐릭터 AI의 특징은 사용자가 답변에 대한 피드백을 통해 캐릭터를 창조해내고 여기서 재미를 느끼는 젊은 사용자가 증가하고 있다는 점이다.

중기적으로 메타버스, 로보틱스 등 유관 산업의 성장도 기대된다. 생성 AI 기수로 인해 콘텐츠 수가 늘어나면 메타버스의 효용이 증대될 수 있고 언어 모델이 발전할수록 로봇의 인지 및 판단 능력이 개선되면서 할 수 있는 일의 범위와 성능이 획기적으로 발전할 수 있기 때문이다.

그림 20. 주요 플랫폼별 월간 방문자 수 추이



자료: Similarweb, 미래에셋증권 리서치센터

III. 반도체

1. AI 추론 시장의 성장성

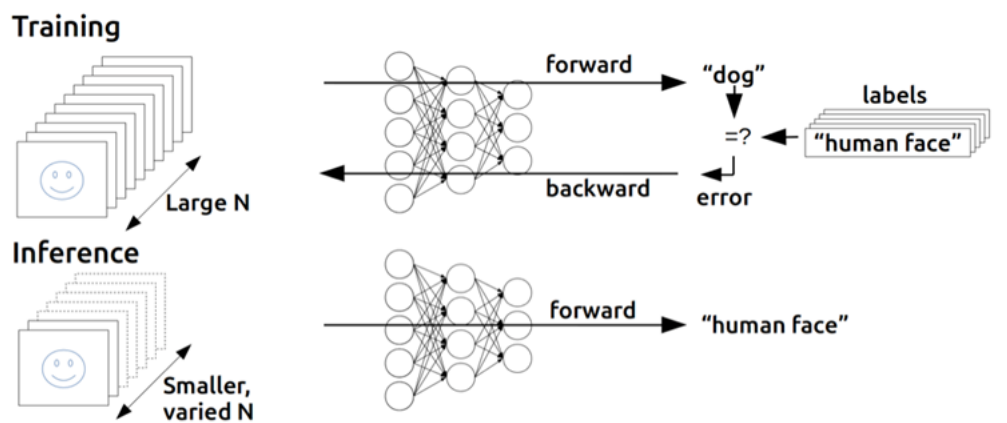
1) 학습(Training)과 추론(Inference)의 계산방식 차이

당 보고서의 독자들은 이미 익숙할 수 있겠으나, AI 모델의 학습(Model Training)과 추론(Inference)의 계산 과정을 간략히 비교해 보고자 한다.

학습은 AI 모델이 주어진 데이터와 그 결과값을 이용해 여러가지 변수들 가운데 최적의 가중치(파라미터)를 찾는 과정이다. 개와 고양이를 판별하는 AI 모델을 예로 들 수 있다. 수만장의 고양이와 개의 사진에 그에 해당하는 결과값을 부여하여 모델의 가중치를 반복적으로 조절하게 된다. 학습 모델은 각 사진마다 개인지 고양이인지 예측하는 시도를 하고, 그 예측과 사전 부여한 결과값(라벨)을 비교한다. 결과가 틀렸다면 오차를 바탕으로 가중치를 업데이트한다. 이와 같은 계산을 여러 번 반복하며 모델은 점점 더 정확한 예측을 할 수 있도록 개선된다. 이 과정에서 가중치값(파라미터)의 개수를 늘릴 수록 모델의 정확도가 향상되는데, GPT-3 모델의 파라미터는 1,750억개, GPT-4 모델의 파라미터는 1.7조개로 알려져 있다.

추론은 기 학습된 모델을 사용하여 새로운 입력 데이터에 대한 결과값을 예측하는 과정이다. 상기 모델의 학습이 완료된 후 새로운 사진을 제시했을 때 그 사진이 고양이인지 개인지 판별하는 것이 일종의 추론 과정이다. 학습 과정에서 수많은 연산을 통해 최적화된 가중치는 고정되어 있기 때문에, 추론에서는 새로운 사진을 입력으로 받아 예측값을 빠르게 출력하는 데 필요한 연산만 수행하게 된다.

그림 21. 학습(Training) vs. 추론(Inference)



자료: Google image, 미래에셋증권 리서치센터

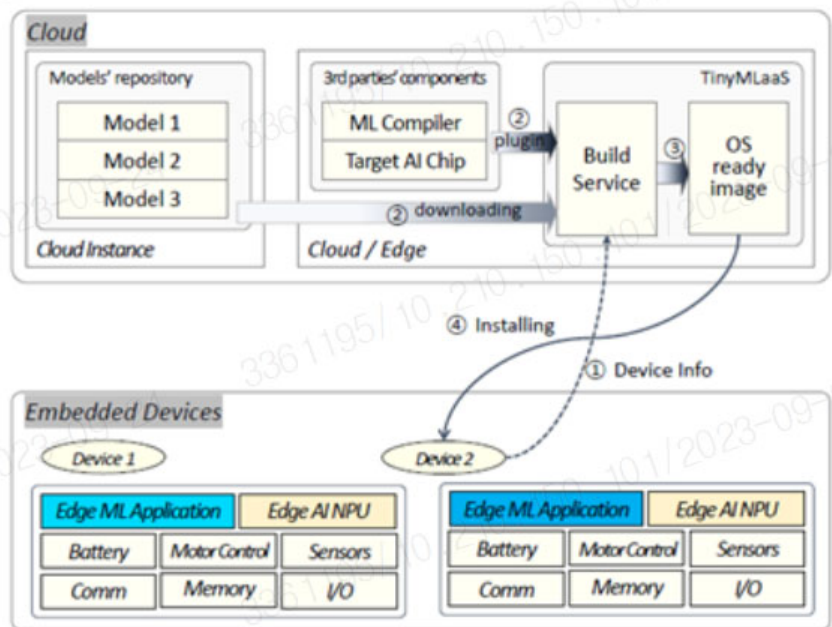
2) 추론시장의 성장성이 큰 이유

AI 데이터 학습은 새로운 모델을 만들거나 기존 모델을 업데이트할 때 필요한 과정이다. 이는 대규모의 데이터와 하드웨어 리소스가 필요하며, 주로 데이터 센터와 같은 대형 인프라에서 이루어진다. 학습모델 시장 성장과 동시에 추론모델 시장 또한 빠르게 성장하고 있다. 추론시장은 응용분야와 적용기기 측면에서 각각 살펴볼 수 있다.

응용분야의 측면에서는 스마트폰 앱에서의 얼굴 및 음성 인식, 자율주행 차량의 주행 경로 예측 등 다양한 서비스 영역에서 AI 추론 기능이 활용되고 있다. 추론 기반의 서비스는 기 학습된 모델을 기반으로 수많은 응용분야에 적용될 수 있기 때문에 확장성과 다양성 측면에서 잠재력이 크다.

적용기기 측면에서는 스마트폰, IoT 디바이스 등과 같은 엣지 디바이스에서 AI 추론을 직접 수행하는 비중이 증가하고 있다. 데이터를 클라우드로 전송하여 처리하고 다시 피드백을 받는 대신, 개별 디바이스 내에서 직접 추론 연산을 수행함으로써 시간 지연을 최소화하고 데이터 보안을 지킬 수 있다. 모델 최적화 기법들이 발전하면서, 성능을 유지하면서 모델의 크기를 줄이거나, 추론 속도를 높이는 연구들이 활발히 진행중이다.

그림 22. 엣지 AI 컴퓨팅 구조 및 구현 메커니즘



자료: SKBY, 미래에셋증권 리서치센터

3) 학습과 추론의 반도체 측면에서의 차이

연산의 종류가 학습인지 추론인지의 여부에 따라 적용되는 반도체의 종류와 구조가 달리 적용된다. 프로세서의 측면에서는 학습 모델의 경우 일반 목적의 GPU(General Purpose GPU)를 주요 프로세서로 사용하는 경향이 크다. 모델의 종류가 다양하고 딥러닝의 방식이 지속적으로 진화하기 때문이다. 반면, 추론 연산은 특화된 전용 프로세서에서 수행될 수 있다. 모델 아키텍처가 상대적으로 고정되어 있기 때문이다.

한편, 메모리의 소요량 측면에서도 차이가 있다. 모델의 규모가 큰 학습 모델들은 많은 메모리를 필요로 한다. 학습의 대표적인 계산 과정인 역전파를 위해 중간 활성화를 저장해야 하기 때문이다. 반면, 추론 모델의 경우 중간 활성화는 오랫동안 저장할 필요가 없어 메모리 사용량은 학습 모델보다 일반적으로 낮다.

추론 작업을 위해서는, 최적화된 전용 프로세서인 TPU(Tensor Processing Units) 또는 FPGA도 사용될 수 있다. 특히, 엣지 디바이스는 저전력, 효율적인 추론을 위해 NPU(Neural Processing Units)와 같은 전용 칩을 사용할 수 있다. 다만, 추론용 프로세서의 경우 특정 모델 추론에 특화되어 있어 구조가 다른 모델을 제대로 실행할 수 없는 문제가 있다. 범용성이 떨어진다는 점이다.

그럼에도 불구하고, 추론용 AI 전용 반도체가 가진 강점은 엔비디아 AI 반도체(GPU)보다 우수한 비용 효율성과 적은 전력 소비량 등이다. AI 서비스를 위한 인프라 투자비용이 비싸지고 있고, 환경적 측면에서 데이터센터의 전력효율성이 점차 중요해지는 자금의 시기에 추론 분야에서의 Non NVIDIA 전용 반도체가 주목받을 수 있다.

시장조사업체 Gartner에 따르면 AI 반도체 시장 규모는 2021년 347억 달러(약 42조 원)에서 연평균 16%씩 급성장해 2026년 861억 달러에 이를 것으로 전망된다. 추론용 AI 반도체의 시장 수요가 절대적으로 많다. 시장조사업체 트랙티카는 오는 2025년 AI 반도체 시장에서 추론용 AI 반도체의 점유율이 전체 AI 반도체의 78% 점유율을 차지할 것으로 예측했다.

2. 주요 기업들의 AI 대응 전략

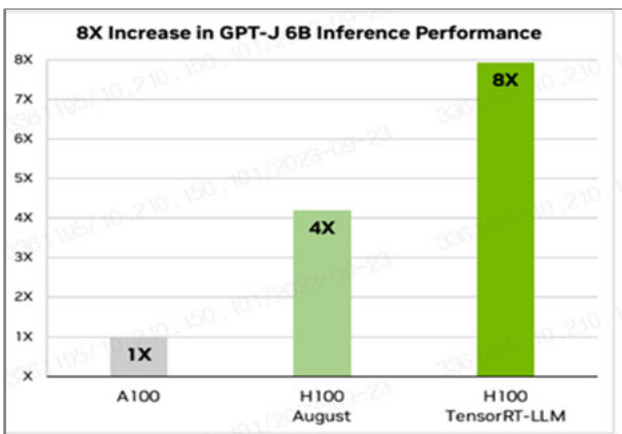
1) NVIDIA

Inference향 대표 제품
A100 Tensor Core GPU
H100 Tensor Core GPU

A100 Tensor Core GPU는 가속화가 지속되고 있는 데이터 센터 플랫폼에서 전 세대 대비 10배 높은 추론 성능과 새로운 수준의 정밀도 및 가속을 가능하게 하는 3세대 Tensor Core 기술을 제공한다. MIG(Multi-GPU Instance) 기능은 단일 A100을 7개의 인스턴스로 분할함으로써 각 인스턴스에서 서로 다른 신경망이 실행 가능하며 **구조 희소성(Structural Sparsity)** 이라는 새로운 특성을 활용하여 추론 AI의 추가적인 가속화를 지원한다.

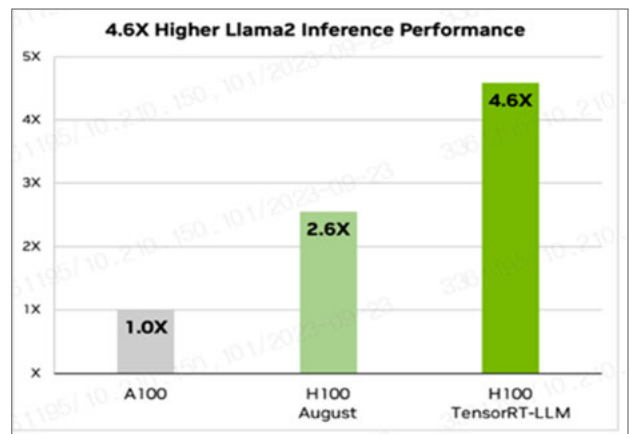
NVIDIA H100 Tensor Core GPU는 모든 워크로드에 대해 전례 없는 성능, 확장성 및 보안을 제공하며 NVIDIA 4세대 NVLINK 활용 시 워크로드 가속화와 전용 Transformer Engine의 조 단위 Parameter의 언어 모델에 대한 지원 병행이 가능하다. H100 PCIe GPU 구성에는 NVIDIA AI Enterprise 소프트웨어 제품군이 포함되어 AI 워크로드의 개발 및 배포 과정 간소화가 가능하며, NVIDIA Hopper 아키텍처 활용을 통해 이전 세대에 비해 대규모 언어 모델의 속도를 30배 향상시킨 대화형 AI 제공이 가능하다. H100 NVL GPU를 장착한 시스템의 경우 1,750개 Parameter의 LLM(large language model)까지 이전 세대 시스템보다 12배 더 높은 Throughput을 기반으로 GPT3-175B에 대한 추론을 지원한다.

그림 23. GPT-J 6B 모델 Tensor LLM 추론 성능 비교



자료: Nvidia, 미래에셋증권 리서치센터

그림 24. Llama2 모델 Tensor LLM 추론 성능 비교



자료: Nvidia, 미래에셋증권 리서치센터

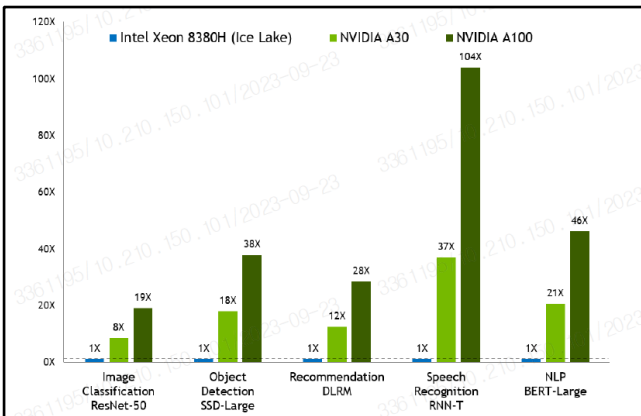
NVIDIA 신제품 및 신기술 동향
 CUDA
 GH200

Nvidia가 반도체 AI 분야의 선두주자로 꼽히는 것은 GPU 성능 외에도 Nvidia가 자체 개발한 'CUDA'라는 프로그램 덕분이다. CUDA는 프로그래머가 병렬처리 알고리즘을 쉽게 만들 수 있도록 개발된 프로그램이며 현재 대부분의 AI 알고리즘은 CUDA 플랫폼을 기반으로 한다. CUDA는 Nvidia GPU에서만 작동하기 때문에 Nvidia GPU가 AI 반도체 시장에서 점유율을 높이는 데 큰 부분을 CUDA가 견인했다.

실제로 CUDA는 고도로 튜닝된(highly-tuned) Math 라이브러리, 데이터 구조와 알고리즘용 Core 라이브러리, 응용 프로그램의 확장을 위한 통신 라이브러리 등 다양한 라이브러리 프레임워크를 제공한다. 또한 C++, Fortran, Python 등 프로그램 언어를 지원하기 때문에 TensorFlow나 PyTorch와 같은 소프트웨어 라이브러리와 연동성을 통한 높은 확장성을 갖추고 있다.

Nvidia의 GH200은 Nvidia의 현재 최고급 AI 칩인 H100과 동일한 GPU를 가지고 있지만 141Gb의 최첨단 메모리 및 72코어 ARM 중앙 프로세서와 페어링이 가능한 제품이다. AI Inference 용 칩으로 설계된 GH200은 기존 제품 대비 더 많은 메모리 용량을 가지고 있으며, 이에 더 큰 AI 모델이 단일 시스템에 들어갈 수 있다. Nvidia의 GH200의 메모리 Capa는 H100의 80GB 메모리 대비 거의 두배에 육박하며 2Q24 Nvidia의 유통업체를 통해 판매될 계획이다.

그림 25. 주요 제품별 per-chip basis 추론 성능 그래프



자료: Nvidia, 미래에셋증권 리서치센터

그림 26. 주요 제품별 per-chip basis 추론 성능 도표

Table 1. Raw Data Showing per-chip Inference Performance Across all Workloads

Cell Heading	Intel Xeon 8380 (Ice Lake)	NVIDIA A30	NVIDIA A100
Image Classification ResNet-50	1X	8X	19X
Object Detection SSD-Large	1X	18X	38X
Recommendation DLRM	1X	12X	28X
Speech Recognition RNN-T	1X	37X	104X
NLP BERT-Large	1X	21X	46X

자료: Nvidia, 미래에셋증권 리서치센터

Inference향 대표 제품
 Goya 추론 프로세서
 Habana Gaudi2

2) Intel

Goya 추론 프로세서는 하바나의 Tensor-Processing Core(TPC)의 아키텍처를 기반으로 8개의 프로그래밍 코어로 구성된 클러스터를 포함한다. TPC는 딥러닝 워크로드를 지원하도록 설계된 하바나의 독점 코어로서 명령어-셋-아키텍처와 딥러닝 워크로드를 효율적으로 서비스하도록 맞춤화된 하드웨어를 갖춘 VLIW SIMD 벡터 프로세서다. TPC는 기본적으로 FP32, INT32/16/8, UINT32/16/8의 혼합 정밀도 데이터 유형을 지원하며 이를 기반으로 한 Goya는 엣지컴퓨팅에 사용되는 대표적인 NPU(통합 신경 처리 장치)이다.

Habana Gaudi2의 경우 생성 AI 및 대형 언어 모델을 훈련하려면 서버 클러스터가 대규모 컴퓨팅 요구 사항을 충족해야 한다. Gaudi2는 384대의 가속기로 GPT-31에 311분이라는 인상적인 훈련 시간을 제공했으며 GPT-3 모델에서 256개에서 384개의 가속기로 거의 선형적인 95% 확장, 컴퓨터 비전(ResNet-508 가속기 및 Unet3D 8 가속기)과 자연어 처리 모델(BERT 8 및 64 가속기)에 대한 우수한 교육 결과를 보여줬다. Gaudi2 11월 제출물 대비 BERT 모델과 ResNet 모델의 성능이 각각 10%, 4% 증가하며 소프트웨어의 성능도가 상승하고 있는 것을 반증했다.

그림 27. Intel Goya High-level Architecture

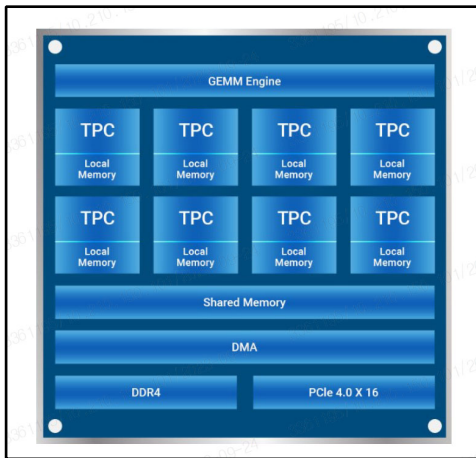
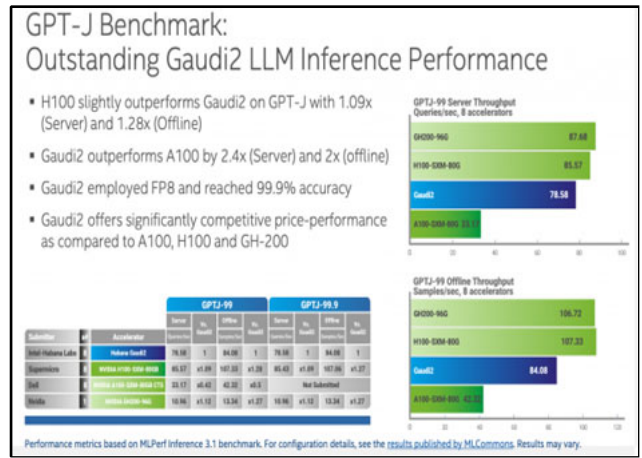


그림 28. GPT-J 모델 Gaudi2 LLM 추론 성능



자료: Intel, 미래에셋증권 리서치센터

자료: Intel, 미래에셋증권 리서치센터

Intel 신제품 및 신기술 동향
Intel Core Ultra Processor

코드네임 메테오 레이크(Meteor Lake)의 PC Inference향 **Intel Core Ultra Processor**는 인텔의 첫 번째 NPU로서 12월 14일 출시 예정이다. Core Ultra는 Foveros 패키징 기술에 의해 만들어진 최초의 클라이언트 Chiplet 디자인 제품으로서 인텔의 클라이언트 프로세서 로드맵의 변곡점을 기록할 것으로 보인다. Core Ultra는 NPU와 전력 효율적인 성능의 주요한 발전에 더해 내장된 인텔 Arc그래픽 기반 독보적인 수준의 그래픽 성능을 제공할 것으로 보인다.

그림 29. Intel Core Ultra Processor 사진

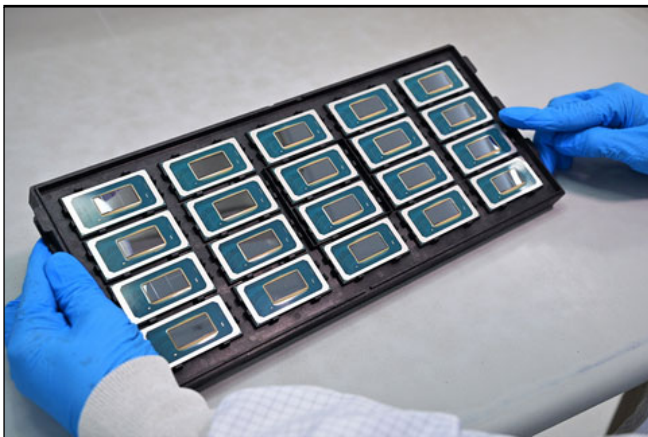


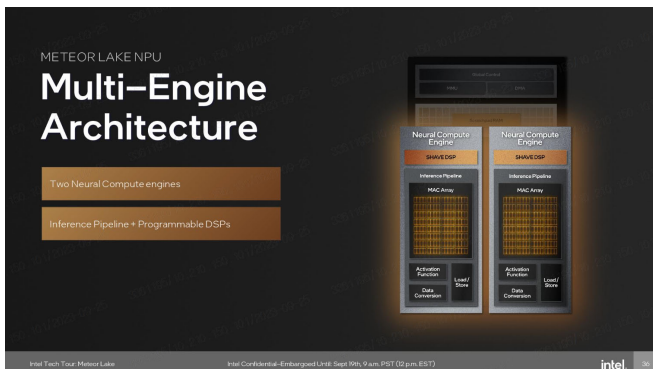
그림 30. Foveros 패키징 비트당 피코줄(pJ/b) 비교

Interconnect	Picojoules per Bit (pJ/b)
NVLink-C2C	1.3 pJ/b
UCle	0.5 - 0.25 pJ/b
Infinity Fabric	~1.5 pJ/b
TSMC CoWoS	0.56 pJ/b
Foveros	Sub-0.3 pJ/b
EMIB	0.3 pJ/b
Bunch of Wires (BoW)	0.7 to 0.5 pJ/b
On-die	0.1 pJ/b

자료: Intel, 미래에셋증권 리서치센터

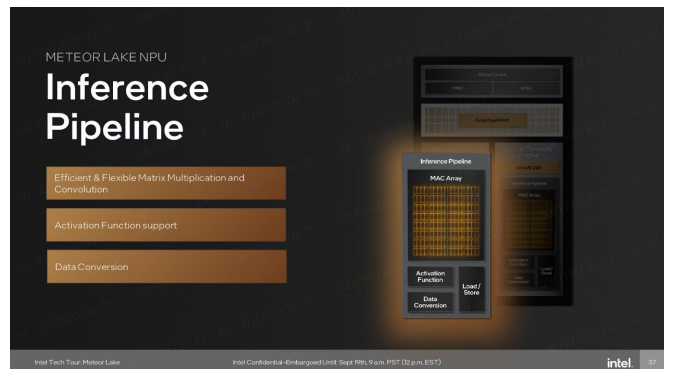
자료: Intel, 미래에셋증권 리서치센터

그림 31. Meteor Lake – Multi-Engine Architecture



자료: Intel, 미래에셋증권 리서치센터

그림 32. Meteor Lake – Inference Pipeline



자료: Intel, 미래에셋증권 리서치센터

3) AMD

Inference향 대표 제품

AMD EPYC 9004 Series 프로세서

AMD EPYC 97x4 Series 프로세서

AMD EPYC 9004 Series 프로세서는 YOLOv5s 워크로드에서 "Zen 4c" 코어로 구동되며 이전 4세대 AMD EPYC 프로세서의 성능과 필적할 뿐만 아니라 이를 능가하여 스트림당 최대 33ms(30FPS)의 최대 지연 시간으로 더 많은 이미지 객체 검출 스트림을 구현했다. 더욱 놀라운 것은 4세대 AMD EPYC 프로세서가 보여주는 세대별 개선이다. 3세대 이전 프로세서와 비교할 때 AMD EPYC 프로세서는 2.4배의 성능 향상을 달성했으며 이는 데이터 센터가 DeepSparse를 사용할 때 발생하는 랙/전력/통화 예산의 전체 비용 감축 및 에너지 효율성 최적화를 통해 이루어진 것으로 보인다.

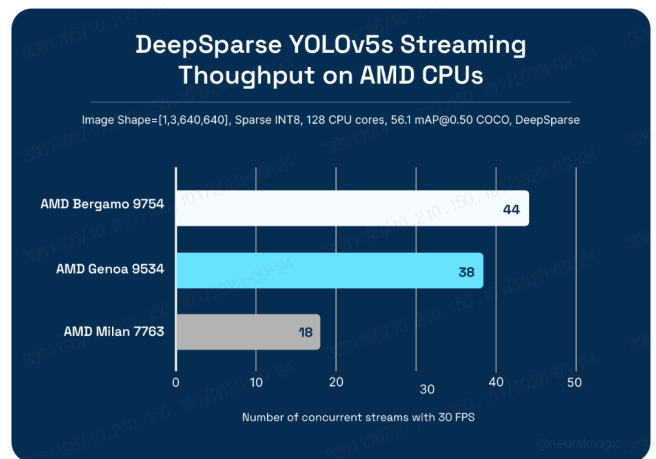
AMD EPYC 97x4 Series 프로세서는 동일한 공간에 더 많은 계산 능력을 담을 수 있는 전력 효율적인 Architecture 채용을 통해 레거시 데이터 센터의 설계 구조로 발생하는 병목현상을 해결한다. 레거시 데이터 센터는 물리적 조건에서 더 많은 고객 워크로드를 처리해야하는 반복적인 문제가 발생하는데 AMD EPYC 97x4 Series 프로세서는 이러한 제약을 해결한다. 또한 DDR5 메모리가 서버 비용의 상당 부분을 차지하므로 단일 소켓 설계에서 메모리 레인에 대한 필요성 감소는 코어당 비용 절감에 기여하며 DeepSparse와 결합된 AMD EPYC 97x4 Series 프로세서는 데이터 센터에서 달러당 성능을 극대화할 수 있다.

그림 33. 4세대 AMD EPYC 시스템 구성

CPU	2 x AMD EPYC 9654
Cores	96 (192 threads) per CPU
Base frequency	2.40 GHz
L3	384 MB
Socket	SP5
Memory	1.58 TB
Settings	SMT enabled
OS	Ubuntu 20.04.2 LTS
Test results (Dense YOLO v5s, 57.1% accuracy with mAP@0.50)	Latency < 33 ms with 20 streams Test 1: 32.9591 Test 2: 32.7182 Test 3: 33.1120 Average: -32.9298 ms
Test results (94% Pruned YOLO v5s, 52.5% accuracy with mAP@0.50)	Latency < 33 ms with 70 streams Test 1: 32.7843 Test 2: 32.6904 Test 3: 32.6253 Average: -32.7000 ms
Test results (Dense oBERT-Base, 95.4% accuracy)	Queries/second Batch size: 256 Test 1: 616.6593 Test 2: 616.4927 Test 3: 616.9794 Average: -616.3771
Test results (90% Pruned oBERT-Base, 93.7% accuracy)	Queries/second Batch size: 256 Test 1: 5670.4560 Test 2: 5668.9001 Test 3: 5660.8382 Average: -5666.7300

자료: AMD, 미래에셋증권 리서치센터

그림 34. AMD DeepSparse Streaming Throughput 개선정도



자료: AMD, 미래에셋증권 리서치센터

AMD 신제품 및 신기술 동향
MI300X
ROCm

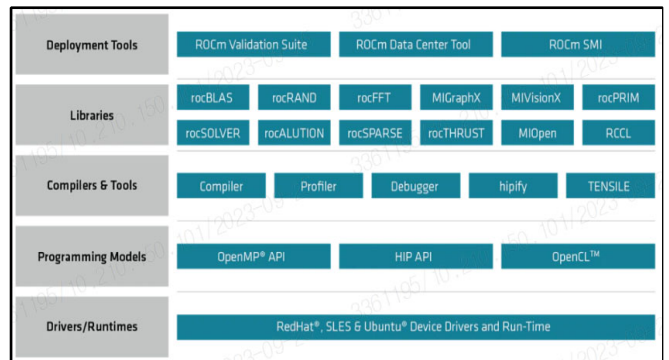
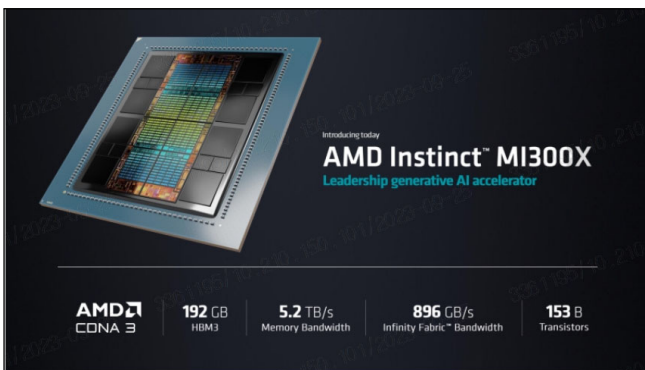
MI300X는 올해 6월에 출시한 AI 애플리케이션용 AMD의 최신형 GPU이다. MI300X는 차세대 AMD CDNA 3 가속기 아키텍처를 기반으로 하며 생성 AI 학습 및 추론과 대형 언어 모델 훈련에 필요한 컴퓨팅 및 메모리 효율성을 제공하기 위해 최대 192GB의 HBM3 메모리를 지원한다. AMD Instinct MI300X의 대용량 메모리로 고객들은 이제 MI300X 가속기5 하나에 400억 parameter 모델인 Falcon-40과 같은 LLM을 장착할 수 있다.

AMD는 MI300A의 Zen4 CPU Chiptlet 3개를 추가 CDNA-3 GPU Chiptlet 2개로 교체하여 GPU 전용 M1300X를 만들었으며 AI 추론 및 훈련 솔루션을 위해 8개의 MI300X 가속기를 업계 표준 설계로 통합하는 AMD Instinct 플랫폼을 개발하였다.

AMD는 NVIDIA의 CUDA 소프트웨어 플랫폼과 경쟁하기 위해 ROCm 소프트웨어 개선 중에 있다. CUDA(Compute Unified Device Architecture)는 NVIDIA의 GPU를 활용하여 성능을 가속화할 수 있도록 하는 NVIDIA의 병렬 컴퓨팅 플랫폼으로서 업계 전문가들은 NVIDIA의 성숙한 소프트웨어 솔루션에 필적하는 AMD의 능력에 대해 회의적인 것으로 보인다. 특히 AMD는 NVIDIA의 소프트웨어에 비해 성능과 AI Tool 적용성 측면에서 열등한 것으로 보인다. 이에 대응하기 위해 AMD는 오픈 소스 전략을 채택하며 소프트웨어 Tool 개발에 집중하고 있다.

그림 35. AMD Instinct MI300X

그림 36. ROCm 5 구조도



자료: AMD, 미래에셋증권 리서치센터

자료: AMD, 미래에셋증권 리서치센터

Inference향 대표 제품
Cloud AI 100

4) Qualcomm

Cloud AI 100은 고급 신호 처리 및 최첨단 전력 효율성으로 데이터 센터, Cloud Edge, Edge Appliance 및 5G 인프라를 포함한 여러 환경을 위한 인공지능(AI) 솔루션을 지원한다. Qualcomm은 저전력, 고성능 및 더 높은 성능/TCO\$ 구동에 대한 AI Edge 시장의 요구 사항을 해결하기 위해 35-55W 열 설계 전력(TDP)으로 구성 가능한 Qualcomm Cloud AI 100 PCIe-Lite Processor를 출시하였다.

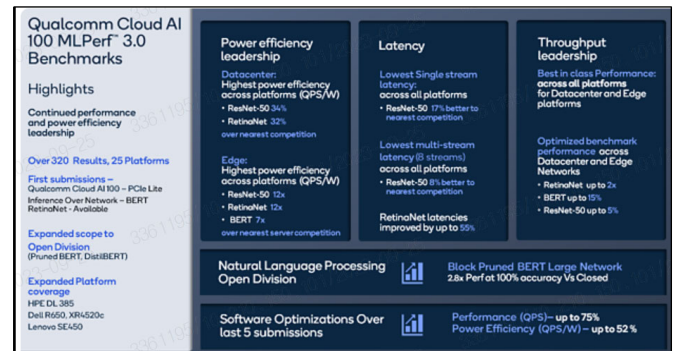
항간에서는 Qualcomm의 Cloud AI 100이 Nvidia의 H100을 성능면에서 앞섰다는 분석이 제기된다. 두 기업의 AI Chip이 이미지 식별과 관련해 와트당 수행하는 쿼리의 수를 테스트한 결과, AI 100이 197.6 쿼리를 수행한 반면 H100은 108.4 쿼리에 그쳤으며 이에 더해 Qualcomm은 객체 탐지 부문에서도 Nvidia에 앞선 것으로 나타났다.

그림 37. Qualcomm 주요 Inference향 제품 라인업

Form Factor	HHHL-Pro	HHHL-Std	HHHL-Lite	Dual M.2	Dual M.2e
Power (TDP)	75 W	75 W	35, 45, 55 W (Configurable)	15 W - 25 W	15 W - 25 W
ML Capacity (INT8)	Up to 400 TOPS	Up to 350 TOPS	Up to 137 TOPS	Up to 200 TOPS	Up to 70 TOPS
On-Die SRAM	144 MD	126 MD	81 MD	126 MD	81 MD
On-Card DRAM	32 GB LPR4x 137 GB/s	16 GB LPR4x 137 GB/s	16 GB LPR4x 137 GB/s	16 GB LPR4x 137 GB/s	8 GB LPR4x 68 GB/s
Host Interface	PCIe Gen 4 8 lanes	PCIe Gen 4 8 lanes	PCIe Gen 4 8 lanes	PCIe Gen 3 4 lanes	PCIe Gen 3 4 lanes

Qualcomm Cloud AI, Snapdragon, and Qualcomm Kyro are products of Qualcomm Technologies, Inc. and/or its subsidiaries.

그림 38. Qualcomm Cloud AI 100 MLPerf 3.0 Benchmark



자료: Qualcomm, 미래에셋증권 리서치센터

자료: Qualcomm, 미래에셋증권 리서치센터

Qualcomm 신제품 및 신기술 동향
Oryon CPU

Qualcomm은 지난 11월 차세대 CPU 모델인 **Oryon**을 발표했다. 이번 PC향 Snapdragon 8cx 4세대 프로세서에 부속 부품으로서 첫 출하가 이루어 질 것으로 보인다. Oryon CPU에 탑재된 새로운 프로세서들은 Hamoa라는 코드네임하에 10월에 공개될 예정이며 1H24 제품 발매가 예상된다.

Oryon CPU는 5nm 공정 사용, 12개의 프로그래밍 코어로 구성된 클러스터를 보유하고 있으며 고전력, 저전력 이분화 기능이 가능하도록 2개의 클러스터로 분리된 구조를 보유하고 있다. 고전력 클러스터의 경우 Nuvia의 Arm Architecture를 채택하였으며 저전력 클러스터의 경우 Arm Cortex-A55 코어를 사용할 것으로 보인다. Oryon CPU는 32Gb LPDDR5와 PCIe 4세대 Storage까지 지원이 가능할 것으로 보인다.

그림 39. Oryon CPU 컨셉 이미지



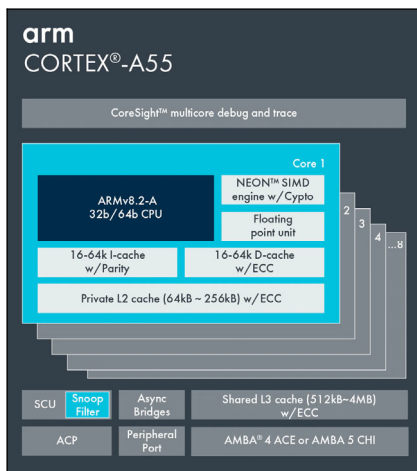
그림 40. Snapdragon 8cx 4세대 프로세서 컨셉 이미지



자료: Gizmochina, 미래에셋증권 리서치센터

자료: Wccftech, 미래에셋증권 리서치센터

그림 41. ARM Cortex-A55 코어



자료: ARM, 미래에셋증권 리서치센터

그림 42. Snapdragon 파트너 업체

자료: FoneArena, 미래에셋증권 리서치센터

Inference향 대표 제품
Google Cloud TPU

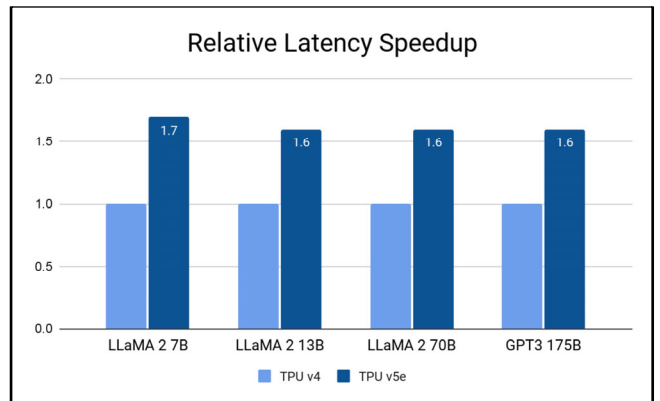
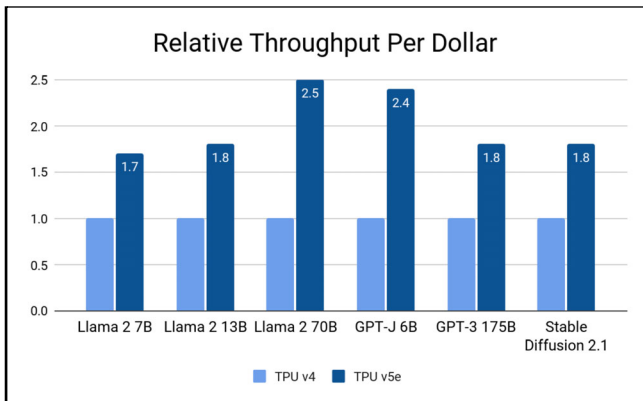
5) Google

Google Cloud TPU(Tensor Processing Unit)는 번역, 사진, 검색, 어시스턴트, 지메일과 같은 구글 제품들의 동력을 공급하는 목적에 따라 만들어진 기계 학습 가속기 칩이다. LLM과 생성형 AI의 획기적인 발전은 AI 모델을 훈련하고 서비스하기 위해 방대한 연산량을 필요로 하며 증가하는 연산 수요를 비용 효율적으로 충족시키기 위해 구글은 Cloud TPU v5e를 맞춤 설계, 구축 및 배포했다.

새로운 Cloud TPU v5e는 최신의 대형 언어 모델(LLM)과 생성 AI 모델을 포함한 광범위한 AI 워크로드에 대한 고성능과 비용 효율적인 추론을 가능하게 한다. TPU v5e는 전세대인 TPU v4 대비 달러당 성능은 2.5배 증가하였으며 각각의 TPU v5e 칩은 최대 393조 int8 연산/초를 제공하여 추론 대기 시간을 1.7배 감소시켰다. TPU v5e 포드는 초고속 링크를 통해 네트워크로 연결된 256개의 칩으로 구성되어 있으며 각 TPU v5e 포드는 초당 최대 10조 int8 연산 또는 100 페타옵스(Peta Ops, Peta Operations Per Second)의 연산 능력을 제공한다.

그림 43. TPU v4 vs v5e Throughput per Dollar 비교

그림 44. TPU v4 vs v5e Relative Latency 비교



자료: Google, 미래에셋증권 리서치센터

자료: Google, 미래에셋증권 리서치센터

Google 신제품 및 신기술 동향
Vertex AI

Google의 **Vertex AI**는 특성 추출부터 모델 학습과 Low-Latency 추론에 이르기까지 데이터 과학자와 AI 개발자가 AI 애플리케이션을 더 쉽게 통합하고 모델의 배포 및 유지 관리를 가속화할 수 있는 관리형 ML 플랫폼이다.

그동안 Google은 수백 가지의 새로운 기능을 더해왔으며 최근 사용가능한 LLM에 Meta의 Llama 2 및 Code Llama, Anthropic의 Claude 2 등을 추가했다. AI 가속기인 Cloud TPU 최신 버전과 Nvidia H100 GPU 인스턴스도 출시하였으며 H100 A3 VM은 이전 세대인 A2 대비 3배 성능 개선이 이루어진 것으로 보인다. Vertex AI 기반 모델, 검색 및 대화 전반에 걸쳐 작동하는 서비스인 그라운드링 서비스와 데이터 엔지니어링, 분석, 예측 분석 등을 위한 단일 인터페이스 빅쿼리 스튜디오가 추가되었다.

그림 45. Vertex AI 로고

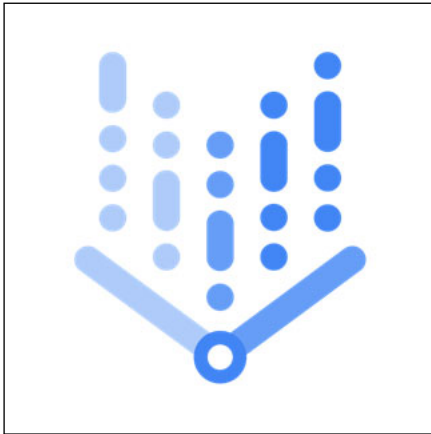
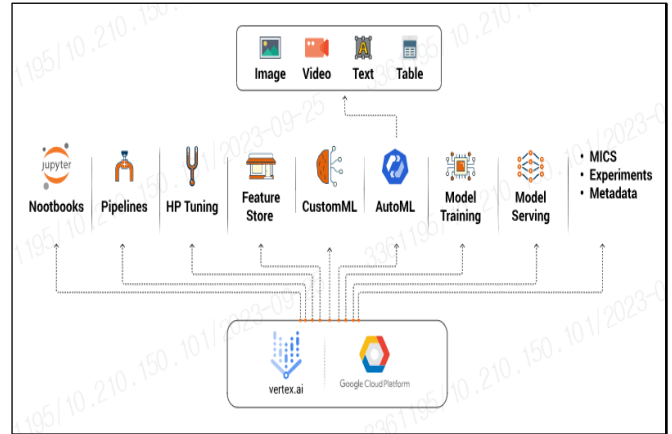


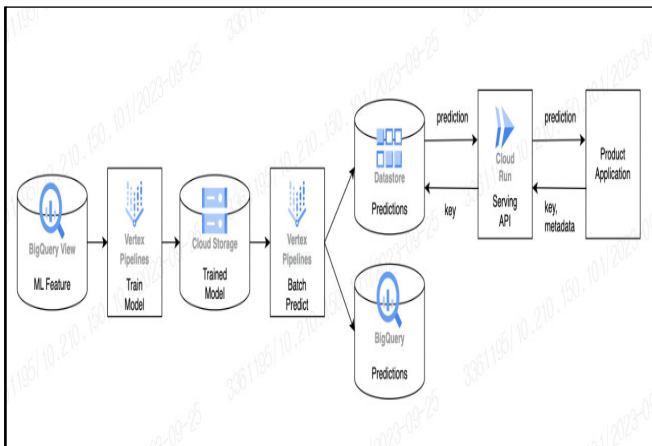
그림 46. Vertex AI 파이프라인



자료: Google, 미래에셋증권 리서치센터

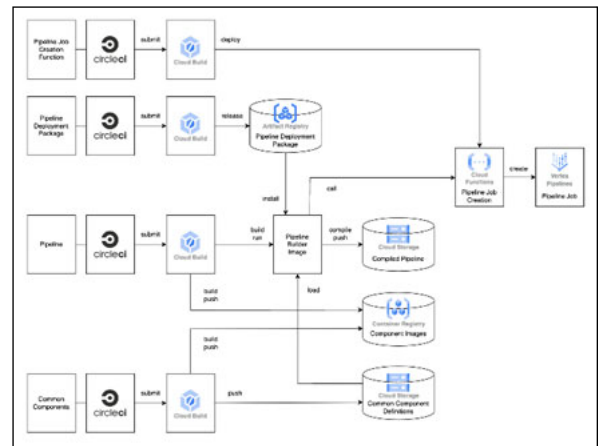
자료: RoyalCyber, 미래에셋증권 리서치센터

그림 47. Vertex Machine Learning Process 도식



자료: AnyMind, 미래에셋증권 리서치센터

그림 48. Vertex AI Pipeline Process 도식



자료: AnyMind, 미래에셋증권 리서치센터

Inference향 대표 제품
ANE(Apple Neural Engine)

6) Apple

ANE(Apple Neural Engine)는 Apple 기기에서 Deep Neural Network를 에너지 효율적으로 실행하기 위해 최적화된 컴퓨팅 코어 클러스터이다. ML과 AI 알고리즘을 가속화해 메인 CPU나 GPU보다 빠른 속도와 높은 메모리, 전력 효율성의 장점을 제공한다.

2022년, M1 Ultra는 UltraFusion이라 불리는 애플의 맞춤형 Interconnect를 사용하여 두 개의 M1 Max 칩을 단일 패키지로 결합했다. ANE 코어(32개)의 두 배를 가진 M1 Ultra는 ANE 성능을 기존 대비 두 배인 초당 22조 동작으로 늘렸다. 2022년 Apple A16은 TSMC의 향상된 N4 노드를 사용하여 제조되었으며, A15의 ANE에 비해 ANE 성능이 약 8% 개선되었다. 2023년 M2 MAX 제품에 있어 M2와 동일한 ANE를 사용한 것으로 보이며 최근 기술 동향 특이사항은 없는 것으로 보인다.

그림 49. Apple 제품별 사용 부품 및 성능 연혁

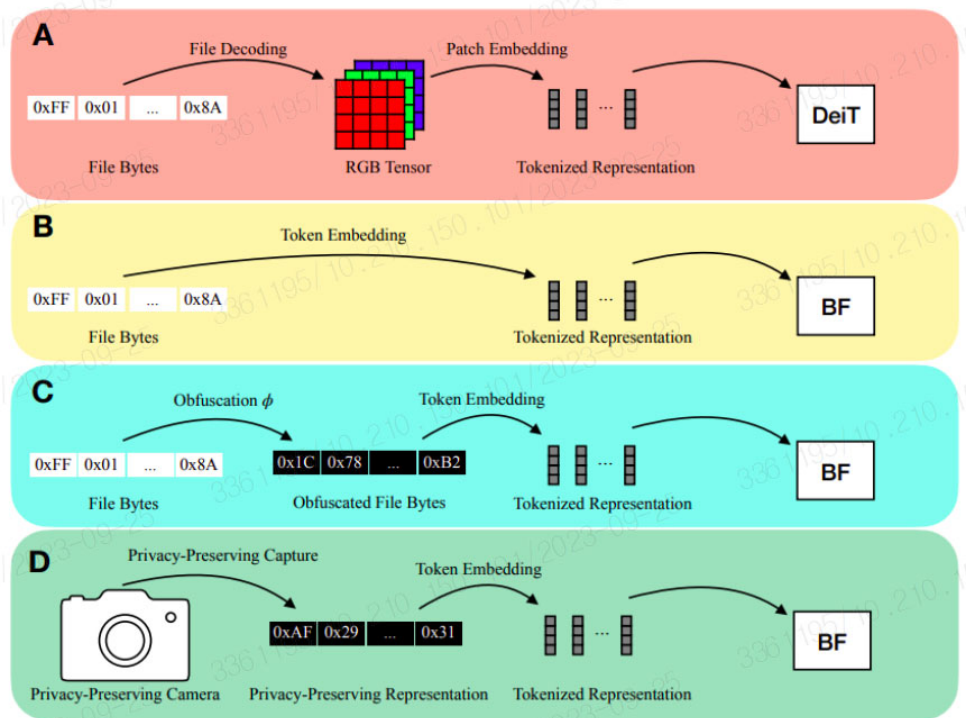
Apple Silicon	Semiconductor Process Node	Launch Date	ANE Cores	Operations Per Second	Additional Notes
A11 Bionic	10nm TSMC FinFET	2017	2	600 billion	Apple's first ANE
A12 Bionic	7nm TSMC FinFET	2018	8	5 trillion	9x faster than A11, 90% lower power consumption
A13 Bionic	7nm TSMC N7P	2019	8	6 trillion	20% faster than A12, 15% lower power consumption
A14 Bionic	5nm TSMC N5	2020	16	11 trillion	Nearly 2x faster than A13
A15 Bionic	5nm TSMC N5P	2021	16	15.8 trillion	40% faster than A14
A16 Bionic	5nm TSMC N4	2022	16	17 trillion	8% faster than A15, better power efficiency
M1	5nm TSMC N5	2020	16	11 trillion	Same ANE as A14 Bionic
M1 Pro	5nm TSMC N5	2021	16	11 trillion	Same ANE as A14 Bionic
M1 Max	5nm TSMC N5	2021	16	11 trillion	Same ANE as A14 Bionic
M1 Ultra	5nm TSMC N5	2022	32	22 trillion	2x faster than M1/M1 Pro/M1 Max
M2	5nm TSMC N5P	2022	16	15.8 trillion	40% faster than M1
M2 Pro	5nm TSMC N5P	2023	16	15.8 trillion	Same ANE as M2
M2 Max	5nm TSMC N5P	2023	16	15.8 trillion	Same ANE as M2

자료: Apple, 미래에셋증권 리서치센터

Apple 신제품 및 신기술 동향
ByteFormer

Apple 개발자들이 개발한 **ByteFormer**는 Vision-Language Transformer AI 모델로서 TIFF 형식으로 저장된 데이터를 사용하며 ImageNet 분류에 있어 약 77.33%의 정확도를 보였다. 해당 모델은 DeiT-Ti Transformer Backbone Hyperparameters를 사용하며, RGB 입력에서 72.2%의 정확도를 기록했다. 또한 아키텍처의 수정 혹은 Hyperparameter의 조정 없이 음성 명령 v2에서 최첨단 수준인 98.7%의 근사치인 95.8%의 정확도에 더해 특유의 input 방식으로 인해 높은 보안성까지 보유했다.

그림 50. ByteFormer와 기존 DeiT 활용 추론 프로세스 비교 도식



자료: Marktechpost, 미래에셋증권 리서치센터

Inference향 대표 제품
Ascend 310 프로세서
Ascend 910 프로세서

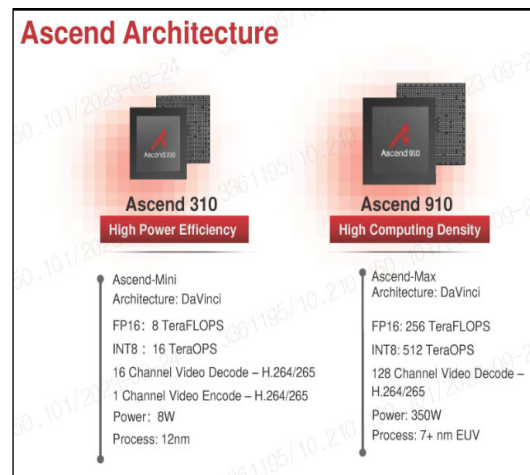
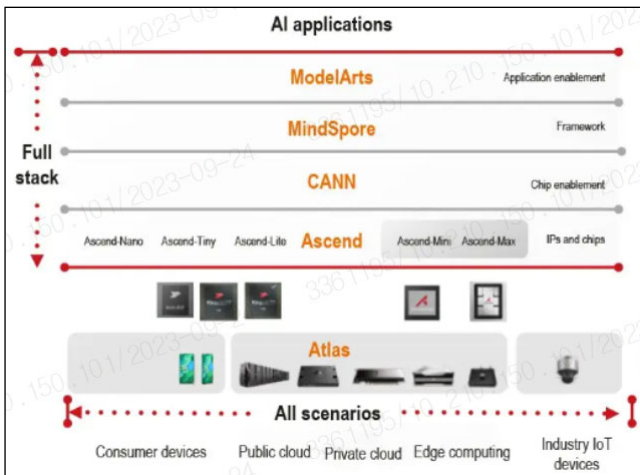
7) Huawei

Ascend 310 프로세서는 효율적이고 유연하며 프로그램 가능한 설계를 특징으로 16개의 TOPS@INT8 및 8개의 TOPS@FP16을 출력하는 전형적인 구성으로 단 8W의 전력을 소비하도록 최적화 되어있다. 이는 Huawei 고유의 다빈치 아키텍처가 주요하게 작용하며, 다양한 범위의 컴퓨팅 유닛을 통합하여 AI 칩의 적용 범위를 확장한다. 모든 AI 서비스 프로세스를 가속화함으로써 Ascend 310은 전체 AI 시스템의 성능을 향상시킬 뿐만 아니라 배포 비용 또한 크게 절감한다.

Ascend 910 프로세서는 AI 컴퓨팅 능력의 경계를 재정 의하는 고통합 SoC 프로세서로서 320 TFLOPS@FP16 및 640 TOPS@INT8 컴퓨팅 성능을 자랑한다. Ascend 910은 최대 310W의 전력 소비를 보유했으며 이 또한 높은 컴퓨팅 능력의 잠재력을 최대한 활용하기 위해 AI 코어, CPU, DVPP 및 작업 스케줄러를 원활하게 통합하는 화웨이 의 독점 다빈치 아키텍처에 의해 가능하다. Ascend 910은 HCCS, PCIe 4.0 및 RoCE v2를 포함함으로써 유연하고 효율적인 스케일업 및 스케일아웃 시스템을 구축할 수 있는 기반을 조성한다.

그림 51. Huawei 풀스택 AI 솔루션 구조

그림 52. Huawei Ascend 310, Ascend 910 제품 사양



자료: Huawei, 미래에셋증권 리서치센터

자료: Huawei, 미래에셋증권 리서치센터

Hwawei 신제품 및 신기술 동향
Pangu Model 3.0

Pangu Model 3.0은 사전 학습된 AI 모델 시스템으로서 여러 산업 분야에 적용될 수 있도록 개발되었다. Pangu는 대규모 데이터 세트 및 ML Algorithm을 통해 기상예보, 신약 개발, 광산, 열차 내 결함 식별 등 다양한 분야에 활용될 것으로 예상되며 화웨이는 지난 7월 산둥에너지그룹(Shandong Energy Group)과 Pangu 광산 모델(Pangu Mining Model)의 첫 상용 사례를 발표하였었다. 이외에도 화웨이는 Pangu-Weather AI예측 모델이 기존 수치 기반 예보 방식보다 더 높은 정확도를 보여준다 밝혔다.

그림 53. Huawei 개발자 컨퍼런스 – Pangu Model 3.0 발표



자료: Huawei, 미래에셋증권 리서치센터

Inference향 대표 제품
Hanguang 800

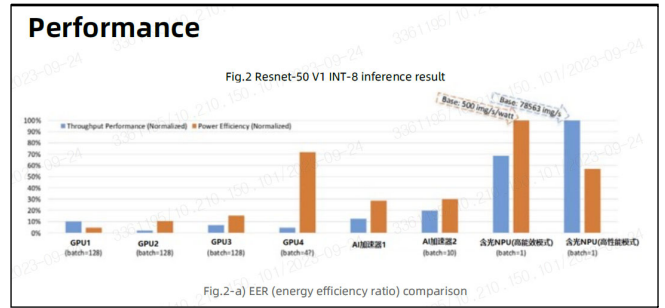
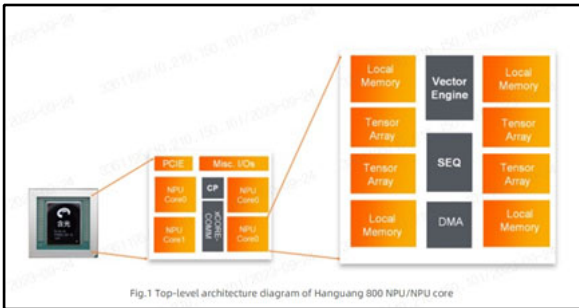
8) Alibaba

Hanguang 800은 데이터 센터의 AI 응용 프로그램을 위해 설계된 170억 개의 트랜지스터를 통합한 12m 고성능 인공지능 추론 칩이다. 아키텍처는 도메인별 컴퓨팅 엔진 및 실행 장치, 192M 로컬 메모리(SRAM) 및 빠른 데이터 액세스를 위한 코어 간 통신 등 AI 추론에 맞춤형으로 제공되어 컴퓨팅 성능과 지연 시간이 모두 최적화 되어있다. MLPerf(산업 표준 AI 벤치마크) Resnet-50 테스트에서 Hanguang 800은 78563의 추론 Throughput을 달성했다

Hanguang 800은 TensorFlow, MXNet, ONNX 등 주류 딥러닝 프레임워크를 지원할 수 있으며 추론 작업을 업계 최고 수준의 성능과 전례 없는 효율성으로 처리할 수 있어 빠르게 성장하는 데이터센터에 적합하다. Hanguang 800 AI 칩은 커맨드 프로세서, 코어간 통신(CORE-COMM) 및 복수의 NPU 코어를 포함한다. 단일 NPU 코어는 일반 추론(ResNet50 v1 등)을 처리하는 업계 최고의 성능과 효율성을 가지고 있을 뿐만 아니라, 코어 간 통신을 통해 여러 NPU 코어가 긴밀하게 작동하여 더 크고 복잡한 작업(ResNet101, Mask R-CNN 등)을 빠른 속도로 처리할 수 있다.

그림 54. Hanguang 800 NPU, NPU Core 설계도

그림 55. Hanguang 800 NPU 성능



자료: Alibaba, 미래에셋증권 리서치센터

자료: Alibaba, 미래에셋증권 리서치센터

Alibaba 신제품 및 신기술 동향
통이치엔원(Tongyi Qianwen)

알리바바는 6월 알리바바 Cloud가 자체 개발한 LLM인 **통이치엔원**을 디지털 어시스턴트 **팅우(Tingwu)**에 통합하였었다. 통합 이후 **통이 텡우(Tongyi Tingwu)** 라는 명칭 하에 새로 배포되며 AI 기반 멀티미어 콘텐츠 분석, 영상 및 오디오 파일 등 다양한 형태의 텍스트 요약문 생성 작업을 수행하였다.

이에 더해 알리바바는 오픈 소스 전략을 채택함에 따라 통이치엔원을 무료로 배포할 예정이라 밝혔으며 이에 **중국** 승인을 받은 것으로 알려졌다. 또한 향후 적극적인 LLM 개발 확대 목적으로 **OPPO, Taobao, DingTalk, Zhejiang University** 등 산학협력을 맺었다 WeChat 공식 채널을 통해 발표했다.

그림 56. 통이치엔원(Tongyi Qianwen) 로고

그림 57. 통이치엔원(Tongyi Qianwen) 통합 Alibaba 앱



자료: Cloudbooklet, 미래에셋증권 리서치센터

자료: Codeandhack, 미래에셋증권 리서치센터

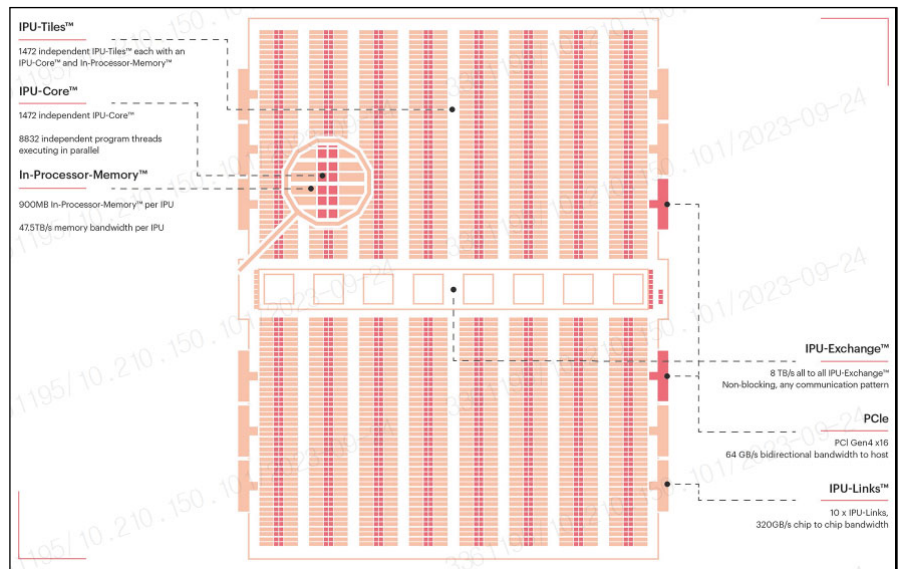
Inference향 대표 제품
Colosus MK2 GC200 IPU

9) Graphcore

Colosus MK2 GC200 IPU는 2세대 Colosus MK2 IPU 프로세서로서 Poplar SDK와 공동 설계된 새로운 종류의 대규모 병렬 프로세서이다. 1세대 Colosus IPU 출시 이후 동사는 실리콘 및 시스템 아키텍처에서 컴퓨팅, 통신, 메모리 분야에서 MK1 IPU 대비 성능을 8배 향상시킨 Colosus MK2 GC200 IPU를 출시했다.

594억개의 트랜지스터를 갖추고 가장 최신 TSMC 7nm 공정을 사용하여 제작된 Colosus MK2 GC200 IPU에는 1,472개의 프로세서 코어가 있으며 약 9,000개의 독립적인 병렬 프로그램 스레드를 실행한다. 각 IPU는 FP16.16 및 FP16.SR에서 250 테라플롭스의 AI 컴퓨팅이 가능한 전례 없는 900MB In-Processor-Memory를 보유하고 있다. GC200은 타 프로세서 대비 독보적으로 많은 FP32 컴퓨팅을 지원한다.

그림 58. Graphcore Colossus MK2 GC200 IPU 구조



자료: Graphcore, 미래에셋증권 리서치센터

Graphcore 신제품 및 신기술 동향
BOW IPU

BOW IPU는 세계 최초로 웨이퍼 온 웨이퍼(WoW) 3D 스택킹 기술을 사용하는 프로세서로 개별 IPU의 성능을 한 단계 끌어올린 제품이다. 연산 아키텍처와 실리콘 구현, 통신 및 메모리의 획기적인 발전을 특징으로 하는 BOW IPU는 최대 1.4페타플롭스의 AI 연산을 제공하며, 이전 세대의 IPU 대비 40%의 성능 개선과 전력 효율을 16% 증가시켰다.

BOW 제품 라인업에는 Bow Pod16, Bow Pod64, Bow Pod256이 있다. Bow Pod16은 4페타플롭의 연산성능을 보유하고 있으며 IPU 시스템을 통해 높은 성능과 적용성을 보유하고 있다. Bow Pod64는 학습 및 추론 워크로드에 모두 사용 가능한 16페타플롭의 AI 연산 성능을 보유함에 대규모 AI에 적합하게 설계되었다. BOW Pod64의 경우 16페타플롭의 AI 연산 성능을 제공하는 Bow Pod64는 학습 및 추론 워크로드에 모두 사용 가능하도록 설계되었다.

그림 59. Graphcore BOW IPU 구조

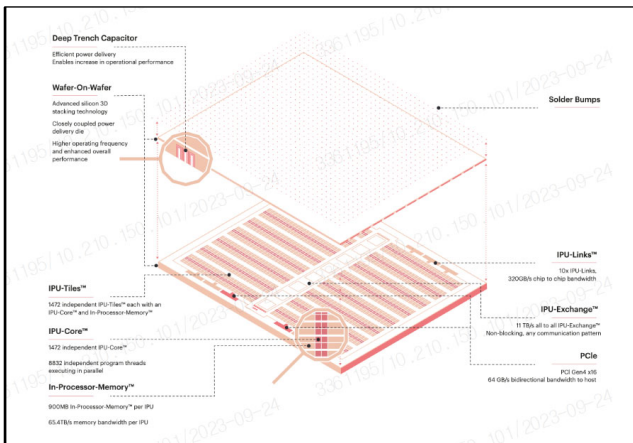
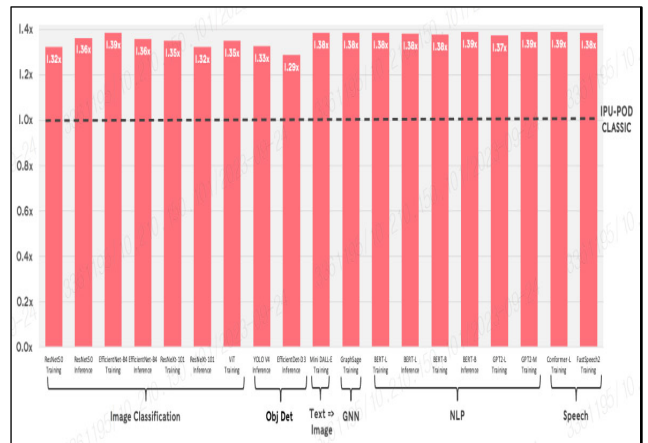


그림 60. BOW POD 및 IPU-POD CLASSIC 성능 비교



자료: Graphcore, 미래에셋증권 리서치센터

자료: Graphcore, 미래에셋증권 리서치센터

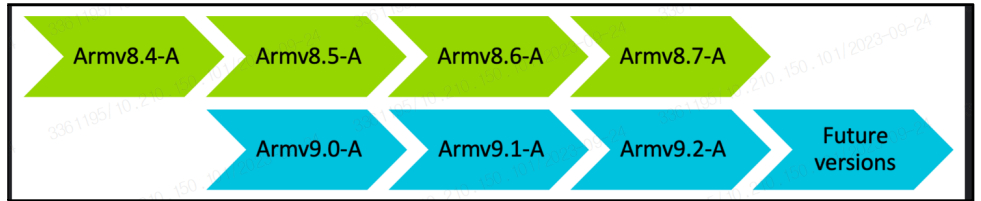
Inference향 대표 제품
Armv9-A

10) ARM

Armv9-A는 이전 세대 Armv8-A를 기반으로 만들어진 Arm Architecture의 최신 버전으로서 Scalable Vector Extension 2(SVE2), Transactional Memory Extension (TME), Branch Record Buffer Extension (BRBE), Embedded Trace Extension(ETE), Trace Buffer Extension(TRBE)와 같은 새로운 기능들을 추가했다.

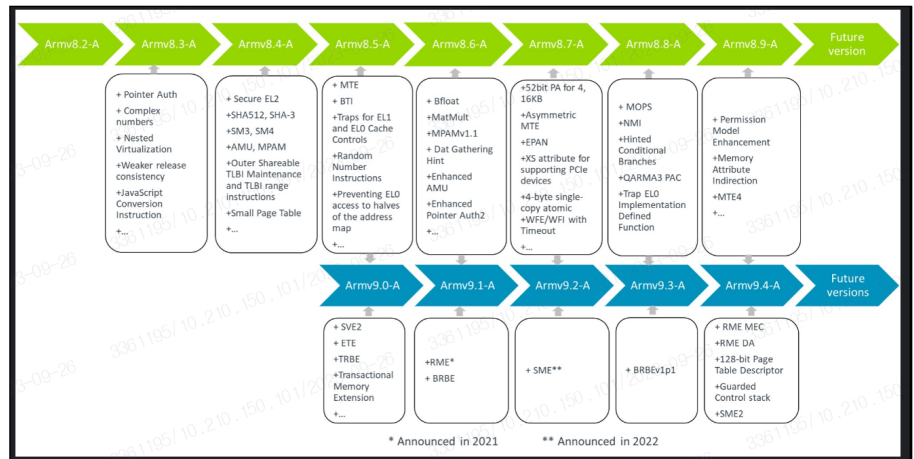
Arm은 아키텍처에 대한 매년 업데이트를 게시하며 새로운 명령어와 기능을 추가한다. Armv9.0-A는 Armv8.5-A와 일치하며 Armv8.5-A의 모든 기능을 이어받고 새로운 기능을 추가하며 Armv9-A의 초기 출시 이후에는 Armv8-A와 Armv9-A가 함께 업데이트된다.

그림 61. ARMv9.0-A, Armv8.5-A 업데이트 과정 도식



자료: ARM, 미래에셋증권 리서치센터

그림 62. ARMv9.0-A, Armv8.5-A 업데이트 과정 도식(세부)



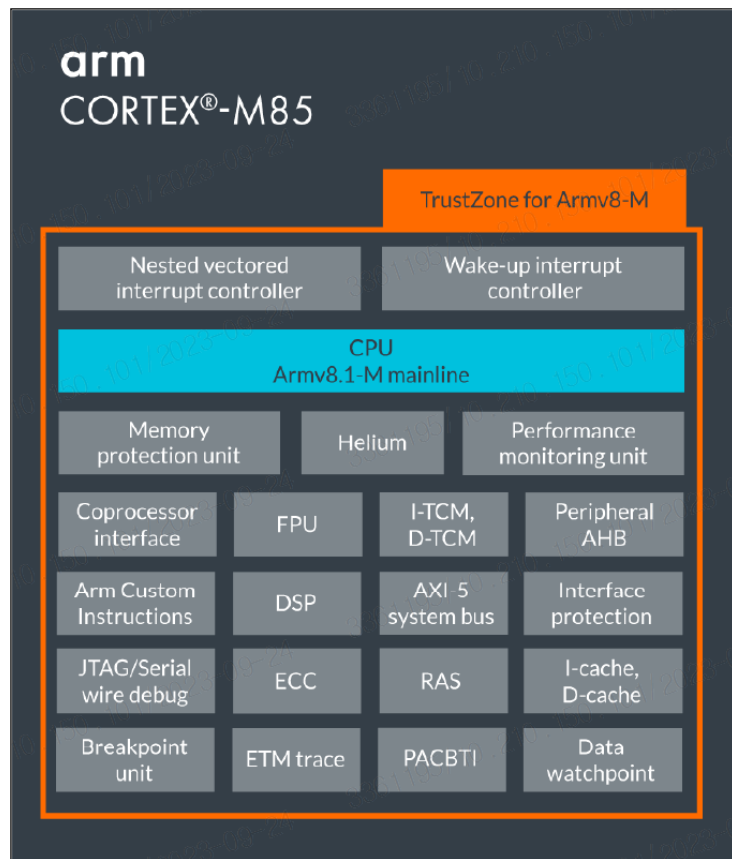
자료: ARM, 미래에셋증권 리서치센터

ARM 신제품 및 신기술 동향
Ethos NPU
Cortex M Series

Ethos NPU 제품들은 확장성이 뛰어난 高효율 2세대 ML 추론 프로세서로서 여러 시장 부문에서 저렴한 비용으로 프리미엄 AI 솔루션 구축에 활용된다. 특히 확장성과 효율성이 뛰어난 Arm의 2세대 NPU Ethos-N78은 다중 코어 기술을 통해 싱글 코어 성능을 2.5배 증가시켰다.

Arm Cortex-M85는 Arm Helium 기술이 적용된 최고 성능의 Cortex-M 프로세서이며 Cortex-M 기반 자연스러운 업그레이드 경로를 제공한다. Arm TrustZone 기술 외에도 Cortex-M85는 새로운 보안 및 안전 기능을 기반으로 응용 프로그램 연결 및 활용 시 Cortex-M에서 가장 높은 수준의 보안을 제공한다.

그림 63. Arm Cortex-M85 제품 아키텍처



자료: ARM, 미래에셋증권 리서치센터

3. AI 반도체 투자포인트

1) 추론 분야도 NVIDIA의 CUDA 생태계가 주도할 가능성

(NVIDIA 점유율 기반) NVIDIA가 GPU 높은 점유율을 유지할 수 있는 이유는 프로세서 자체의 경쟁력 뿐 아니라 자사 GPU 기반의 CUDA(쿠다) 소프트웨어 생태계의 요인이 크다고 알려져 있다. CUDA를 활용한 cuDNN(CUDA Deep Neural Network)이 AI 라이브러리를 선점함에 따라 이후 대부분의 AI 개발자는 이 라이브러리 위에 AI 모델을 구축하기 시작했다. Google의 Tensor Flow, Meta의 PyTorch 등 AI 선두업체의 프레임워크마저 cuDNN 라이브러리를 채택하는 등 점차 개발자들은 CUDA 기반 라이브러리와 NVIDIA GPU에 익숙해지기 시작했다.

(CUDA 탈피 시도) 최근 AMD와 Intel이 각각 ROCm와 원API를 출시하며 CUDA 기반 GPU 소프트웨어 생태계를 대체하겠다고 나서고 있으나, NVIDIA가 십수년간 쌓은 영역을 후발주자가 단기간에 따라잡기 쉽지 않은 상황이다. NVIDIA 또한 DNN 라이브러리 뿐 만 아니라 다수의 CUDA 기반 소프트웨어 생태계를 구축하는데 대규모 자본을 투입하는 중이다.

(추론 시장 가능성) 다만, 추론 인프라는 NVIDIA GPU 대체의 여지가 일부 있다. Intel이 인수한 Habana랩스는 'Gaudi2'의 Mlperf 벤치마크 테스트에서 Nvidia의 A100 성능을 앞섰는데, 이는 cuDNN을 Intel의 원API로 교체했을 때 성능에 문제가 없음을 의미한다. 그러나, 하드웨어 제조사마다 칩에서 지원하는 기능이 동일할 수 없으므로 학습 단계부터 추론 인프라를 고려해 AI 모델을 설계하는 수밖에 없다. 이런 설계역량을 갖춘 주체는 극소수다.

(CUDA 탈피 시도) OpenAI는 오픈소스 언어 Triton을 이용해 GPT 모델을 개발했다. Triton은 학습 단계부터 DNN 라이브러리나 드라이버 등을 CUDA가 아닌 자체 소프트웨어로 쓰는 것이다. Meta와 Google, MSFT, AWS 등도 자체적인 AI 스택을 보유하고 있거나 구축중인 것으로 알려져있다. 그러나, 상기 빅테크 업체들이 추론용 자체 AI 소프트웨어 스택을 공개할 가능성이 점차 줄고 있다. AI 주도권을 확보하기 위한 경쟁이 심화되고 있기 때문이다. NPU 프로세서 개발사들 또한 하드웨어 설계와 함께 소프트웨어 스택 개발에도 나선 사례가 있으나, GPU에 대한 완전한 독립은 불투명하다.

(결론) 당사는 AI 추론시장에서도 NVIDIA의 GPU 및 생태계 시장 점유가 지속될 가능성이 높다고 판단하며, NVIDIA에 대한 긍정적인 전망을 유지한다. 비록 빅테크 업체들의 자체 프레임워크와 라이브러리를 구축할 의지와 역량이 있지만, 헤게모니 확보 경쟁 가운데 개발 소스를 상호 오픈하기는 구조적으로 어려워 보인다. 빅테크사의 자체 추론 모델에의 NVIDIA 의존도는 떨어질 수 있겠으나, 그 외의 대부분의 추론 모델 운영주체에게는 CUDA외의 대안이 마땅치 않은 것으로 판단한다.

2) AI 추론용 IP 개발사의 확장성에 주목

앞서 언급했듯이, 추론시장 개화에도 Cloud 및 Datacenter의 AI 프로세서는 Nvidia의 경쟁력이 지속될 것으로 판단한다. 다만, 추론용 반도체 시장에서의 구조적 변화로부터 IP(반도체 설계 자산) 업체들의 수혜를 예상한다.

첫째, NVIDIA의 높은 점유율에도 불구하고 빅테크 및 탑티어 칩메이커들의 경우 자체 추론 프로세서를 만들고자 하는 시도는 지속될 것이며, 이와 관련한 상세한 예시를 전술한 바 있다. 당사는, 이러한 변화 가운데 반도체 IP 설계 업체들의 중요도가 증대될 것으로 판단한다. NVIDIA와 경쟁하는 과정에서 막대한 시간과 비용을 필요로 함에 따라 반도체 설계 분업의 니즈가 커지기 때문이다. 통상 SoC 개발기간 3.5~5년인데, IP 하는데 1~2년, 이후 디자인하우스를 거치는 기간이 2년정도로 총 8~9년여가 소요된다. 여러 회사들이 분업을 하게 되면 기간이 단축되고 비용을 절감할 수 있다.

둘째, 엣지디바이스 시장에서의 AI 추론시장 개화는 칩메이커 보다는 IP 업체에 더 큰 기회가 될 수 있다. AI 추론 시장은 점차 엣지디바이스로 확장될 가능성이 높다. 데이터를 클라우드로 전송하여 처리하고 다시 피드백을 받는 대신, 개별 디바이스 내에서 직접 추론 연산을 수행함으로써 시간 지연을 최소화하고 데이터 보안을 지킬 수 있기 때문이다. AI 추론이 적용되는 대표적인 엣지디바이스는 자율주행차, 보안카메라, 스마트폰 등이다.

자율주행차나 Non-Mobile 디바이스에는 별도의 AI 추론 프로세서가 탑재될 가능성이 충분하다. 전력이나 공간상 제약이 상대적으로 덜하기 때문이다. 그러나, 스마트폰이나 IoT등 Mobile 디바이스에는 신규 프로세서의 추가 탑재는 사실상 제한적이다. BoM cost의 문제도 있지만, 무엇보다 공간상 제약이 크다. 스마트폰을 예로 들더라도 새로운 AI 프로세서가 탑재되는 것은 사실상 쉽지 않다. 따라서, 이러한 Mobile 디바이스에는 AI 추론과 관련된 IP가 기존 프로세서에 추가될 가능성이 높다.

(결론) 당사는 Datacenter 용 Non NVIDIA 추론 프로세서의 지속적인 시도와 엣지디바이스에서의 제한적 여건가운데, AI 추론능력 탑재의 과정에서 IP 업체들의 수혜가 기대되며 이에 긍정적인 전망을 제시한다. 대표적 AI관련 IP 업체로는 나스닥의 ARM과 국내의 칩스앤미디어, 오픈엣지테크놀로지 등이 있다.

IV. 소프트웨어/통신

1. 빅테크 생성AI 비즈니스 현황 - 승자는?

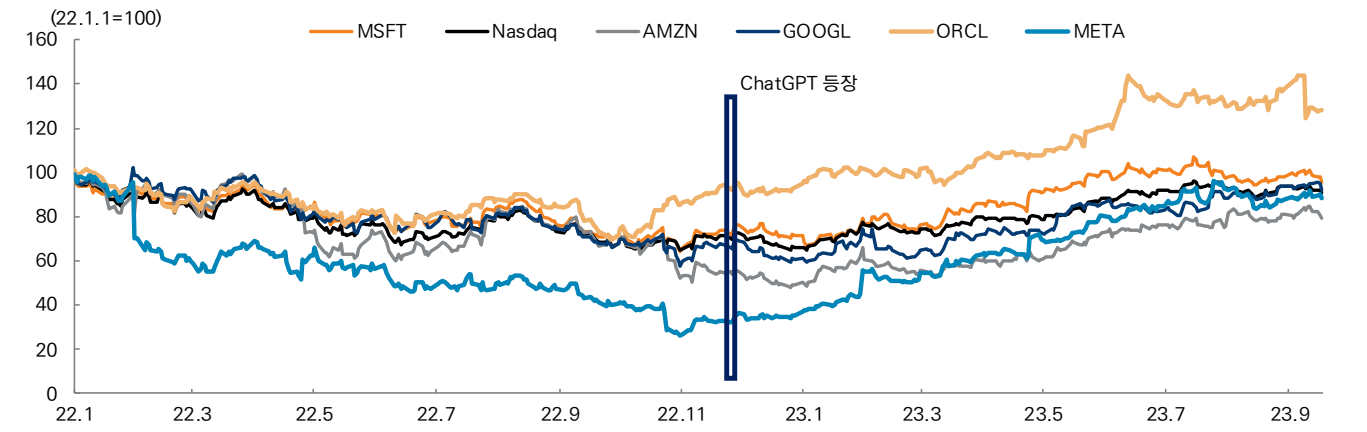
빅테크의 Generative AI(생성AI) 주도권 싸움이 치열하다. OpenAI와 함께 'GPT-3.5'로 제일 먼저 생성AI를 수면위로 가져온 마이크로소프트(MSFT US)부터, AI 관련 연구가 가장 앞서 나간 구글(GOOG US), 클라우드 인프라를 기반으로 생성AI 어플리케이션 개발 플랫폼에 집중하는 아마존 AWS(AMZN US), LLM을 오픈소스로 공개해 생태계 주도권을 잡으려는 메타(META). 이에 더해 클라우드 산업에서 데이터분석에 탁월한 경쟁력을 가진 오라클(ORCL US)도 점유율을 확대하는 중이다.

생성AI는 새로운 산업 생태계를 만들어가고 있다. 인터넷 시대에는 검색 엔진과 인터넷 커머스가 부상했고, 스마트폰 시대에는 애플이 주도권을 잡았으며, 클라우드 시대에는 아마존이 치고 나갔듯이, 생성AI 시대 생태계의 중심이 되기 위해 빅테크는 속도전을 펼치고 있다. 생성AI를 확장하기 위해서는 반드시 클라우드가 필요하다. 따라서 각 클라우드 기업들은 자사의 머신러닝 플랫폼에서 생성AI를 확대 활용할 수 있는 기능들을 계속해서 출시하고 있다. 특히 주요 고객이 엔터프라이즈(기업)인 만큼, 이를 위한 제품을 런칭하기 시작했다.

그리고 어플리케이션의 폭발적인 성장을 앞두고, 앱을 더 간편하게 만들도록 돕는 미들웨어의 중요성이 강조되면서 현재는 관련 서비스들이 지속적으로 출시되고 있다. 9월 들어 어플리케이션의 Monetize도 본격화되는 양상이다. 마이크로소프트와 어도비가 각각 정식 제품 판매를 시작했다. 생성AI 시장은 계속해서 숨가쁘게 진화 중이다.

이제 생성AI 기대감이 현실로 바뀌는 시점에 왔다. 정말 매출이 발생하는 기업 중심으로 주가도 우상향 할 것이다. 아직까지 어플리케이션이 본격적으로 등장하지 않은 시점 이므로, 어플리케이션 개발 플랫폼에 집중하는 **구글(GOOG US), 아마존(AMZN US)**의 성장 우위를 예상한다. 어플리케이션 중에서는 이미지 생성AI 시장을 주도하는 **어도비 (ADBE US)**가 성과를 내기 시작할 것으로 전망한다.

그림 64. 생성AI 시대 도래, 대표 클라우드 기업 주가와 나스닥 주가 비교

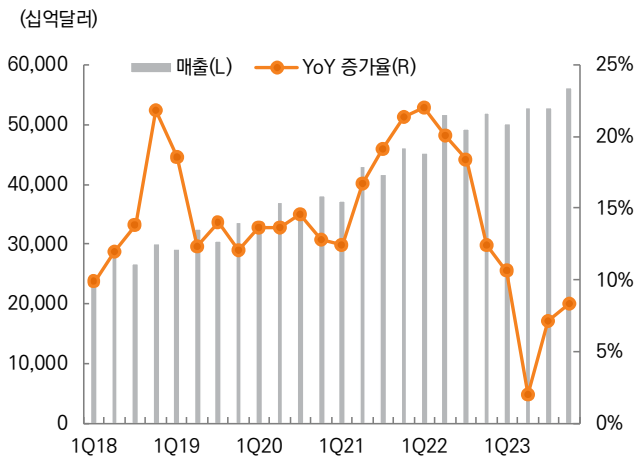


자료: 블룸버그, 미래에셋증권 리서치센터

마이크로소프트 (MSFT US)

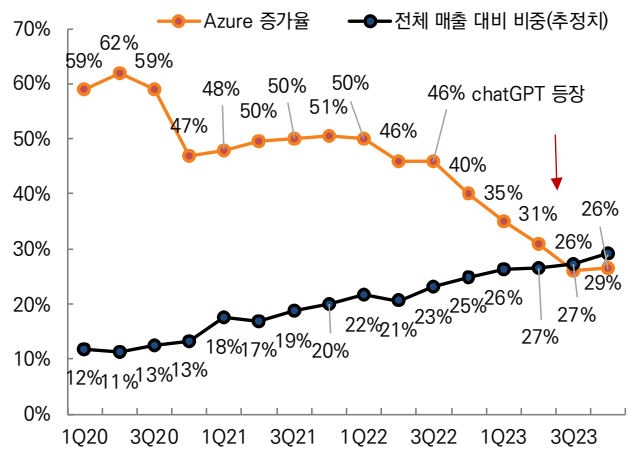
OpenAI에 Azure 인프라를 투자한 뒤 ChatGPT를 출시하며 주도권을 잡은 마이크로소프트는 9월 Microsoft Copilot 출시를 본격화했다. MS의 생성AI 이름을 모두 '마이크로소프트 코파일럿(Microsoft Copilot)'으로 통합했다. MS는 빅테크 기업들 가운데 생성AI 어플리케이션으로 가장 먼저 수익화를 실현할 수 있을 것으로 기대한다. 다만 그 시장 규모나 파급력을 아직까지 예측하기 어려운 상황이다. Copilot의 매출 인식은 올해 말부터 시작되어 내년부터 반영될 것으로 전망한다. 따라서 CY3Q23 실적 이후 성과를 통해 투자 판단을 내리는 것을 추천한다.

그림 65. 마이크로소프트 매출



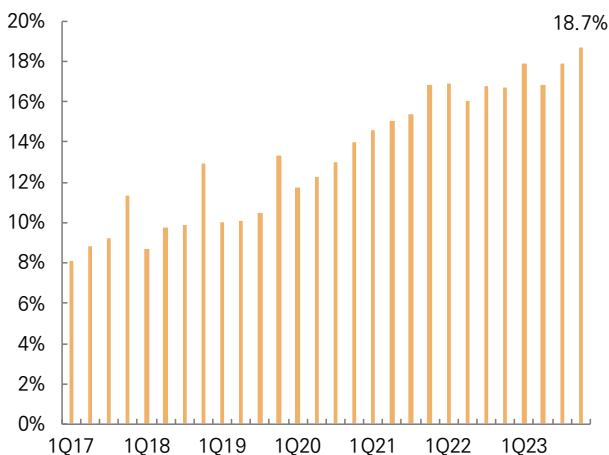
자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 66. Azure 매출 증가율



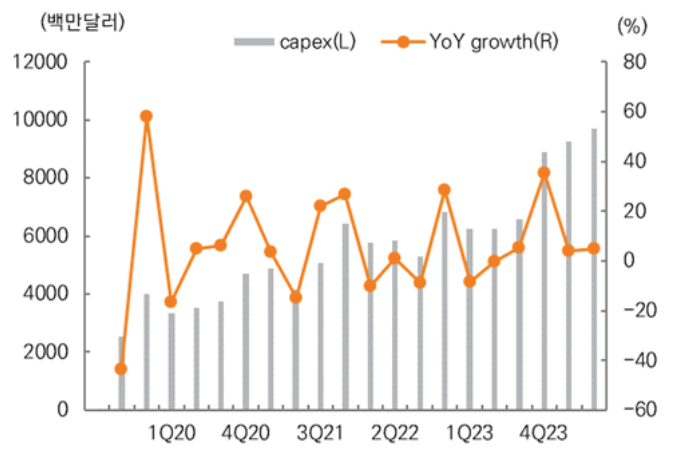
자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 67. Intelligent cloud opm



자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 68. CapEx



자료: 마이크로소프트, 미래에셋증권 리서치센터

1) Azure OpenAI Service

MS는 자사가 투자한 OpenAI가 개발한 GPT LLM을 통해 생성AI 솔루션을 개발하도록 하는 클라우드 툴 'Azure OpenAI' 서비스를 제공 중이다. Azure OpenAI Service를 통해 GPT-4, GPT3.5-turbo, embeddings 모델 등을 웹 기반 인터페이스(REST API, 파이썬 SDK 등)로 이용할 수 있다.

2) Copilot

① Microsoft 365 Copilot

MS는 9월 21일 'Microsoft Copilot'의 정식 런칭을 발표했다. Copilot은 AI 비서로 생산성 솔루션에서 사용된다. 가장 처음 Copilot이 쓰인 서비스는 2022년 개발 도구 Github Copilot이었으며, 올해 3월에는 오피스에서 사용될 'Microsoft 365 Copilot'과 보안에서 사용될 'Microsoft Security Copilot'도 공개되었다. 5월에는 'Windows Copilot'에 대한 언급됐고, 실제 제품 출시일이 9월에 발표된 것이다. Copilot은 윈도우 11, 마이크로소프트 365, 엣지 및 Bing 등에 적용될 예정이다.

우선 가장 기대되는 제품은 Microsoft 365 Copilot이다. 런칭 예정일은 11월 1일이다. 워드부터 사용할 수 있다고 설명했다. 한국어 버전은 내년 하반기로 예상된다. 가격은 1인당 월 30달러로 책정됐다. 현재 Microsoft 365 E3 가격이 36달러인데(E5는 57달러), 비교해서 추가로 30달러가 더해져 66달러로 83% 인상되는 것이다. 따라서 시나리오 분석에 따르면 이용자가 10% 증가할 경우, 오피스 매출 40~50억달러 증가가 예상된다.

② Microsoft Security Copilot

보안에서는 우선 챗봇 형태의 Copilot이 적용된다. 이 제품이 필요한 이유는, 사이버공격이 늘어나는 상황에서 보안 전문가를 챗봇이 대체할 수 있기 때문이다. 또 문제가 발생했을 때, 문제 발생 이유와 해결 방법을 챗봇에게 물어 해결책을 찾아낼 수 있다. 챗봇 형태로 다양한 보안 기능을 제공해줄 수 있다.

③ Windows Copilot

9월 25일 윈도우 11 업데이트를 통해 150개 이상의 신규 기능이 제공되는데, 이 때 코파일럿 연동, AI 기능이 추가된 그림판, 노트북패드, 자연스러운 음성을 지원하는 내레이터 등 AI가 추가된 기능이 대거 등장했다.

④ Bing과 Edge 브라우저

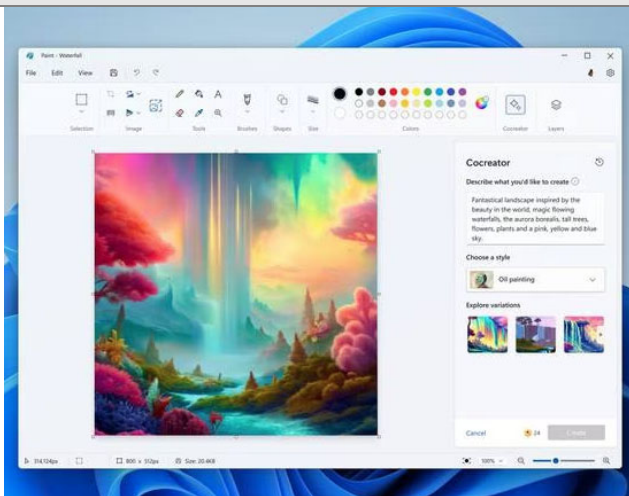
이미 3월부터 GPT를 Bing에서 사용할 수 있었으나, 이제 이미지 AI인 Dall-E3도 지원된다. 텍스트 검색을 통해 이미지를 생성할 수 있는 것이다. 또 모든 이미지에는 생성된 시간과 날짜를 포함해 암호화 된 디지털 워터마크가 추가된다. 이를 콘텐츠 자격 증명(content credentials)라 한다. 마이크로소프트 엣지 모바일 앱에서도 AI 기능을 사용 가능하며, Bing 엔터프라이즈에서는 멀티모달 비주얼 서치와 이미지 크리에이터 기능이 추가된다.

표 6. 마이크로소프트 사이버보안 제품

제품명	기능	대표 peer	대형 보안 기업 출시 제품
Microsoft Security Copilot	생성시로 위협 발견, 조사 및 대응	클라우드스트라이크(CRWD US) 외 등장 중	
Microsoft Sentinel	SIEM	Splunk(SPLK US), Elastic, Rapid7	아마존, 포티넷, IBM, 포티넷, 팔로알토네트웍스, 오라클,
Microsoft Defender	XDR	클라우드스트라이크(CRWD US), 센티널원(S US), 트런드마이크로(4704 JP)	포티넷, 시만텍
Microsoft Intune	모바일 디바이스 관리 보안	Jamf(JAMF US)	VM웨어, 시스코, Sophos, 시트릭스, AWS, IBM
Microsoft Entra	ID 보안	OKTA(OKTA US)	포티넷, 아마존, 구글, IBM
Microsoft Purview	데이터/거버넌스 보안	informatica(INFA US)	IBM, SAP

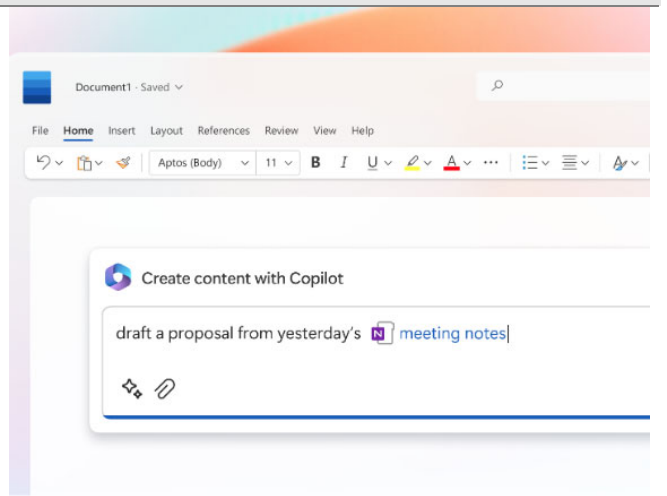
자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 69. 윈도우 그림판에 AI 기능 추가



자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 70. Microsoft Copilot



자료: 마이크로소프트, 미래에셋증권 리서치센터

표 7. Microsoft 365 Copilot 매출 시뮬레이션

(백만달러, 백만명)

	1Q24F	2Q24F	3Q24F	4Q24F
커머셜 구독자(백만명)	262.4621	265.0868	273.8504	274.7171
커머셜 구독자 * 월 30달러 * (a) = Copilot 이용 가정시 추가 발생 가능한 매출 (백만달러)				
(a)				
5%	394	398	411	412
10%	1,181	1,193	1,232	1,236
15%	5,315	5,368	5,545	5,563
기존 오피스 매출 예상치(백만달러)	14,629	14,928	15,573	16,102
YoY growth(%)	15.2%	12.6%	10.8%	9.6%
구독자 10% 증가시 매출 (백만달러)	15,810	16,121	16,806	17,338
YoY growth(%)	24.5%	21.6%	19.5%	18.0%

자료: 마이크로소프트, 미래에셋증권 리서치센터

3) OpenAI

현재 OpenAI의 기업가치는 8~900억달러 수준으로 알려졌다. 8월 기준으로 월 8,000만달러 매출이 발생하는 것으로 추정하고, 24년까지 연간 매출 10억달러를 목표로 하고 있다. OpenAI는 텍스트 모델에서 GPT-3.5에 이어 지난 4월 GPT-4까지 발표한 상황이다. 이 밖에 이미지 AI DALL-E, 음성 AI Whisper를 서비스 중이다. 기능적으로는 사전학습 된 모델에 추가 학습할 수 있는 파인튜닝, 텍스트를 숫자 형식으로 변환할 수 있는 임베딩, speech to text, plugin 등을 제공한다.

OpenAI는 올해들어 LLM을 수익화 할 수 있는 비즈니스 모델을 적극적으로 출시하고 있다. 우선 1) ChatGPT를 월 20달러 유료로 판매 중이고, 2) 복미 유료 이용자에 한해 Plugin 기능을 제공하고 있다. 이를 통해 plugin 어플리케이션으로부터 수수료를 받을 수 있다. 3) 지난 8월에는 ChatGPT Enterprise를 공개했다. 데이터 프라이버시, 서버 사용 등을 고심하는 엔터프라이즈에게 최적화 된 LLM을 제공하겠다는 것이다.

그림 71. ChatGPT 플러그인 대표 앱 Link Reader 사용 사례: 링크 url을 붙이면, 해당 링크 내용을 요약해줌

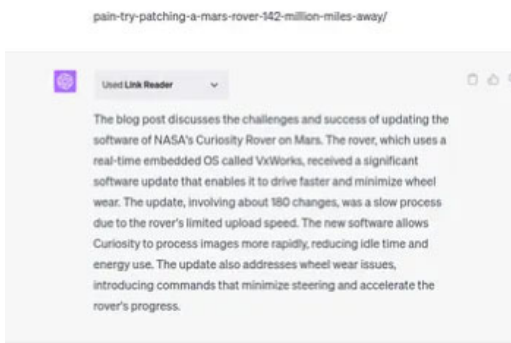
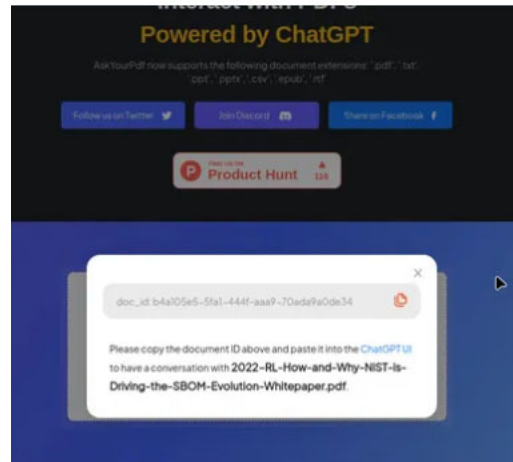


그림 72. ChatGPT 플러그인 대표 앱 Ask Your PDF 사용 사례: Ask Your PDF에 PDF를 업로드하면 PDF를 요약해줌



자료: ChatGPT, 미래에셋증권 리서치센터

자료: ChatGPT, 미래에셋증권 리서치센터

표 8. ChatGPT Plugin에서 가장 인기있는 대표 10가지 어플리케이션

기능	설명
Expedia	"7월 10일부터 14일까지 맨해튼 중심부에서 가장 좋은 3성급 호텔 요금은 얼마입니까?" 라는 질문에 Expedia 정보를 기반으로 대화 형태 답변
Ask your PDF	이 사이트에 PDF 업로드하고 문서ID를 받으면, ChatGPT한테 문서ID 입력 후 내용에 대해 질문 가능
Link Reader	링크를 제공하고, 링크 내용에 대해 요약 등 질문 가능
Open Table	"현재 A 지역에 예약 가능한 레스토랑 리스트를 달라"는 질문에 답변
Asking the right question	내가 원하는 내용을 적으면 완벽한 형태의 프롬프트로 수정해줌
VoxScript	유튜브 링크를 제시하면, 동영상의 스크립트를 제공해줌
WolframAlpha	수학, 물리학, 금융 등 전문 영역에 대한 질의응답 가능
What to Watch	내가 보고 싶은 콘텐츠를 어떤 OTT 등에서 볼 수 있는지 알려줌
Zapier	구글 캘린더, 슬랙 등을 기반으로, 플랫폼마다 정보를 이전시켜줌
Advanced-Data Analytics	CSV 파일, Excel Spreadsheet, SQL Database 등의 파일을 ChatGPT에 업로드/다운로드 한 뒤, 데이터 분석/차트 생성/수학 문제 해결 등을 할 수 있음

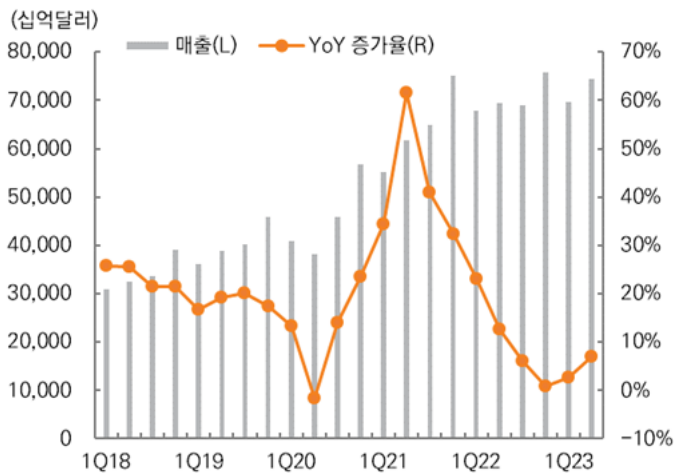
자료: OpenAI, 미래에셋증권 리서치센터

알파벳 (GOOGL US)

구글은 가장 뛰어난 AI 기술을 가지고 있는 기업이다. 현재 생성AI의 근간이 되는 Transformer 알고리즘 논문을 발표한 기업도 구글이었다. 구글은 AI 산업에서 주도권을 잡기위해 모든 방면에서 AI 관련 제품을 출시 중이다. **1) 검색 광고, 2) 클라우드, 3) 엔터프라이즈 솔루션 전 영역에서 관련 점유율 확대를 목표로 할 것으로 예상된다.**

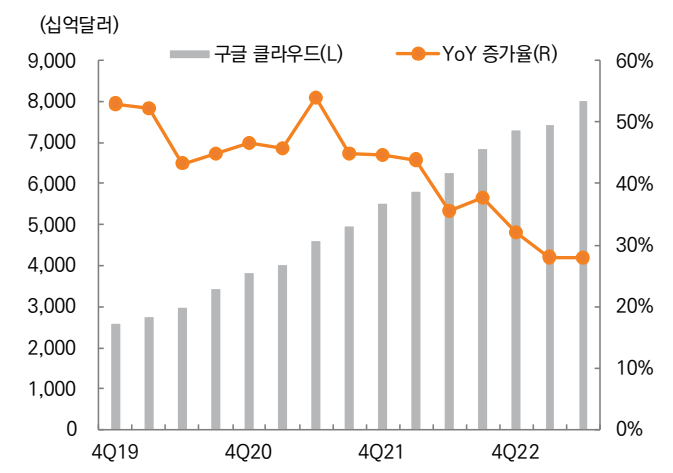
따라서 생성AI 시대에 가장 강력한 경쟁력을 가진 기업으로 판단하며, 검색 광고에 생성AI가 적용되면서 매출의 대부분을 차지하는 광고 수익 확대가 본격화될 수 있을 것으로 기대한다. 뿐만 아니라 구글 클라우드 매출의 반등과 영업이익의 지속적인 개선으로 생성AI 개발 플랫폼으로써의 역할도 강조될 수 있을 것으로 전망한다.

그림 73. 총 매출 추이



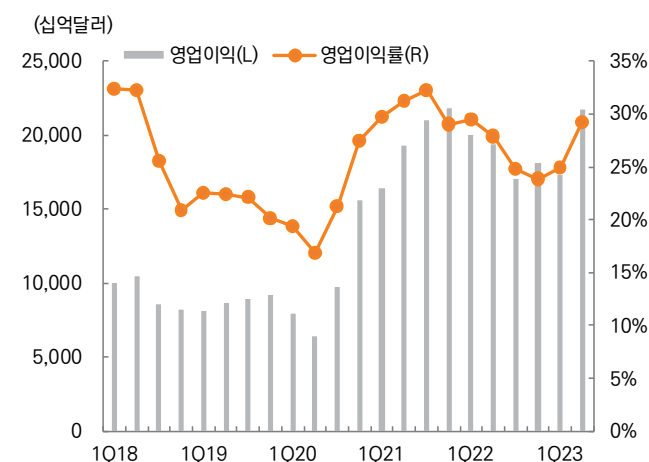
자료: 구글, 미래에셋증권 리서치센터

그림 74. 클라우드 매출 추이



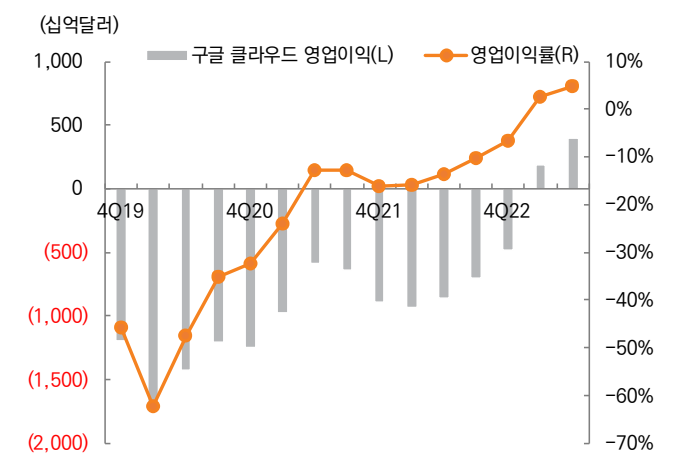
자료: 구글, 미래에셋증권 리서치센터

그림 75. 총 영업이익 및 영업이익률 추이



자료: 구글, 미래에셋증권 리서치센터

그림 76. 구글 클라우드 영업이익 및 영업이익률 추이



자료: 구글, 미래에셋증권 리서치센터

1) Palm2와 Bard, 그리고 Gemini

OpenAI의 GPT-4 등장에 긴장한 구글은 5월 ‘PaLM2’를 출시했다. 자사의 대화형 AI 서비스 ‘바드(Bard)’의 파운데이션 모델을 LaMDA에서 PaLM2로 교체한 것이다. PaLM2 파라미터수는 5,400억개고(이전 버전과 동일), TPU v4 인프라에서 구축됐다. 주목할 만한 점은 도메인별로 미세 조정된 ‘Sec-PaLM’과 ‘Med-PaLM2’이 함께 발표됐다는 점이다. Sec-PaLM은 보안에 특화된 모델로, 악성 스크립트 탐지에 사용될 수 있다. Med-PaLM2는 의학 지식으로 미세 조정(파인튜닝)을 한 모델이다. 기존 모델보다 의학 지식과 관련한 부정확한 추론을 9배나 감소시켰고, 임상 전문가에 근접한 답변을 줄 수 있는 것으로 알려졌다. 또 의료 영상 정보 합성 기능도 제공한다. 그 결과 미국 의사면허시험(USMLE) 질문의 85%를 맞췄고, 인도 의학 시험 질문 MedMCQA 데이터의 72.3%를 맞춰 합격선을 넘었다.

그 다음은 ‘Gemini’다. Gemini는 멀티모달 AI로 텍스트, 이미지와 기타 다른 데이터 종류를 모두 통합해 학습될 것이다. Gemini는 올 하반기 출시 예정으로 알려졌다. 10월부터 멀티모달 AI 경쟁이 본격화 될 것으로 전망한다.

표 9. 구글의 생성AI 모델

생성 AI 모델	분야	주요 특징	공개시기
LaMDA	LLM	구글의 대화형 언어 모델로, 2021 Google I/O Keynote에서 발표되어 초기 버전인 Meena부터 현재의 두번째 세대 모델인 LaMDA2까지 진화함 자연어 이해 및 생성 능력을 통해 다양한 주제에 대한 대화를 전달하는데 초점이 맞춰짐 Bard의 초기 버전에서 사용됨	05/2021
PaLM	LLM	구글의 대형 언어 모델로, 다양한 생성 AI 응용 프로그램에 사용됨 기존의 LaMDA 모델보다 규모가 크고 상식적인 추론 및 코딩 문제와 같은 작업을 부다 잘 다룰 수 있음 다중 대화 및 콘텐츠 생성과 같은 다양한 작업에 최적화된 모델과 요약, 분류 등과 같은 일반적인 작업에 최적화된 모델을 제공함 크기와 능력 측면에서 효율적인 모델을 시작으로, 나중에 다른 모델 및 크기를 추가할 예정 (PaLM 2.0)	03/2023
PaLM 2.0	LLM	Google I/O 2023에서 발표한 PaLM의 다음 세대 언어 모델 100개 이상의 언어 및 특화된 프로그래밍 언어 학습 통해 기존 모델보다 향상된 다국어 및 코딩 지원 여러 모델과 다양한 크기를 제공하며, 각 모델마다 다른 용도에 최적화됨 의료 분야 응용: 의료 전문 팀에 의해 학습된 Med-PaLM 2는 의학 텍스트로부터 정보를 추출하고 의학 분야의 질문에 답변할 수 있으며, 엑스레이와 유방 촬영과 같은 의료 정보를 종합하여 환자 치료를 개선하는 데 기여함 보안 분야 응용: Sec-PaLM은 보안 분야에 특화된 모델로, 사이버 보안 분석을 지원하고 악성 스크립트를 분석하여 보안 위협을 신속하게 탐지하는 데 기여함 개발자 및 클라우드 지원: PaLM 2 API는 개발자에게 제공되며, 클라우드 고객들은 PaLM 2를 이용하여 엔터프라이즈 수준의 보안과 프라이버시를 유지하며 사용 가능	05/2023
Gemini AI	멀티모달	Google DeepMind에서 개발중인 GPT-4 및 AlphaGo 기술을 결합한 LLM 다중 모달 기능으로 텍스트와 이미지를 처리하고 생성할 수 있음 AlphaGo에서 영감을 받아 강화 학습과 트리 탐색 기술을 활용하여 개발됨 PaLM 2.0과 마찬가지로 다양한 크기와 기능으로 출시되어 사용자의 다양한 요구에 맞게 선택 가능함	개발중

자료: 구글, 미래에셋증권 리서치센터

2) Vertex

Vertex AI는 기존 구글 클라우드의 머신러닝 플랫폼이다. 지난 7월에 생성AI 기반의 프로토타입 애플리케이션을 개발할 수 있게 해주는 서비스가 추가됐다. 제품명은 'Vertex AI Model Garden'과 'Generative AI Studio'다.

표 10. Vertex AI에서 사용 가능한 AI 모델

모델 유형	설명			
first party 모델	기본 모델	비전, 대화, 코드 생성, 코드 완성 전반에 걸쳐 구글의 멀티모달 모델을 활용	세부 모델	- 텍스트용 PaLM, 채팅용 PaLM - 텍스트 이미지 변환 위한 Imagen - 코딩용 Codey - 음성 텍스트 변환용 Chirp
	사전 학습된 API	구글이 사전 학습해 둔 API로 AI 어플리케이션 빌드 후 배포	세부 기능	- Speech to Text - 자연어 처리 - 번역 - 비전
오픈소스 모델		다양한 엔터프라이즈 지원 가능한 오픈소스 모델에 액세스	이용 가능 모델	- Meta의 Llama2 - TI의 Falcon - BERT, T-5 FLAN, ViT, EfficientNET
third-party 모델		기본 모델이 있는 파트너의 서드 파티 모델 지원	이용 가능 모델	엔트로픽의 Claude2

자료: 구글, 미래에셋증권 리서치센터

표 11. Vertex AI Model Garden의 기능

기능	설명
기본 모델 및 API 액세스	모델 가든에서 여러 모델 유형과 크기에 액세스 가능 다양한 모델도 선택 가능
자체 데이터로 모델 조정 가능	프롬프트 디자인 및 파인튜닝 가능한 인터페이스 제공
프롬프트 디자인으로 모델 실험 및 테스트 가능	프롬프트를 수동으로 만들어 모델 응답을 맞춤설정하는 프롬프트 디자인 기능 제공 예를 들어 응답의 '온도'를 설정 가능
end-to-end ML 도구와 통합	ML 지식 없이도 단순한 코드로 애플리케이션 제작 가능
엔터프라이즈급 데이터 거버넌스 및 보안	데이터를 사용해 모델을 맞춤 설정할 때, 완전하게 보호되고 또 비공개로 데이터 관리 가능 데이터의 위치와 방법, 사용 여부 등을 사용자가 제어 가능 입력 프롬프트는 원본 기본 모델 조정에 사용되지 않음 모든 데이터가 암호화 됨

자료: 구글, 미래에셋증권 리서치센터

표 12. Generative AI Studio 기능

이름	생성AI 기능	설명
Codey	코딩	실시간 코드 완성 및 생성 기능 구글 standard SQL, Java, JavaScript, 파이썬 등 20개 이상 코딩 언어 지원 개발자가 붓과 대화하여 디버깅, 문서작성 등 도움 받음
Imagen	이미지 생성 및 맞춤설정	대규모 이미지 생성 및 수정 300개 이상 언어로 마스크 없는 편집, 이미지 업스케일링, 이미지 캡션 기능 활용 가능
Chirp	범용 음성 모델	음성 작업에 대규모 모델 성능 제공 수백만 시간 분량의 오디오 학습한 Chirp은 100개 이상의 언어 지원 영어 정확도 98%, 1,000만명 미만의 화자 보유

자료: 구글, 미래에셋증권 리서치센터

3) 광고 알고리즘 적용 계획

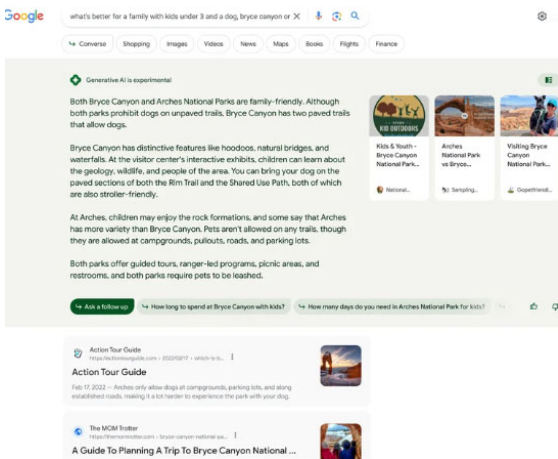
Search Generative Experience (SGE)

전세계 검색 광고 점유율의 90%를 지배하는 구글이 인공지능을 검색엔진에 탑재하겠다고 밝혔다. 우선 공개한 Search Generative Experience(SGE)의 방식은 MS의 검색엔진 Bing Chat과 유사하다. 챗봇 형태의 검색 엔진이다. 현재 'Search Labs'에서 테스트 해 볼 수 있다.

Google Ads에 새 기능 출시 - Video View Campaigns and Demand Gen

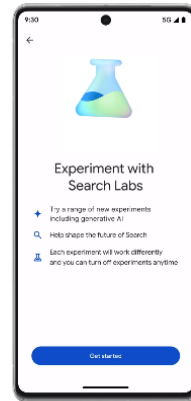
8월 Google Ads에 새 기능을 출시했다. Video View Campaigns은 비디오, 쇼츠 광고, 인피드 광고, 인스트림 광고 등 다양한 포맷을 결합해 한번에 제공하는 광고 포맷이다. 여기에 Demand Gen 기능을 통해 시를 접목해 타겟 고객에 따른 할인율 차별 제공, 구독 프로그램 제안 등의 고객 최적화 전략을 구사할 수 있게 한다. 구글은 시를 접목한 마케팅 초기 테스트에서 평균 40% 더 많은 조회수를 달성했다고 밝혔다.

그림 77. SGE 사례



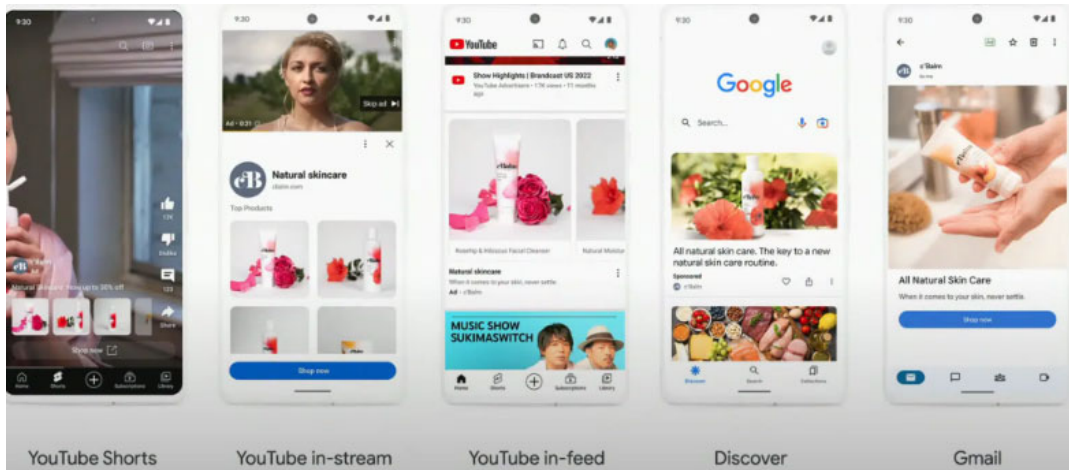
자료: 구글, 미래에셋증권 리서치센터

그림 78. Search Labs



자료: 구글, 미래에셋증권 리서치센터

그림 79. Google Ads의 Demand gen



자료: 구글, 미래에셋증권 리서치센터

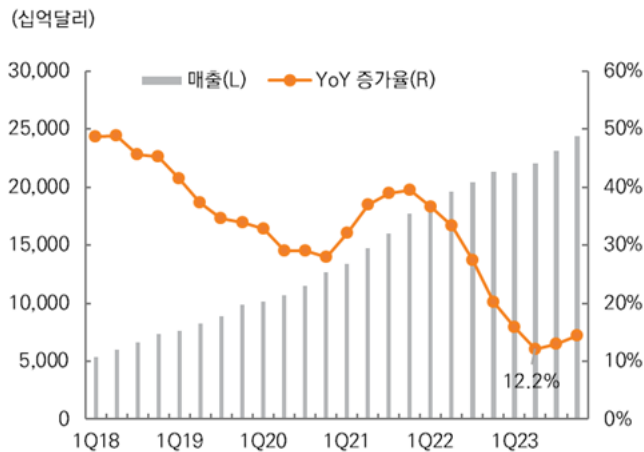
아마존 (AMZN US)

아마존은 LLM 보다는 자사의 강점인 클라우드 인프라를 통해 AI 개발 플랫폼 역할을 확고히 하려는 전략을 펼치는 중이다. 아마존은 기존 머신러닝 개발 플랫폼인 'Sagemaker'를 통해서도 생성 AI 개발을 할 수 있지만, 추가로 'AWS bedrock'을 선보여 생성 AI 에 최적화된 개발 플랫폼을 공개했다. Bedrock 의 특징은 여러 LLM 중에 개발자가 원하는 것을 선택할 수 있다는 점이다.

이 밖에 코드 생성 솔루션 'Codewhisperer'를 공개했고, 올 9 월 신제품 발표회에서 자사의 기존 AI 비서 '알렉사(Alexa)'에 LLM 을 접목시킨 버전을 공개했다. ChatGPT 및 바드와 가장 큰차이점은 텍스트 프롬프트 대신 음성으로 대화를 주고받을 수 있다는 점이다. 알렉사는 아마존의 '파이어 TV'와 스피커 '에코 8' 등에 탑재되며 IoT 하드웨어에서의 역할을 확장해 나갈 것으로 예상된다.

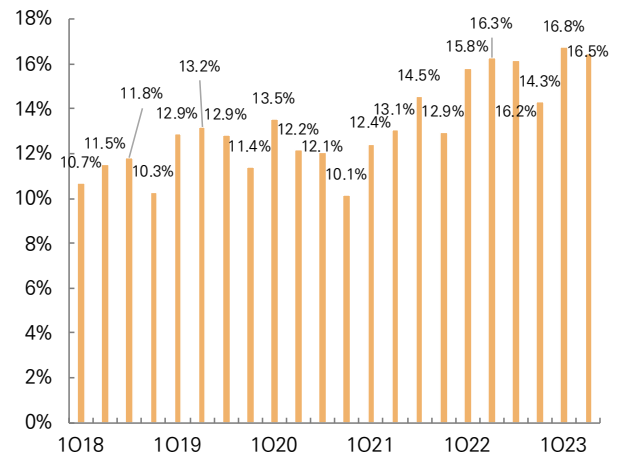
아마존의 전략이 아마존 AWS 사업부 매출 성장을 반등으로 연결될 것으로 기대한다.

그림 80. 아마존 AWS 매출



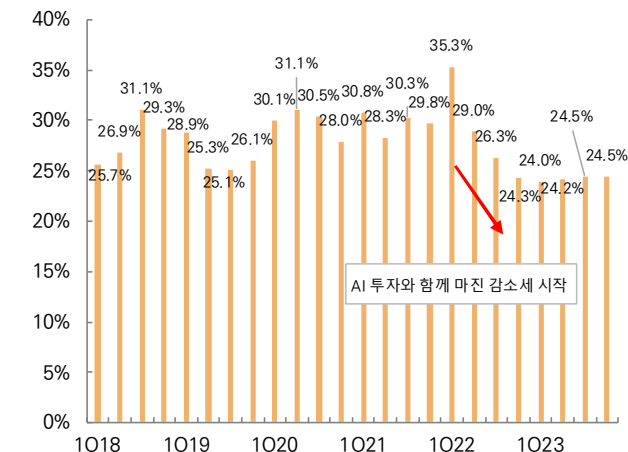
자료: 아마존, 블룸버그, 미래에셋증권 리서치센터

그림 81. 매출 대비 비중



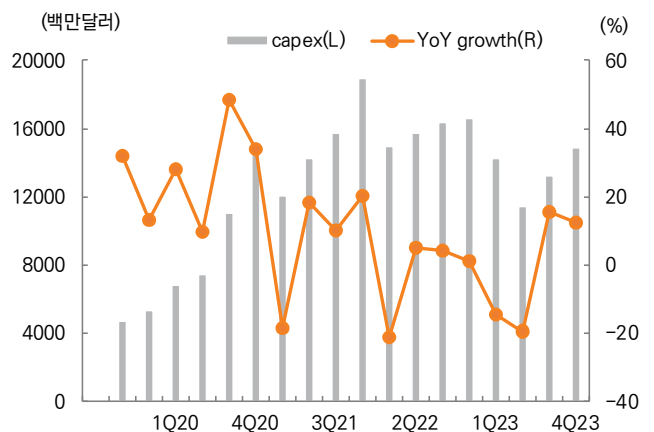
자료: 아마존, 블룸버그, 미래에셋증권 리서치센터

그림 82. 아마존 AWS opm



자료: 아마존, 블룸버그, 미래에셋증권 리서치센터

그림 83. CapEx



자료: 아마존, 미래에셋증권 리서치센터

표 13. 아마존 AWS 생성AI 솔루션 비교

제품명	주요 기능	특징
Amazon SageMaker	머신러닝(ML) 개발, 훈련, 구축 플랫폼으로, 아마존 bedrock보다 좀 더 데이터부터 모델까지 이용자에게 권한이 있음	
	머신러닝 라이프사이클 지원	데이터 라벨링부터 모델 구축, 모니터링까지 전부 지원
	알고리즘과 프레임워크 빌트인	가장 적합한 알고리즘 빌트인 가능
	오토매틱 모델 튜닝	수천개의 다른 알고리즘 파라미터 콤비네이션 조정을 통해 자동으로 모델 튜닝
	세이지메이커 스튜디오 통한 훈련/추론	웹 기반 인터페이스로 구축, 훈련, 추론 등 가능
	세이지메이커 JumpStart	ML 모델 수백 개 기본 알고리즘과 사전 훈련 모델이 있는 저장소
아마존 bedrock	생성AI 전용 클라우드 플랫폼	
	다양한 파운데이션 모델 사용 가능	Amazon Titan, Jurassic-2(멀티링구얼 텍스트 생성) Claude 2(대화 생성, 콘텐츠 생성), Stable Diffusion(이미지 생성)
	API로 통합할 수 있는 에이전트 기능	
	서버리스	서버리스로 API 이용 가능
	파인튜닝	S3 서버에서 이용 가능
	데이터 보호	가상 개인 PC(VPC)에 데이터 저장되고, 데이터는 모두 암호화됨
	기타 서비스 통합 가능	SageMaker Experiments(다른 모델 테스트), Pipelines
CodeWhisperer	코드 생성 서비스로 4월 런칭	월 19달러 부터이며, 지원되는 언어는 파이썬, 자바, 자바스크립트, C++ 등

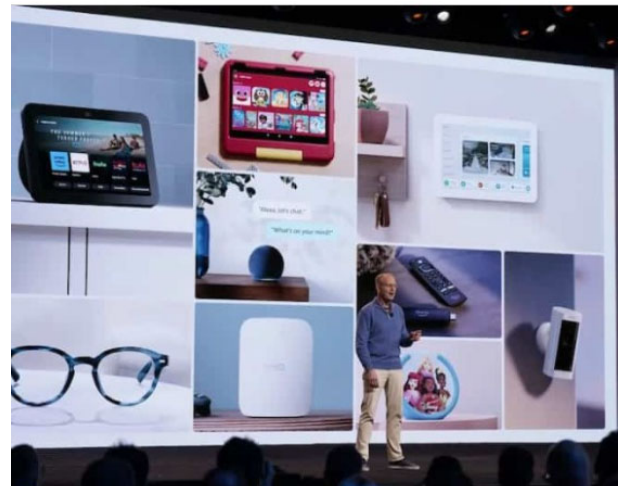
자료: 아마존, 미래에셋증권 리서치센터

그림 84. 아마존 음성비서 알렉사에 생성AI 탑재 발표(9.21)



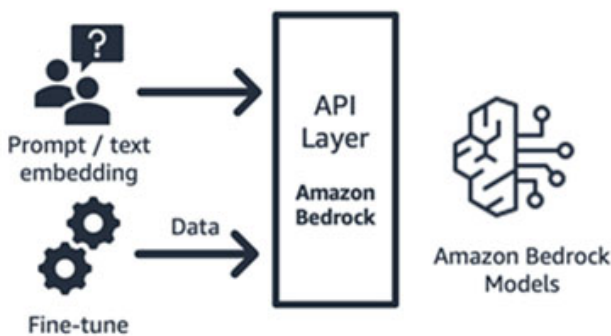
자료: 아마존, 미래에셋증권 리서치센터

그림 85. 아마존의 다양한 IoT 기기 통해 생성AI 구동 예정



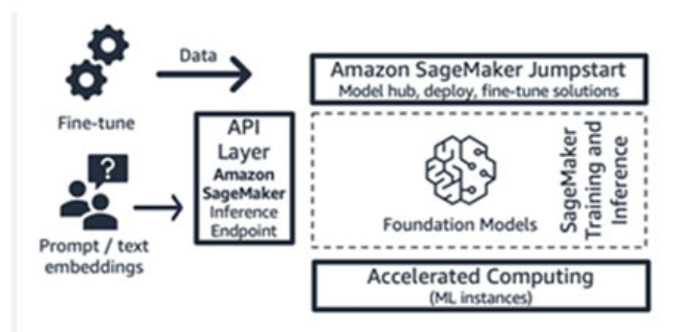
자료: 아마존, 미래에셋증권 리서치센터

그림 86. 아마존 bedrock으로 생성AI 이용 프로세스



자료: 아마존, 미래에셋증권 리서치센터

그림 87. 아마존 sagemaker 생성AI 이용 프로세스



자료: 아마존, 미래에셋증권 리서치센터

메타 Meta (META US)

메타는 페이스북 창업 초기인 2006년 뉴스피드에서 AI를 도입했었고, 2013년에는 AI 연구소 FAIR(Facebook Artificial Intelligence Research)를 출범했다.

메타는 현재 AI 산업에서 가장 중요한 위치에 놓여있다. 지난 7월 'LLaMa2'를 **상업적 사용이 가능한 오픈소스로 공개해버리며 생성AI 산업에 큰 반향을 불러왔기 때문이다.** LLaMa2는 3개의 모델 - 70억개, 130억개, 700억개-를 선택해서 사용할 수도 있다. 기존 버전보다 40% 많은 토큰으로 학습됐으며, 컨텍스트 길이는 2배 증가한 4,095로 설정됐다.

메타가 LLaMa2를 오픈소스로 공개한 이유는 1) 스탠다드의 주도권 확보, 2) 통합 가능성 확대, 또 3) 기업의 비용 절감, 4) 에코 시스템 구축, 5) 경쟁 제거 등의 뜻이 포함되어 있다.

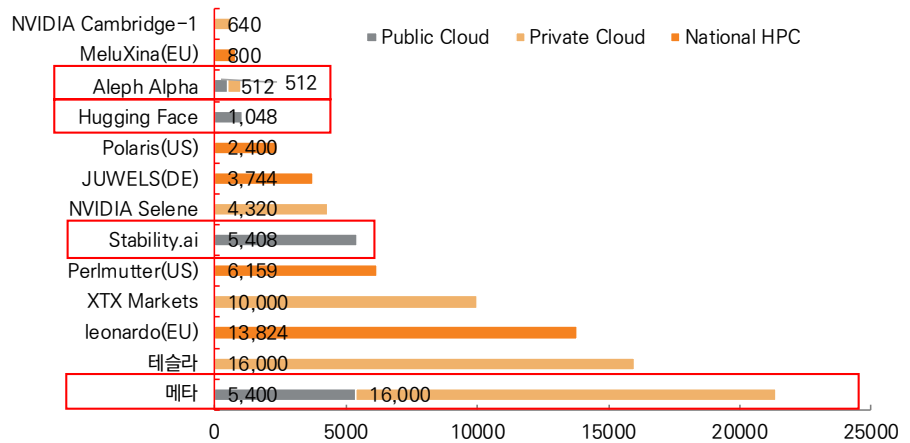
현재 LLaMa2는 아마존 AWS Sagemaker, MS Azure AI Studio, Hugging Face, Google Vertex 플랫폼에서 모두 개발 가능하다. 파인튜닝 역시 가능하기 때문에 원하는 애플리케이션 개발이 가능하다.

이어 메타는 8월 AI 코딩 비서 '코드라마(CodeLlama)'를 공개하며, 그동안 유료로 제공되던 AI 코딩서비스도 무료로 이용할 수 있게 만들었다. 코드라마는 커뮤니티 라이선스 기반의 오픈소스로 제공되고, 상업적 활용이 가능하다. 프롬프트 기반으로 코드를 생성하고, 특정 코드 문자열을 통해 코드를 완성하고 버그도 할 수 있다.

또 8월에 다중언어모달 AI인 '심리스M4T(SeamlessM4T)'도 공개했다. 음성 인식과 다른 언어로 된 음성을 텍스트로 번역하는 '음성-텍스트 번역', 입력된 음성을 번역해 음성으로 출력하는 '음성-음성 번역', 구글 번역과 유사한 기능인 '텍스트-텍스트 번역' 및 '텍스트-음성 번역' 기능이 가능하다. 텍스트 기능은 약 100개의 언어를 지원하며 음성 출력 기능은 약 36개의 언어를 지원한다.

이 AI는 총 47만시간 이상의 데이터로 구성된 멀티모달 데이터셋으로 학습됐다. 텍스트 데이터는 위키피디아, 뉴스 소스, 대본화 된 연설 등에서 추출된 문장 데이터를 사용했고 음성 데이터는 공개된 웹데이터에서 얻은 4백만 시간의 원본 오디오로부터 수집했다고 밝혔다.

그림 88. 기업별 GPU 개수 (2023년 9월 1일 업데이트)



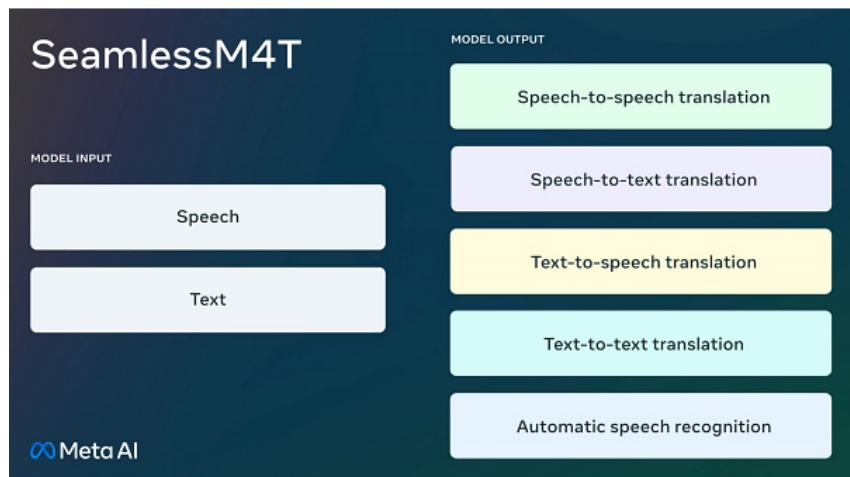
자료: state of AI Report Compute Index(2023.9.1 업데이트), 미래에셋증권 리서치센터

표 14. 메타의 생성AI

생성 AI 모델	분야	주요 특징	공개시기
LLaMa	LLM	<ul style="list-style-type: none"> - 메타의 대형 언어 모델로, 다양한 모델 크기로 출시되었으며, 7, 13, 33, 65 billion 파라미터 크기의 모델이 있음 - GPT-3(175B 파라미터)보다 더 우수한 성능을 보이며, 가장 큰 모델은 PaLM 및 Chinchilla와 같은 최신 모델과 경쟁할 만한 능력을 가짐 - 대규모 학습 데이터를 기반으로 학습되어, 이 데이터에는 다양한 공개 데이터 소스로부터 수집된 텍스트 정보가 포함되어 있으며, 이를 통해 모델은 다양한 주제 및 언어를 이해하고 처리할 수 있음 	03/2023
LLaMa-2	LLM	<ul style="list-style-type: none"> - Meta AI가 개발한 LLaMa의 다음 세대 모델로, 7, 13, 70 billion 파라미터 크기의 모델로 출시됨 - 마이크로소프트 Azure에서 사용 가능 - LLaMa-1과 거의 동일한 모델 아키텍처를 사용하며, 학습 데이터 양을 40% 늘림 - 대화용 모델인 "LLaMa-2 Chat"를 비롯하여 다양한 모델을 제공하며, 이 중 일부 모델은 상용 활용을 위해 무료로 제공됨 - 기본 모델로만 제공되었던 LLaMa1과 달리 fine-tuning을 통해 향상된 성능을 제공 - 대화의 일관성을 향상시키기 위해 'ghost attention' 기술 사용 - 'Code Llama'와 같은 특화된 모델을 통해 AI 코딩보트로서의 기능도 탑재 	07/2023
CM3Leon	이미지 생성 AI	<ul style="list-style-type: none"> - 텍스트와 이미지를 생성하기 위한 최첨단 생성 모델로, 이미지 슈퍼-해상도 및 이미지에서 텍스트 생성과 같은 작업을 수행할 수 있음 - Shutterstock와 같은 라이선스 이미지 및 캡션 데이터를 활용하여 보다 복잡한 텍스트와 객체를 파생하여 사용자 입력을 더 잘 따르도록 설계됨 	07/2023
Audiocraft	오디오 생성 AI	<ul style="list-style-type: none"> - 텍스트에서 고품질 음악 및 오디오 생성을 위한 AI 도구로, MusicGen, AudioGen 및 EnCodec의 세 가지 모델로 구성됨 - MusicGen은 메타 소유 및 특별히 라이선스 받은 음악을 기반으로 음악을 생성하며, AudioGen은 공개 사운드 효과 데이터를 활용하여 오디오를 생성함 - EnCodec 디코더는 음악 생성 품질을 향상시키며, 모델 가치치와 코드의 공개 통해 사용자의 참여도를 높임 	08/2023
Code LLaMa	AI 코딩봇	<ul style="list-style-type: none"> - Llama 2를 코드 관련 데이터로 추가 훈련한 모델로, 총 500조 개의 코드 토큰을 기반으로 Python, C++, Java, PHP, TypeScript, C#, Bash와 같은 주요 프로그래밍 언어를 지원함. 코드 생성, 자연어로 된 코드 설명, 코드 완성 및 디버깅을 위한 용도로 활용됨 - Code Llama - Python: Python 코드에 특화된 모델로, 100조 개의 Python 코드 토큰을 기반으로 함 - Code Llama - Instruct: 사용자 지시 및 안전한 응답 생성에 특화되어 있으며, 자연어로 된 사용자 입력에 따라 응답을 생성하는 데 주로 사용됨 	08/2023
SeamlessM4T	멀티모달	<ul style="list-style-type: none"> - 다중 언어 및 다중 작업 모델로, 텍스트 및 음성 간 번역 및 텍스트와 음성에 대한 전환을 자동으로 수행함 - 이 모델은 자동 음성 인식, 음성-텍스트 및 텍스트-텍스트 번역, 음성-음성 및 텍스트-음성 번역을 지원하며, 100개 이상의 언어를 다루는 데 활용됨 - 공개적으로 배포되어 연구 및 개발 활용을 위한 커뮤니티 라이선스를 제공하며, SpeechAlign과 같은 대규모 다중모달 번역 데이터셋도 함께 공개되었음 	08/2023

자료: 메타, 미래에셋증권 리서치센터

그림 89. 8월 새롭게 공개한 멀티모달 인공지능 심리스M4T



자료: 메타, 미래에셋증권 리서치센터

2. 수익화의 기로 - 1) LLM 효율화 진행 중

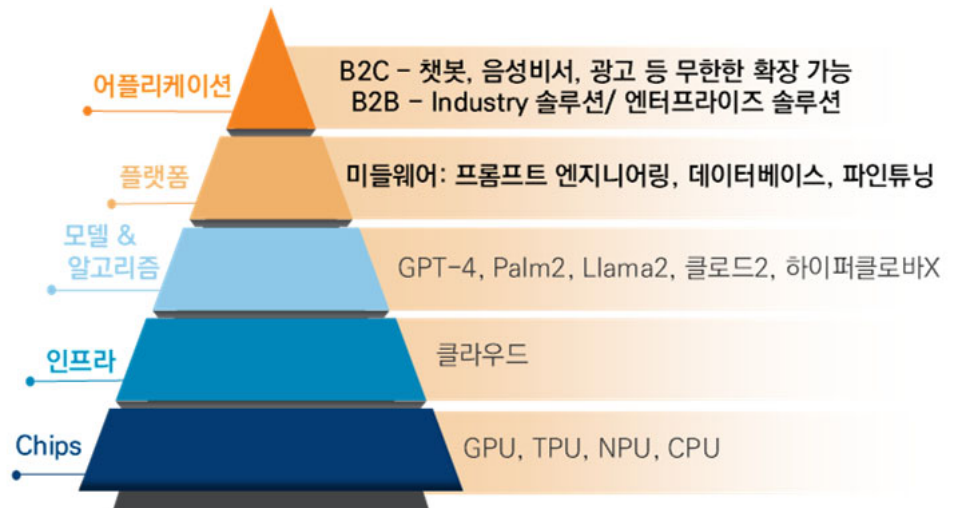
현재 기업들의 가장 큰 고민은 생성AI의 '수익화'다. 이를 위해 테크 기업들은 이용자의 니즈를 충족할 수 있는 비즈니스 모델을 만들어 나가는 것에 집중하고 있다. 생성AI 밸류체인은 LLM(Large Language Model 초거대언어모델)부터 시작되는데, 이를 응용 애플리케이션으로 만들기 위해서는 중간에 LLM을 구동 시켜 줄 수 있는 인프라 플랫폼이 필요하다. 바로 클라우드 컴퓨팅과 미들웨어다. 결국 실질적으로 클라우드 컴퓨팅과 미들웨어가 애플리케이션의 비용과 성능에 직결되기 때문에, 이들을 얼마나 효율적이고 적합하게 사용하느냐가 애플리케이션을 개발해 나가는 현 시점에 핵심이 될 수 있다.

애플리케이션은 가시적인 사용처를 찾을 수 있는 B2B 영역에서 먼저 등장하고 있다. 기업에서는 주로 사람들이 하는 업무를 자동화하거나, 절차를 간소화하는 것 등에서 생성AI를 활용하려고 시도 중이다. 뿐만 아니라 각 산업별로 반드시 필요한 특정 업무에서 생성AI를 적극적으로 도입하고 있다.

기업에서 생성AI LLM을 사용할 때 가장 염두하는 부분은 크게 세 가지다. 1) 파인튜닝 과 2) 보안, 3) 비용이다. 기업에서는 특정 업무나 산업에 특화된 용도에 필요한 AI의 역할을 찾는다. 그러다보니 범용적인 데이터 세트를 학습한 LLM은 불충분하다. 산업에 특화된 데이터를 추가로 학습해야 한다. 이 때 파인튜닝이 필요하다. 기업용 솔루션을 위해서는 파인튜닝이 필수적이다. 두번째로는 보안이다. 기업에서 LLM을 사용할 때 기업의 내부 정보를 프롬프트에 작성할 수도 있고, 내부 데이터 자체를 파인튜닝 해야 할 수도 있다. 이 때 클라우드 기반의 LLM에 데이터가 유출되는 위험을 기업들은 우려한다. 그리고 기업들은 무엇보다 비용에 민감하다. 최근의 초거대 모델들은 비용이 많이 들 수밖에 없다. 따라서 기업들의 니즈를 충족하는 제품들이 등장하기 시작했다.

파인튜닝
(Fine-Tuning, 미세조정)
사전 학습된(pre-trained) 인공지능 모델을 이용해 필요한 과제에 맞게 모델을 다시 학습시키는 것을 말함. 사전 학습된 모델의 가중치를 새로운 데이터에 맞게 세밀하게 조정하여 성능을 향상시키고 학습 시간을 줄이는 과정

그림 90. 인공지능 밸류체인

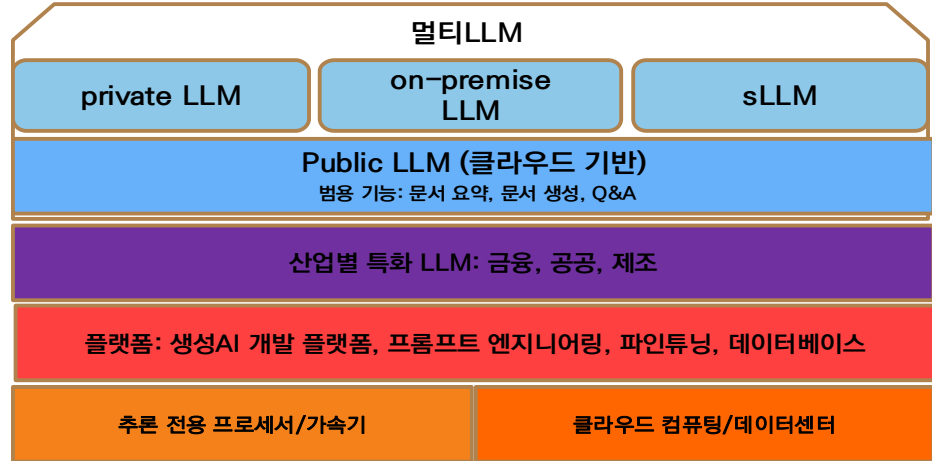


자료: 미래에셋증권 리서치센터

LLM 모델의 다양화

기업들이 원하는 방향으로 모델을 효율화하기 위해 기술적으로 다양한 형태의 LLM이 등장하고 있다. 대표적으로 private LLM, 온프레미스 LLM, sLLM이다.

그림 91. LLM의 밸류체인 다양화로 효율화 진행 중



자료: 미래에셋증권 리서치센터

Private LLM

Private LLM이란 개인 정보 보호 및 데이터 보호를 우선시하도록 설계된 언어 모델이다. 훈련 및 추론 중에 사용자의 데이터 노출을 최소화하는 것을 목표로 하는 기술로 구축됐다. Private LLM으로 데이터를 보호하는 방법으로 4가지가 있다.

- 연합학습 혹은 차등정보보호: Private LLM은 연합학습(federated learning)이나 차등정보보호(differential privacy) 등의 방법으로 데이터를 보호할 수 있다.

연합학습(federated learning)이란, 분산형 데이터 소스에서 모델을 훈련하는 방법이다. 기존 머신러닝은 하나의 고성능 서버에 대용량 데이터를 업로드해서 학습했다. 그러나 연합 학습은 여러 대의 컴퓨팅 자원에서 각각 알고리즘을 학습한 뒤, 그 결과를 통합해 최종 학습 모델을 만드는 방법이다. 대표적인 사례로 각 스마트폰에 저장된 데이터로 각각의 스마트폰에서 기계 학습을 진행한다. 학습된 모델을 압축하고 암호화한 뒤 클라우드에 전송한다. 클라우드에서 여러 스마트폰에서 학습된 정보가 결합되어 학습 모델이 개선될 수 있다. 이렇게 개선된 모델은 다시 개인의 스마트폰에 전송되어 기존 모델보다 더 정밀한 예측 모델이 되는 것이다.

차등정보보호(differential privacy)란 데이터 트레이닝 과정에서 데이터에 노이즈를 추가해 특정 사용자 정보 식별을 어렵게 만드는 방법이다.

- 암호화 및 보안 계산: 훈련 및 추론 단계 전반에서 암호화 및 보안 프로토콜을 반드시 사용한다. 데이터를 모델로 보내기 전에 암호화하고, 안전한 보안 기술을 사용해 데이터를 처리하는 것이다.

- 데이터 최소화: 데이터 수집 및 저장을 최소화 해 양을 줄이면 잠재적인 개인정보 침해 위험이 크게 줄어들 수 있다.

- 정기 감사 및 평가: 정기적인 개인 정보 보호 감사와 평가를 수립해 사립 LLM에 구현된 보안 및 개인 정보 보호 조치를 평가한다.

온프레미스 LLM

모든 데이터 처리 및 모델 작업을 자체 인프라에서 진행하는 LLM이다. 이 LLM은 내부 네트워크에서 작동해 강력한 네트워크 보안 조치도 구현 가능하다. 현재 OpenAI의 GPT 나 구글의 바드 등 우리가 잘 알고 있는 거대 파라미터의 LLM은 온프레미스로 제공될 수 없다. 온프레미스로는 감당할 수 없는 규모이기 때문이다. 따라서 온프레미스 LLM은 범용적인 기능보다는 한정된 기능을 제공하는 소형 LLM을 선택할 때 사용할 수 있다.

그림 92. 연합학습 사례

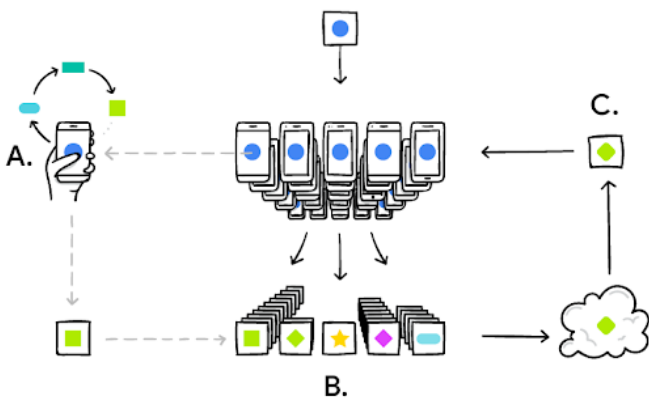


그림 93. 차등정보보호 사례



자료: 구글리서치, 미래에셋증권 리서치센터

자료: 텐서플로우, 미래에셋증권 리서치센터

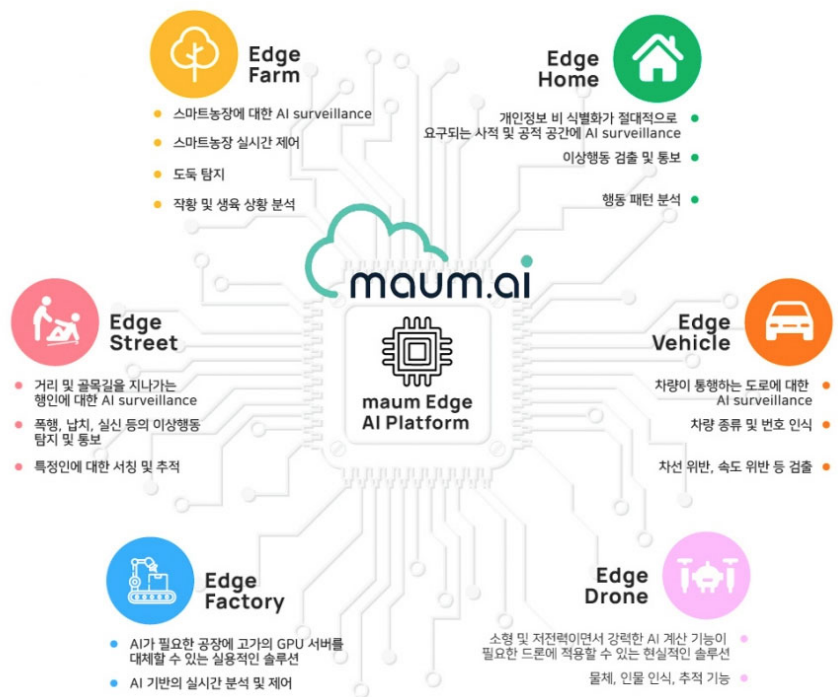
sLLM (엣지LLM)

기업에서는 메모 및 프레젠테이션 초안 작성, 이메일 편집, 질문 답변, 사본 요약 등을 위해 LLM 기반 앱을 채택하는데, LLM을 활용하여 효율성을 높이고 비용을 절감하는 것이 목표지만 사용량이 증가함에 따라 비용이 기하급수적으로 증가하게 된다. 파라미터가 1,000억개 넘는 LLM을 구동시키는 것은 매우 고사양의 인프라를 요구하기 때문에 비용이 많이 들기 때문이다.

이를 위해 등장한 것이 sLLM이다. 다른말로 엣지 LLM이라고도 부른다. 특정 요구 사항을 가진 사용자 정의에 맞게 소규모로 구축하는 LLM을 뜻한다. GPT-4나 Bard의 파라미터가 1,000억개가 넘는 것과 달리 10억개 수준의 파라미터를 가질 수 있다. 그러므로 당연히 데이터 학습에 사용되는 컴퓨팅 비용이 절감된다.

이 LLM은 산업별 데이터 세트를 교육함으로써 특정 요구 사항에 맞게 모델을 맞춤화하고, 이를 통해 AI 투자에서 최대 가치를 창출할 수 있다. 뿐만 아니라 소규모 데이터 세트를 학습하면 편향된 결과나 비윤리적 답변도 오히려 줄일 수 있다. 보안에서도 더 안전해진다. 예를 들어 최근 '구글 I/O'에서 구글은 PaLM2를 Gecko, Otter, Bison, Unicorn 4가지 크기로 세분화해 출시한다고 밝혔다. 사용 사례에 맞게 모델 크기를 선택해 쉽게 배포할 수 있게 한 것이다. Gecko는 가장 작은 사이즈의 모델로 모바일에서도 이용 가능해, 삼성 갤럭시를 통한 테스트에서 초당 20개 토큰 전달이 가능했다.

그림 94. 국내 마인즈랩이 출시한 엣지 LLM 플랫폼



자료: 미래에셋증권 리서치센터

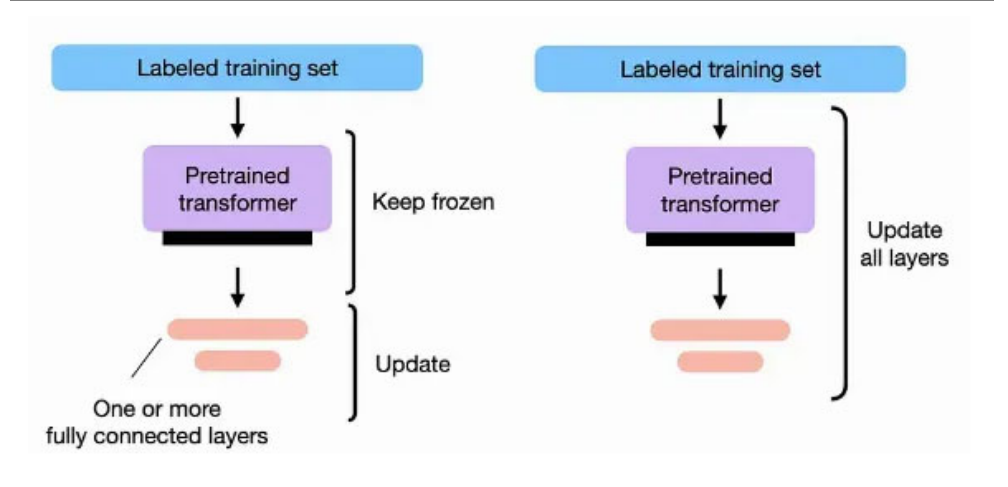
파인튜닝(Fine-tuning, 미세조정)

파인튜닝(Fine-Tuning, 미세조정)이란 사전 학습된(pre-trained) 인공지능 모델을 이용해 필요한 과제에 맞게 모델을 다시 학습시키는 것을 말한다. 사전 학습된 모델의 가중치를 새로운 데이터에 맞게 세밀하게 조정하여 성능을 향상시키고 학습 시간을 줄이는 과정이다.

LLM을 특정 업무나 기능을 위해 사용하기 위해서는, pre-trained model을 특정 태스크에 맞게 파인튜닝 해야 하는 것이 필수적이다. 하지만 파인튜닝을 할 때, 모든 weight를 다시 학습시키게 되면 지금처럼 파라미터가 1,000억개 이상의 초대형 모델의 경우 파인튜닝 작업에만 몇 달이 소요되고 엄청나게 많은 GPU가 소모된다.

그래서 등장한 개념이 PEFT(Parameter Efficient Fine-Tuning)이다. 적은 양의 파라미터만 학습해서 유사한 성능을 가져오는 것을 목표로 하는 방법론이다. 대표적으로 adapter, LoRA, Prompt tuning, Prefix tuning 등의 방법이 있는데, 최근 가장 주목받는 방법 중 하나는 21년 마이크로소프트가 발표한 논문 LoRA다.

그림 95. 파인튜닝 방법



자료: analytics Vidhya, 미래에셋증권 리서치센터

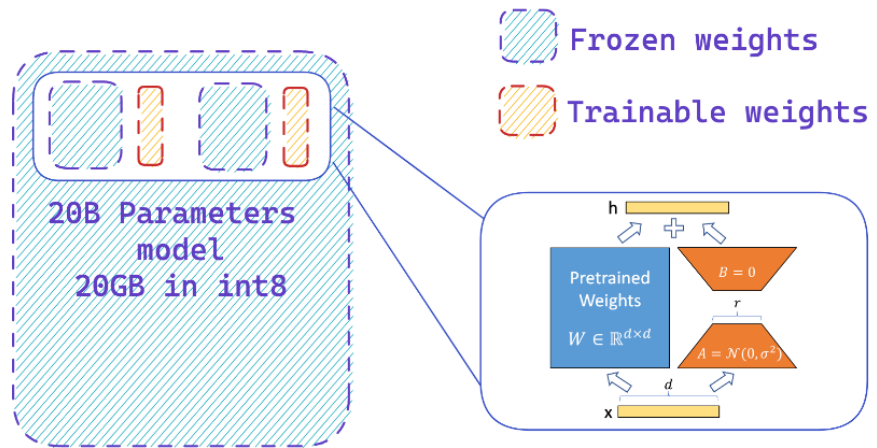
LoRA(Low-Rank Adaptation)

LoRA의 방법은 이렇다. 거대한 파라미터를 가진 LLM을 매번 학습시키는 대신, 이 pre-trained model은 weight를 고정시켜두고, 내가 필요한 데이터는 학습 가능한 랭크 분해 행렬을 사용해 파인튜닝 시켜 추가 학습되는 파라미터수를 줄이는 모델병렬 기술력이다. 이 방법을 쓰면, 전체 LLM 파라미터를 기반으로 파인튜닝 할 때와 유사한 결과값을 얻게 된다.

핵심은, 기존 LLM Model은 RAM에 저장해 CPU inference만 진행하고, LoRA의 가중치만 GPU에 저장해 학습을 진행한다. 이에 LoRA를 사용하면 VRAM 사용량을 대폭 감소시킬 수 있고, 체크포인트 크기를 줄여 GPU 수도 줄일 수 있다. 또 파인튜닝 작업을 수행할 때 그래디언트 계산이 필요하지 않아 효율성이 극대화된다.

GPT-3의 경우, 사용하는 VRAM 사용량이 1.2TB에서(A100 80GB 1200개) 350GB(A100 80GB 4~5개)로 줄었다. 또 25%의 속도 향상이 있었다. 또 추론 과정에서 추가적인 latency가 없다는 점도 특징이다. 한 사례에서는 A100 GPU 하나로 30분만에 튜닝을 했다고 한다.

그림 96. LoRA의 알고리즘



자료: hugging face, 미래에셋증권 리서치센터

어플리케이션 개발 활성화를 위한 미들웨어 지속 등장 중

1990년대 초반에 네트워크, 클라이언트-서버 및 여러 계층의 아키텍처가 등장하면서 통합적으로 발생하는 문제를 해결하고 개발을 가속화하는 솔루션에 대한 필요성이 커졌다. 이에 컴퓨터와 어플리케이션 사이의 계층에서 비즈니스 프로세스와 데이터 통신 및 통합을 촉진하는 미들웨어가 개발되기 시작했다. 생성AI가 클라우드 컴퓨팅, 웹 서비스, API와의 연동 등과 밀접하게 연관되며 미들웨어는 점점 더 중요한 존재로 부상했다. 생성AI가 IT 환경 전반에 걸쳐 여러 어플리케이션과 시스템을 연결하는 강력한 도구 역할을 하기 위해서는 통합, 콘텐츠 분석, 요청-응답, 작업 성능 및 어플리케이션/시스템 업데이트 등이 원활히 이뤄져야하므로, 이를 담당하는 미들웨어가 부상 중이다. 생성AI의 대표적인 미들웨어로 랭체인(LangChain)과 마이크로소프트가 서비스하는 시멘틱 커널(Symantic Kernel)이 있다.

랭체인(LangChain)

랭체인이란 LLM과 어플리케이션의 통합을 돕는 SDK다. 기존 프로그래밍 언어와 통합이 가능하다. 몇 줄의 코드로 기존 프로그래밍 언어와 연결되는 플러그인이 작동한다. 랭체인과 같은 프레임워크가 필요한 이유는 개발자가 AI를 사용해 어플리케이션을 만들 때 직면하는 문제들을 해결해주기 위함인데, 대표적으로 1) LLM에 단기 및 장기 메모리를 추가하는 것은 각 기업 개발자가 선택할 문제, 2) LLM에 대한 규칙을 만들어야 하는 것 - 매개변수 조정, 프롬프트 보강, 응답 조정부터 감정 분석, 분류, 질문과 답변 요약 등 적절한 상황에 맞는 모델 선택 등을 랭체인 프레임워크로 해결할 수 있다. 랭체인의 개발자는 해리슨 체이스로 2022년 10월 말 처음 랭체인을 선보였다. 랭체인이 지원하는 언어는 파이썬과 자바스크립트다.

시멘틱 커널(Symantic Kernel)

시멘틱 커널은 마이크로소프트가 오픈소스로 제공하는 SDK 프레임워크로, 랭체인과 같은 역할을 한다. 랭체인과 몇 가지 차이점이 있는데, 지원하는 언어가 C#, 파이썬, 자바 세 가지다.

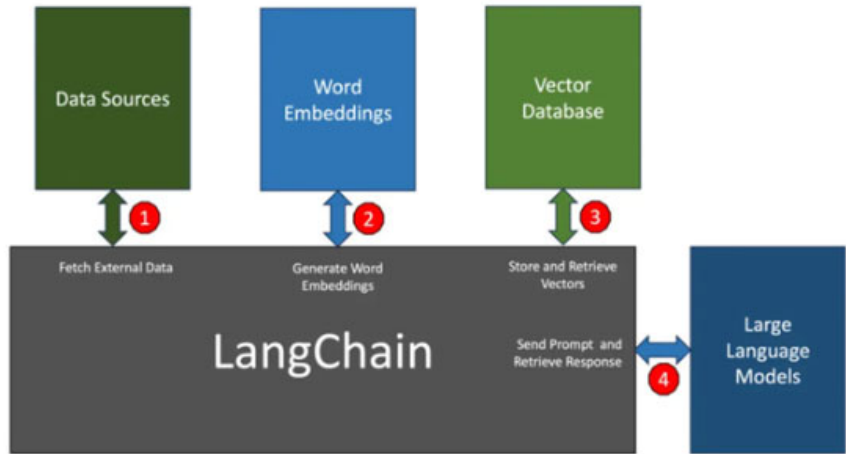
SDK
(Software Development Kit)
소프트웨어를 개발할 수 있는 도구
및 라이브러리 집합

표 15. 랭체인 대표 기능

기능	설명
데이터 소스	어플리케이션이 LLM에 대한 컨텍스트를 구축하기 위해 PDF, 웹 페이지, CSV, 관계형 데이터베이스와 같은 외부 소스에서 데이터를 검색해서 가져와야하는데, 랭체인은 서로 다른 소스에서 데이터에 액세스하고 검색할 수 있는 모듈과 원활하게 통합해줌
단어 임베딩	일부 외부에서 검색된 데이터는 벡터로 변환되어야 LLM과 관련된 단어 임베딩 모델에 전달하게 됨
벡터 데이터베이스	생성된 임베딩은 유사성 검색을 위해 벡터 데이터베이스에 저장됨. 랭체인은 메모리 내 배열부터 파인콘(Pinecone)과 같은 호스팅 벡터 데이터베이스에 이르기까지 다양한 소스에서 벡터를 쉽게 저장하고 검색할 수 있도록 지원
LLM	오픈AI, 코히어(Cohere), AI21에서 제공하는 주류 LLM과 허깅페이스(Hugging Face)에서 제공되는 오픈소스 LLM을 지원

자료: IDC, 미래셋증권 리서치센터

그림 97. 랭체인 의 흐름도



자료: 랭체인, 미래에셋증권 리서치센터

표 16. 랭체인 의 모듈 설명

모듈	설명
모델 I/O	모델 I/O 모듈은 LLM과의 상호 작용을 하는 영역. 기본적으로 효과적인 프롬프트를 생성하고, 모델 API를 호출하고, 결과를 해석함. 생성형 AI의 핵심인 프롬프트 엔지니어링이 랭체인에서 처리되는 것. 또 이 모듈은 LLM 제공자가 노출하는 인증, API 매개변수, 엔드포인트를 요약. 마지막으로, 모델에서 보낸 응답을 애플리케이션서 사용할 수 있는 원하는 형식으로 해석하는 역할을 함
데이터연결	이는 랭체인 의 가장 중요한 구성 요소로, LLM 애플리케이션의 ETL 파이프라인에 해당. 이 모듈은 PDF 또는 엑셀 파일과 같은 외부 문서를 로드하고, 이를 처리하기 위해 일반적으로 단어 임베딩으로 변환한 다음, 임베딩을 벡터 데이터베이스에 저장하고, 마지막으로 쿼리를 통해 검색함
체인	랭체인 의 체인은 구성 요소와 LLM을 활용하여 예상되는 응답을 얻는 효율적인 파이프 라인을 구축하도록 설계됨. 간단한 체인에는 프롬프트와 LLM이 포함될 수 있지만 재귀와 같이 LLM을 여러 번 호출하여 결과를 얻는 등 매우 복잡한 체인을 구축할 수도 있음. 예를 들어, 체인은 문서를 요약한 다음 이에 대한 감정 분석을 수행하는 프롬프트가 포함될 수 있음
메모리	LLM은 상태 비저장형이지만 정확한 응답을 위해 컨텍스트가 필요함. 랭체인 의 메모리 모듈은 모델에 단기 및 장기 메모리를 쉽게 추가할 수 있도록 도와줌. 단기 메모리는 간단한 메커니즘을 통해 대화의 기록을 유지. 메시지 기록은 레디스(Redis)와 같은 외부 소스에 저장되어 장기 메모리를 유지
콜백	랭체인은 개발자에게 LLM 애플리케이션의 다양한 단계에 연결할 수 있는 콜백 시스템을 제공. 이는 로깅, 모니터링, 스트리밍 및 기타 작업에 유용. 파이프라인 내에서 특정 상황이 발생할 때 호출되는 사용자 지정 콜백 핸들러를 작성할 수 있게 해줌
에이전트	LLM은 추론과 행동이 가능한데, 이는 ReAct 프롬프트 기법이라고 불린다. 랭체인 의 에이전트는 LLM을 사용하여 프롬프트를 행동 계획으로 추출하는 ReAct 프롬프트 제작을 단순화함

자료: IDC, 미래에셋증권 리서치센터

그림 98. 랭체인 사용 샘플

Chain #1 - 사용자가 좋아하는 영화 장르에 대해 묻는 LLM 체인입니다. Chain #2 - Chain 1의 결과를 기반으로 선호하는 장르를 사용하여 해당 장르의 영화를 추천하는 다른 LLM 체인입니다.

먼저 chain_one을 생성합니다.

```
template = '''당신은 영화 장르 설문 조사입니다. 사용자에게 선호하는 장르에 대해 물어보세요.
문기:'''
prompt_template = PromptTemplate(input_variables=[], template=template)
chain_one = LLMChain(llm=llm, prompt=prompt_template)
```

그런 다음, chain_two를 생성하여 영화를 추천합니다.

```
template = '''당신은 영화 추천 프로그램입니다. 사용자가 선호하는 장르: {genres}, 이 장르에 해당하는 영화를 추천하.
추천:'''
prompt_template = PromptTemplate(input_variables=["genres"], template=template)
chain_two = LLMChain(llm=llm, prompt=prompt_template)
```

다음으로, 이러한 체인을 SimpleSequentialChain을 사용하여 결합합니다.

```
from langchain.chains import SimpleSequentialChain
overall_chain = SimpleSequentialChain(
    chains=[chain_one, chain_two],
    verbose=True)
```

자료: Kanaries, 미래에셋증권 리서치센터

3. 수익화의 기로 - 2) 결국 Vertical 접근 중요

사이버보안

생성AI의 등장으로 사이버보안 산업의 구조적 확대를 예상한다. 스스로 진화하는 생성AI의 특성 때문에 보안 방법도 늘어나고, 해킹 공격도 늘어나고 있기 때문이다.

대표적으로 하드웨어 잠금을 해킹하는 '탈옥'이 늘어날 수 있는데, 1) 생성AI에 윤리적 제약을 거스르는 응답을 하도록 행동을 극적으로 바꾸도록 지시하는 '스위치 기법', 2) AI 모델이 특정 캐릭터 역할과 행동 및 응답 세트를 가정하도록 요구하는 '캐릭터 플레이' 등의 방법을 사용할 수 있다. 이 밖에도 피싱, 랜섬웨어 등 대표적인 사이버공격이 생성AI를 통해 증가할 수 있다.

공격이 늘어나지만, 그만큼 사이버보안 대응 방법도 다양해질 수 있다. 이미 사이버보안 산업에서 AI를 통해 보안 성능을 향상시키는 기술이 많이 등장했으나, 생성AI 도입으로 한단계 업그레이드될 수 있을 것으로 기대한다. 특히 최근에는 전문가 대신 챗봇을 통해 사이버 보안 문제를 해결하는 솔루션들이 많이 등장하고 있다.

표 17. 사이버보안 공격 증가 사례

기능	생성AI 사용 내용
사회 공학 공격 (Social Engineering Attacks)	개인을 심리적으로 조작하여 민감한 정보를 공개하도록 유도하는 것이며, ChatGPT와 같은 AI 모델은 이러한 유형의 공격에 활용
피싱	신뢰할 수 있는 객체로 가장하고 피해자로부터 민감한 정보를 추출하는 것을 목표로 하는데, ChatGPT와 같은 고급 AI 시스템을 악용하여 더 효과적이고 감지하기 어려운 피싱 시도 가능
자동 해킹 시스템	시스템 취약성을 식별하고 이를 악용하는 전략을 개발하기 위해 AI 모델을 활용할 수 있으며, PentestGPT와 같은 윤리적 목적으로 개발된 AI 도구도 있지만, 이러한 원리조차 악용될 가능성
공격 페이로드 생성	공격자가 ChatGPT와 같은 AI 모델을 사용하여 공격 페이로드를 생성할 수 있는 능력을 갖추고 있으며, 이를 통해 다양한 위협을 빠르게 만듦
랜섬웨어 및 악성코드	ChatGPT와 같은 강력한 AI 모델을 사용하여 자동화될 수 있으며, 악의적인 소프트웨어를 더 빠르게 생성 사례1) - ChatGPT를 사용하여 REvil 공격과 관련된 코드를 생성하려고 시도하였으며, 이를 통해 네트워크 스캔 및 공격 특징과 같은 특정 기능에 대한 코드를 생성하는 능력을 보여줌 사례2) - 애드웨어, 스파이웨어 및 트로이 목마에 대한 코드도 ChatGPT를 통해 생성할 수 있습니다. 애드웨어는 광고를 통해 전송되고 스파이웨어는 사용자에게 제공되며 트로이 목마는 합법적인 소프트웨어로 위장 사례3) - Meltdown 및 Spectre와 같은 CPU 아키텍처를 공격하는 특정 바이러스에 대한 코드 예제도 ChatGPT를 통해 생성 가능

자료: IEEE Access, 미래에셋증권 리서치센터

표 18. 사이버보안 방어 증가 사례

기능	생성AI 활용 내용
사이버 방어 자동화	ChatGPT는 사이버 보안 사고의 분석을 자동화하고 전략적인 권고 사항을 제공하며 보안 교육을 지원하는 데 큰 역할 가능. 보안 운영 센터(SOC) 분석가의 업무 부담을 줄이고 전반적인 보안 수준을 향상시킬 수 있음
사이버 보안 보고	ChatGPT는 사이버 보안 데이터와 사건을 기반으로 자연어 보고서를 생성하여 이를 이해하기 쉽게 다양한 이해 관계자에게 전달
위험 인텔리전스	ChatGPT는 대량의 데이터를 처리하여 잠재적인 보안 위협을 식별하고 실행 가능한 인텔리전스 보고서를 생성함으로써 조직이 보안 수준을 향상하는 데 도움
안전한 코드 생성과 감지	코드 검토에서 보안 버그를 감지하고 안전한 코딩 관행을 제안하며 안전한 코드를 생성하는 데 도움
사이버 공격 식별	보안 관련 데이터를 분석하고 공격 패턴 및 행동의 자연어 설명을 생성하여 조직이 위협에 대응하는 데 도움
윤리적 가이드라인 개발	기존의 윤리적 프레임워크를 요약하고 교육 및 훈련을 위해 윤리적 시나리오를 생성함으로써 AI 시스템에 대한 윤리적 가이드라인 개발을 지원
사이버 보안 기술 향상	침입 탐지 시스템에 통합되어 잠재적인 위협이 감지될 때 실시간 경고와 알림을 제공하여 보안 기술의 효과성을 향상
사고 대응 가이드	보안 사고 중에 즉각적인 지침을 제공하고 자동으로 사건을 문서화하여 사건 대응을 더욱 신속하게 수행하는 데 도움
악성 코드 탐지	코드 동작을 분석하고 보고서를 생성하여 위협을 완화하기 위한 조치를 권장함으로써 잠재적인 악성 코드를 감지하는 데 활용

자료: IEEE Access, 미래에셋증권 리서치센터

사이버보안 시장은 27년까지 연평균 12%의 성장률을 기록할 것으로 예상되며, 23년과 24년에는 14%의 성장률을 기록해 평균적인 IT 지출의 성장률 3~4%를 훨씬 웃돌 것으로 전망한다. 이에 전체 글로벌 IT지출에서 사이버보안 매출이 차지하는 비중도 22년 3.8%에서 27년 4.8%로 증가할 것으로 예상된다. 그 중에서도 클라우드 보안, 어플리케이션 보안, ID 접근 보안 등의 카테고리 성장세가 기대되고, 최근 생성AI 등장으로 데이터 프라이버시 및 보안의 성장률이 가장 가파를 것으로 전망한다.

표 19. 사이버보안 산업 세부 카테고리별 성장률

세부 시장	2021	2022	2023	2024	2025	2026	2027	CAGR 2022~2027
Application Security	4,550.4	5,047.6	5,765.2	6,670.3	7,632.8	8,642.1	9,619.6	13.8%
YoY growth(%)		10.9%	14.2%	15.7%	14.4%	13.2%	11.3%	
전체 대비 비중(%)	3.1%	3.1%	3.1%	3.1%	3.2%	3.3%	3.3%	
Cloud Security	3,619.2	4,487.4	5,616.7	7,002.6	8,656.3	10,625.9	12,811.1	23.3%
YoY growth(%)		24.0%	25.2%	24.7%	23.6%	22.8%	20.6%	
전체 대비 비중(%)	2.4%	2.7%	3.0%	3.3%	3.6%	4.0%	4.4%	
Data Privacy	1,027.9	1,129.2	1,338.7	1,667.3	1,970.8	2,300.8	2,667.5	18.8%
YoY growth(%)		9.9%	18.5%	24.6%	18.2%	16.7%	15.9%	
전체 대비 비중(%)	0.7%	0.7%	0.7%	0.8%	0.8%	0.9%	0.9%	
Data Security	2,532.2	3,072.9	3,692.1	4,333.3	4,989.8	5,596.5	6,171.2	15.0%
YoY growth(%)		21.4%	20.1%	17.4%	15.2%	12.2%	10.3%	
전체 대비 비중(%)	1.7%	1.9%	2.0%	2.0%	2.1%	2.1%	2.1%	
Identity Access Management	12,271.5	13,944.1	16,169.1	18,556.5	20,753.3	22,781.5	24,863.4	12.3%
YoY growth(%)		13.6%	16.0%	14.8%	11.8%	9.8%	9.1%	
전체 대비 비중(%)	8.2%	8.5%	8.6%	8.6%	8.6%	8.6%	8.6%	
Infrastructure Protection	20,089.7	24,089.0	28,359.6	33,319.6	38,172.1	42,832.4	47,396.2	14.5%
YoY growth(%)		19.9%	17.7%	17.5%	14.6%	12.2%	10.7%	
전체 대비 비중(%)	13.5%	14.6%	15.1%	15.5%	15.9%	16.2%	16.4%	
Integrated Risk Management	4,706.4	5,157.3	5,687.1	6,277.7	6,696.8	7,052.4	7,392.6	7.5%
YoY growth(%)		9.6%	10.3%	10.4%	6.7%	5.3%	4.8%	
전체 대비 비중(%)	3.2%	3.1%	3.0%	2.9%	2.8%	2.7%	2.6%	
Network Security Equipment	16,924.4	18,932.5	21,383.6	24,360.1	27,285.6	30,235.1	33,103.7	11.8%
YoY growth(%)		11.9%	12.9%	13.9%	12.0%	10.8%	9.5%	
전체 대비 비중(%)	11.4%	11.5%	11.4%	11.3%	11.3%	11.4%	11.5%	
Other Information Security Software	5,349.1	8,029.8	11,365.4	14,362.8	16,520.4	17,502.5	18,456.7	18.1%
YoY growth(%)		50.1%	41.5%	26.4%	15.0%	5.9%	5.5%	
전체 대비 비중(%)	3.6%	4.9%	6.0%	6.7%	6.9%	6.6%	6.4%	
Security Services	70,641.9	73,394.7	80,835.7	89,996.7	98,884.9	107,717.7	116,255.8	9.6%
YoY growth(%)		3.9%	10.1%	11.3%	9.9%	8.9%	7.9%	
전체 대비 비중(%)	47.4%	44.6%	43.0%	41.9%	41.1%	40.7%	40.3%	
Consumer Security Software	7,234.7	7,443.4	7,901.7	8,406.7	8,847.9	9,300.9	9,759.4	5.6%
YoY growth(%)		2.9%	6.2%	6.4%	5.2%	5.1%	4.9%	
전체 대비 비중(%)	4.9%	4.5%	4.2%	3.9%	3.7%	3.5%	3.4%	
Total	148,947.4	164,728.0	188,114.8	214,953.7	240,410.7	264,587.7	288,497.2	11.9%
YoY growth(%)		10.6%	14.2%	14.3%	11.8%	10.1%	9.0%	

자료: 가트너(2023.7), 미래에셋증권 리서치센터

4. 통신 서비스

통신사들은 AI를 통해 새로운 도약을 꿈꾸고 있다. 5G 도입으로 컨슈머를 대상으로 하는 기존 유무선 사업부의 성장이 한계에 봉착하자 새로운 성장 동력으로 AI에 눈을 돌리고 있다. 글로벌은 물론 국내 통신사들도 생성AI 산업에 대한 대비를 철저히 하는 중이다. 특히 데이터가 핵심인 AI 산업 특성상, 통신 산업은 그 이점을 누릴 수 있다. 국내 통신3사는 모두 생성AI 모델을 직간접적으로 보유하고 있다. SK텔레콤과 KT는 직접 생성AI를 개발했다.

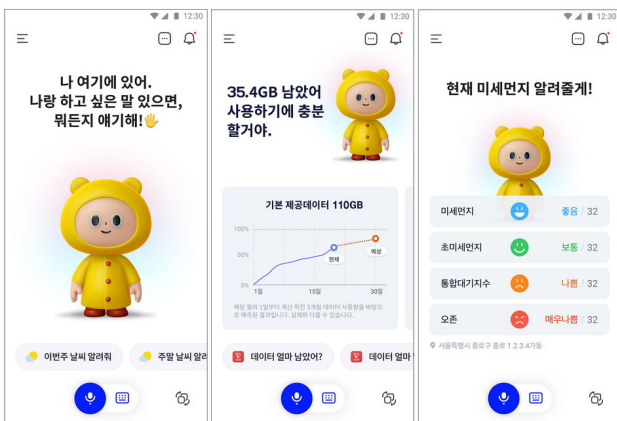
SK텔레콤은 9월 대대적인 ‘AI 피라미드 전략’을 발표하고, 자사 생성AI 플랫폼 ‘에이닷’을 공식 런칭했다. 직접 개발하는 LLM ‘에이닷엑스’도 확정 발표하면서 ‘엔트로픽’ 등 투자한 클라우드 LLM과의 투트랙 전략을 공고화했다. 향후 AI 투자를 전체 투자 대비 33%로 늘리고, 매출 비중도 2030년까지 36%로 늘리겠다고 밝혔다.

KT는 지난해 초거대AI ‘믿음’을 발표, 23년 연내 발표 예정이라고 밝혔다. KT는 생성AI 뿐만 아니라 AI 반도체 개발사 ‘리벨리온’ 투자를 통해 풀스택으로 AI를 제공한다는 목표를 발표했다. LG유플러스는 LG 그룹사가 개발 중인 ‘엑사원’을 활용할 계획이다.

통신사는 일단 생성AI로 가장 대표적인 통신사 서비스인 콜센터에 AI를 적용한 AICC 사업을 확장하는 중이다. 뿐만 아니라 최근에는 음성 통화를 텍스트로 전환하는 서비스 등 기존 가입자 락인 효과를 강화시킬 수 있는 서비스들을 하나씩 출시 중이다. 통신사들은 활용할 수 있는 데이터가 다양한 만큼 여러가지 서비스를 통해 유무선 가입자 해지율 감소 및 신규 가입자 유치, ARPU 증가 계획 등을 발표할 것으로 예상된다.

통신 산업 탐픽으로 SK텔레콤을 추천한다. 에이닷 뿐만 아니라 엔트로픽 투자 등을 통해 기술력을 확보했을 뿐만 아니라, 글로벌 얼라이언스를 통해 해외 판매 전략도 구축하고 있다. 또 사피온으로 AI 반도체 개발을 진행, 데이터센터를 기반으로 한 클라우드 컴퓨팅과 함께 통합적인 AI 서비스가 가능할 것으로 기대한다.

그림 99. SK텔레콤의 생성AI 어플리케이션 에이닷



자료: SK텔레콤, 미래에셋증권 리서치센터

그림 100. KT의 AI 풀스택



자료: KT, 미래에셋증권 리서치센터

V. 인터넷/게임

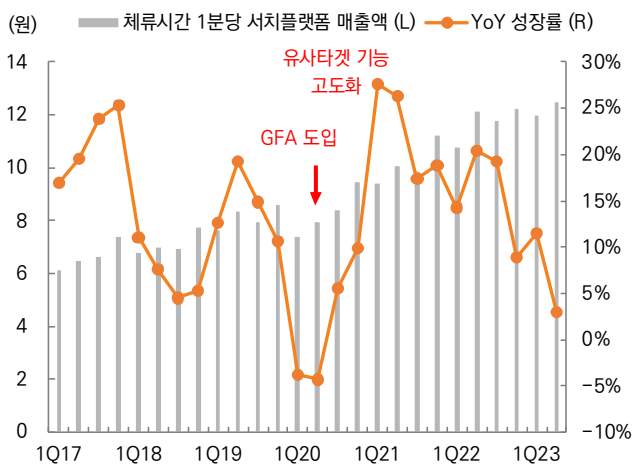
1. 광고: 타겟팅 고도화 및 수수료 BM 등장 예상

타겟팅 고도화에 따른 단가 상승

자연어 단위로 검색이 이루어지는 생성 AI 검색 도입은 광고 플랫폼 업체들의 타겟팅 능력을 고도화시킬 것이며 이는 곧 광고 단가 상승으로 이어질 전망이다. 질문 과정에서 노출되는 구체적인 조건들은 광고 플랫폼들이 검색자가 원하는 정보를 찾아주는 일의 난이도를 대폭 낮춰준다. 생성 AI 검색자는 서비스 또는 물품에 대한 구매의 목적성이 높기에 기존 대비 수배 이상 되는 구매 전환율을 보여주게 될 것이다. 빠른 목적 달성에 따른 검색 쿼리 및 체류시간(Q)이 일부 감소할 수 있으나 광고 단가(P) 상승폭이 이를 충분히 상쇄할 것으로 판단한다.

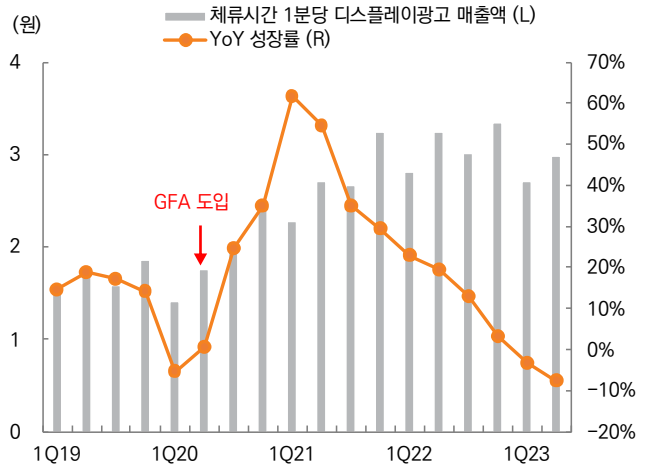
생성 AI 기반 광고를 통한 신규 매출은 1H24부터 본격화될 전망이다. 생성 AI 기반 광고 도입이 늦어지고 있는 글로벌과 달리 NAVER는 8월 CLOVA for AD를 일찌감치 공개하며 광고 상품의 대대적인 변화를 예고했다. CLOVA for AD에서 광고주는 생성 AI를 원하는 방향으로 정보를 학습시켜 검색자와의 대화 중 적절한 맥락에서 광고를 제공할 수 있게 된다. NAVER는 2020년에도 성과형 디스플레이광고(GFA)를 도입하고 타겟 기능을 고도화 하면서 광고 단가를 상승시킨 경험이 있다. 이번 생성형 AI 광고가 도입된 이후에도 광고 단가가 YoY 20% 이상으로 성장세가 가속화될 것으로 전망한다.

그림 101. NAVER 체류시간 1분당 서치플랫폼 매출액 추정



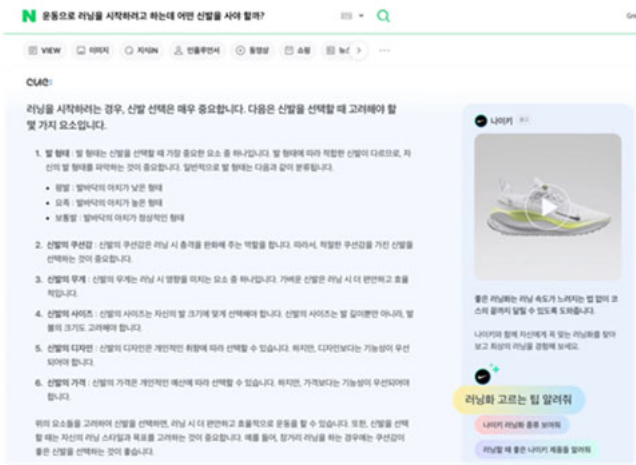
자료: 미래에셋증권 리서치센터 추정

그림 102. NAVER 체류시간 1분당 DA 매출액 추정



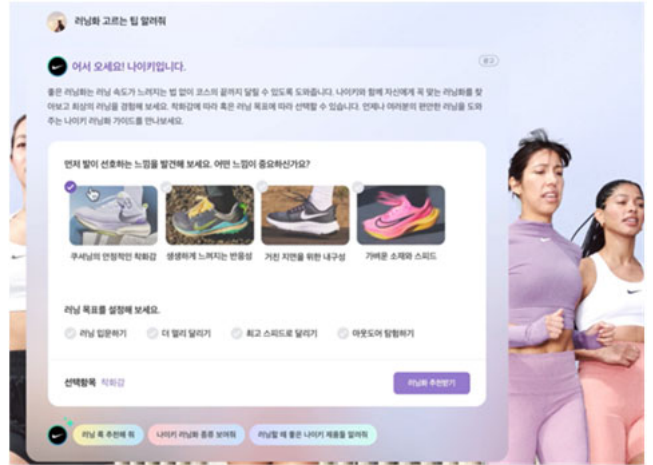
자료: 미래에셋증권 리서치센터 추정

그림 103. CLOVA for AD 예시1



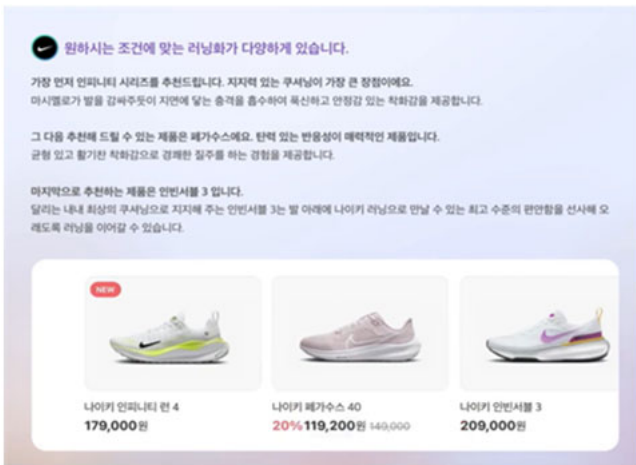
자료: NAVER, 미래에셋증권 리서치센터

그림 104. CLOVA for AD 예시2



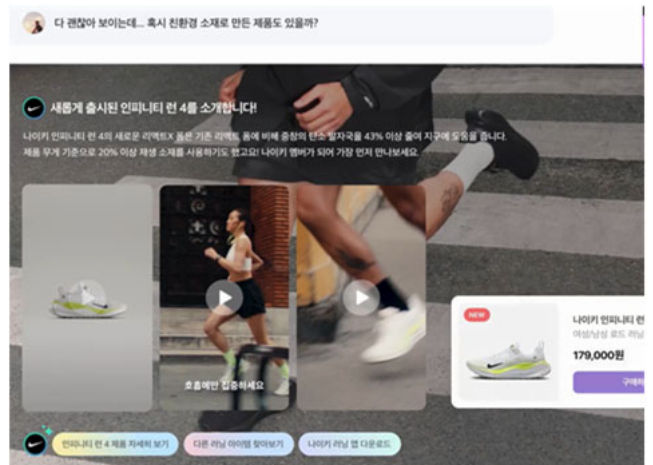
자료: NAVER, 미래에셋증권 리서치센터

그림 105. CLOVA for AD 예시3: 조건에 최적화된 상품 추천



자료: NAVER, 미래에셋증권 리서치센터

그림 106. CLOVA for AD 예시4: 대화를 통한 상품 추천



자료: NAVER, 미래에셋증권 리서치센터

그림 107. CLOVA for AD 예시5: 타겟층 세부 설정



자료: NAVER, 미래에셋증권 리서치센터

그림 108. CLOVA for AD 예시6: 브랜드 말투 학습



자료: NAVER, 미래에셋증권 리서치센터

중개 및 판매 수수료 BM 추가
가능성

챗봇에 광고가 추가되기 시작하면 구독 외에도 중개 및 판매수수료 BM이 추가적으로 가능할 것으로 판단한다. 그리고 다양한 버티컬 서비스를 보유하고 있는 플랫폼의 챗봇 서비스가 가장 큰 경쟁력을 보여줄 것이다. 챗봇을 업무에 보조하는 용도로 주로 사용하는 사람들도 있을 것이나 대다수의 사람들은 챗봇을 어떠한 계획을 세우거나 무엇을 구매하기 위해 사용할 가능성이 높다.

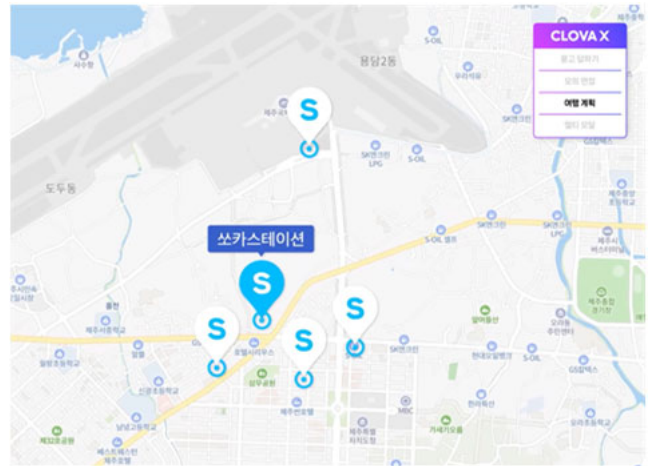
계획은 여행에 대한 계획(행선지, 숙소, 교통편, 티켓 등), 장소 방문에 대한 계획(음식점, 병원, 헤어샵 등) 등이 대표적인 것이다. 계획 역시도 상품 또는 서비스 등 무언가의 구매로 이어진다. 챗봇은 구매로 이어질 가능성이 높은 창구이기에 필연적으로 광고 수요가 강하게 발생할 것이다. 개별 카테고리에서의 신규 광고 매출 외에도 실제로 구매까지 연결되었을 경우 일정 수준의 Take rate를 수취하는 등의 중개 및 판매 수수료가 신규로 발생할 전망이다.

그림 109. 클로바X 예시 1



자료: NAVER, 미래에셋증권 리서치센터

그림 110. 클로바X 예시 2



자료: NAVER, 미래에셋증권 리서치센터

그림 111. 클로바X 예시 3

9월 30일에서 10월 4일 일정으로 제주도 가는데 30만원대 서귀포 숙소 5곳 추천해줘

- 제주도 여행을 계획 중이시군요! 고객님의 여행 일정에 맞는 숙소를 추천해 드리겠습니다.
1. 제주롯데호텔: 시설과 뷰가 좋으며 수영장과 조식이 매우 좋은 곳입니다.
 2. 시절인연: 감성적인 분위기와 자쿠지가 있는 숙소로, 가격도 숙소 퀄리티에 비해 생각보다 괜찮은 편입니다.
 3. 더그랜드 성오름 호텔: 수영장과 룸 컨디션이 좋으며, 부대시설과 가격이 착한 숙소입니다.
 4. 플라밍고 다이 라이 리조트: 숙박 + 왕복 픽업서비스 + 석식 + 조식 구성으로 20만원대 후반으로 예약이 가능하며 독채 숙소와 수영장을 이용할 수 있습니다.
 5. 롯데호텔: 시설과 뷰가 좋으며 소핑몰과 붙어있는 것이 장점입니다. 미리 예약하지 않으면 풀북일 수 있으니 미리 예약하는 것이 좋습니다. 1박 가격은 3월 평일 기준으로 20만원대부터 있습니다.
- 사용자님의 여행 일정과 선호하는 숙소 유형을 말씀해 주시면 더욱 정확한 추천을 해드릴 수 있습니다.

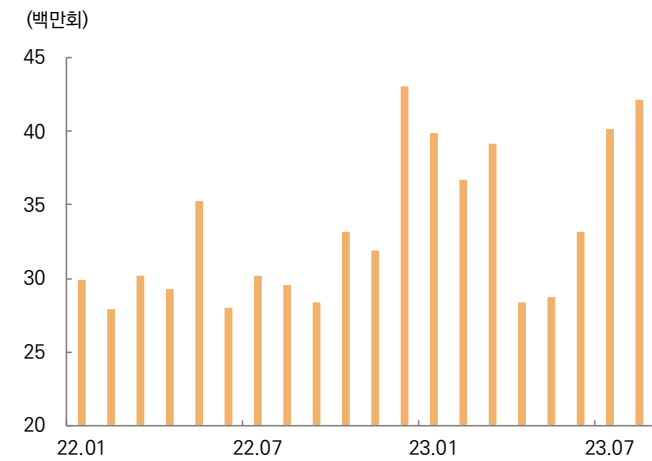
자료: 본인 화면, 클로바X, 미래에셋증권 리서치센터

SNS 플랫폼의 희생 가능성

생성 AI의 도입은 한계에 다다른 침투율, 광고 타겟팅 능력 약화로 인해 부진을 겪고 있는 SNS 플랫폼 업체들에 희망이 될 수 있다. 챗봇 외에도 생성 AI 기반의 이미지 생성, 카메라 촬영 등 다양한 콘텐츠를 플랫폼에 확대해 플랫폼 체류시간 확대 및 인앱 결제 매출 증가를 노려볼 수 있다. 체류시간 증가는 광고 단가 상승으로 이어진다.

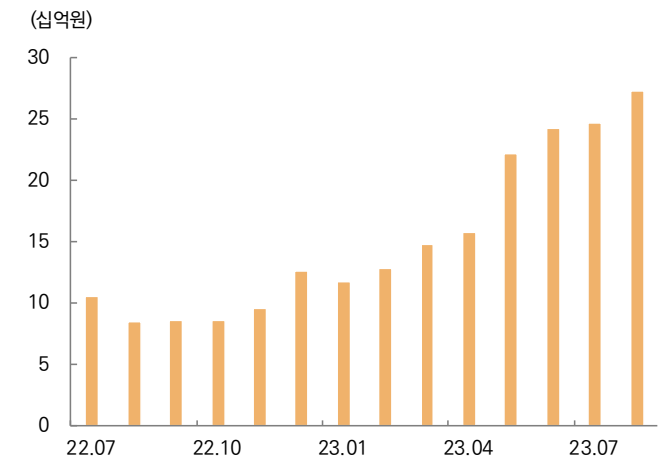
스냅은 지난 3월 자체 챗봇인 마이AI(My AI)를 베타 출시하고 광고를 테스트하고 있으며 8월에는 AI 생성 셀카인 드림스(Dreams)를 출시했다. 마이AI를 출시한 초기 흑평이 이어지기도 했으나 2Q23 기준 1.5억명이 대화에 참여했고 100억개 이상의 메시지가 발송되었다. 정체되었던 스냅챗 다운로드 트래픽도 증가하고 있으며 인앱결제 매출도 견조하게 증가하고 있다. 메타 역시도 자체 챗봇인 페르소나 출시를 앞두고 있다.

그림 112. 스냅챗 글로벌 월간 다운로드



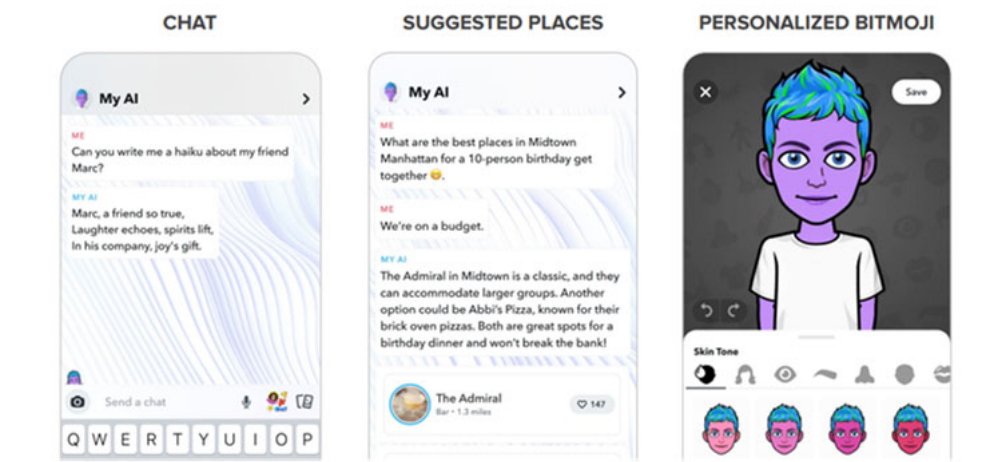
자료: 센서타워, 미래에셋증권 리서치센터

그림 113. 스냅챗 글로벌 인앱결제 월매출



자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 114. 스냅 생성 AI 챗봇 마이AI



자료: 스냅, 미래에셋증권 리서치센터

생성 AI 도입을 통해 SNS 플랫폼 업체들에게 신규 수수료 BM도 생겨날 것으로 전망한다. 자체 앱 이외에도 타사의 다양한 AI 앱들을 온보딩시켜 결제 매출 또는 광고 매출에 대한 수수료 수익이 발생하게 될 것이다. 챗봇들 내에서 커머스 연결도 시작된다면 커머스 매출 발생에 대한 일정 부분의 수수료 수취도 가능해질 것이다.

대형 SNS 플랫폼은 수많은 생성 AI 앱들의 스토어 기능을 겸하게 될 것이다. 이미 범람하기 시작한 생성형 AI 어플리케이션들이 개인에게 주는 효용 차이는 점차 적어질 가능성이 높고 개별 독립 앱으로서의 존재보다는 이미 거대 MAU를 갖춘 플랫폼에 종속될 가능성이 높다. 플랫폼 입장에서는 신규 수익도 수익이지만 플랫폼 트래픽이 증가하는 일석이조의 효과를 누린다.

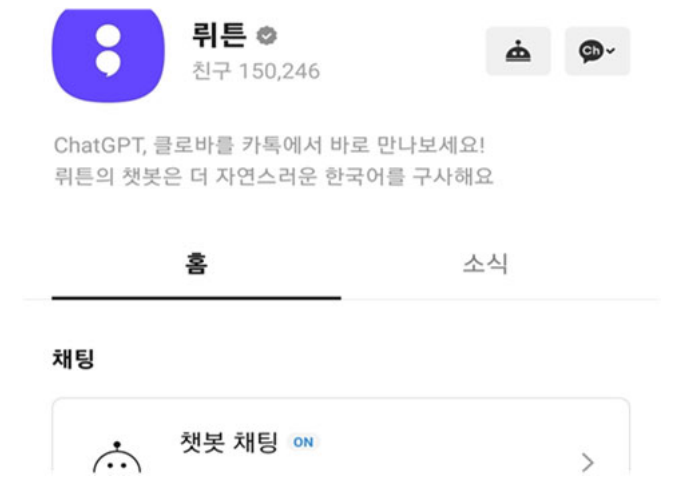
국내는 카카오톡이 대표적이다. 국내에서 인기를 끌고 있는 주요 생성형 어플리케이션들은 이미 카카오톡 내에 입점하는 방식으로 카카오톡 내 버티컬 서비스 강화에 도움을 주고 있다. 생성형 AI 기반 챗봇 서비스로 국내 시장에서 돌풍을 일으키고 있는 챗봇 서비스 Askup과 뽀빠이 카카오톡 채널을 통한 서비스 확장을 택한 점이 대표적이다. 추후 이러한 어플리케이션 모델의 주요 BM이 광고가 될 수도 있고 구독형 요금제가 될 수도 있겠지만 어떤 방식이 되더라도 플랫폼에 수수료를 나누게 될 것이다.

그림 115. 카카오톡에 채널에 입점한 AskUp



자료: 카카오톡, 미래에셋증권 리서치센터

그림 116. 카카오톡에 채널에 입점한 뽀빠이



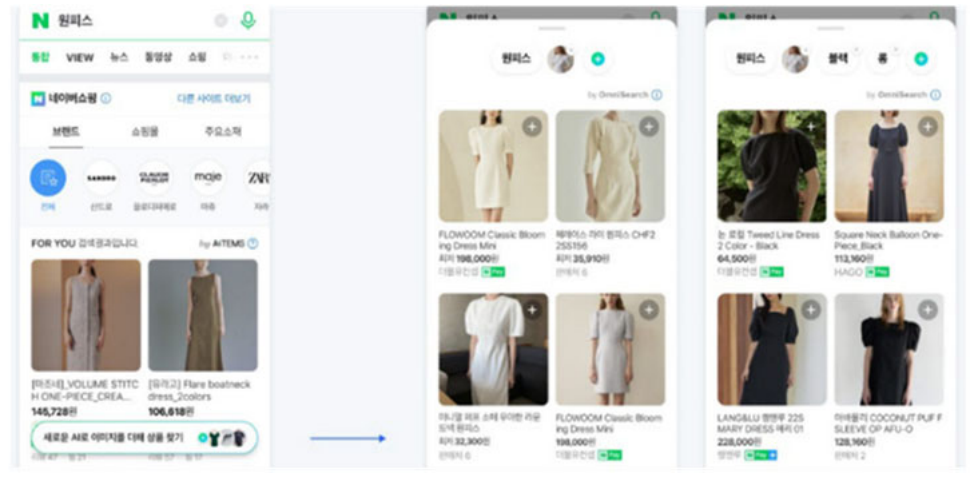
자료: 카카오톡, 미래에셋증권 리서치센터

2. 커머스: 구매 전환을 상승, 라이브커머스 활성화

AI 기반 멀티모달 서비스 도입

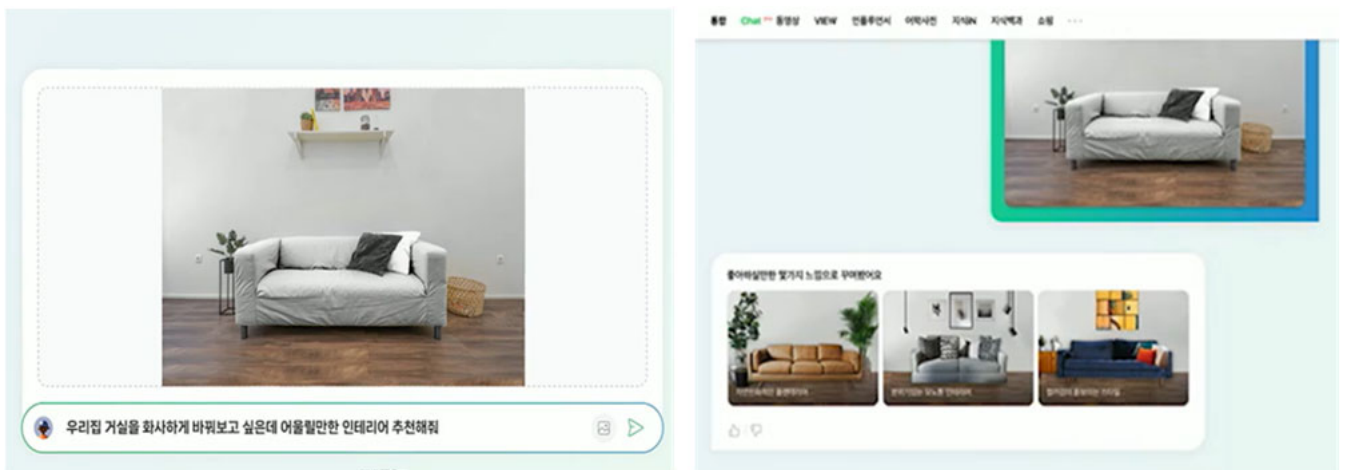
AI 기반으로 이미지, 음성, 텍스트를 동시에 인식하는 멀티모달 서비스가 도입되면서 이용자가 원하는 상품을 보다 효율적으로 찾을 수 있게 변할 것이다. 멀티모달 서비스 도입으로 쇼핑 이용자가 고민할 만한 상품을 보기까지 걸리는 시간이 대폭 단축될 것이다. 그 결과 쇼핑 피로도도 낮아지고 구매 전환율이 상승하게 된다. 또한 검색 이용자는 배경, 공간 등에 어울리는 상품 추천까지 받을 수 있을 것으로 보인다. 더 나아가면 공간에 대한 코디와 더불어 개인의 사진에 대한 AI 커스터마이징이 가능해질 것으로 기대된다.

그림 117. NAVER 멀티모달 AI 서비스 옴니서치



자료: NAVER, 미래에셋증권 리서치센터

그림 118. NAVER 이미지+텍스트 결합형 검색 예시: "우리집 거실을 화사하게 바꿔보고 싶은데 어울릴만한 인테리어 추천해줘"



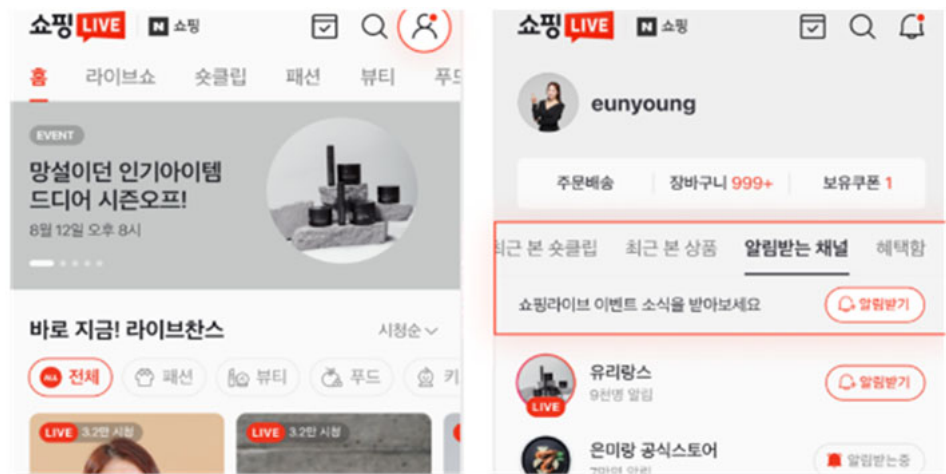
자료: NAVER, 미래에셋증권 리서치센터

라이브커머스 성장 가속화

생성 AI의 도입은 라이브커머스 시장의 성장을 가속화할 전망이다. 생성 AI의 도입으로 1) 특정 품목에 관심이 있을 잠재 고객을 타겟팅 할 수 있게 되고, 2) 대본 생성을 통한 판매자 생산성이 향상되며, 3) AI 휴먼 라이브 방송이 가능해지고, 4) 다국어 라이브 스트리밍이 가능해질 전망이다.

AI 기반 유저 타겟팅이 고도화되면 될수록 TV홈쇼핑 시장의 붕괴는 가속화되고 라이브 커머스 침투율이 올라갈 것이다. TV홈쇼핑이 쇠퇴 원인은 TV 시청자의 급감도 있지만 홈쇼핑 채널들이 고객들에 대한 능동적인 타겟팅이 불가능하다는 점에 있기 때문이다. AI 기반 메시지 마케팅은 특정 상품을 사길 원하는 구매 의지가 높은 고객을 해당 방송 시간에 맞춰 데려올 수 있다. 그 결과 판매자의 판매 효율과, 구매자의 시간 및 구매 효율성 모두가 향상된다.

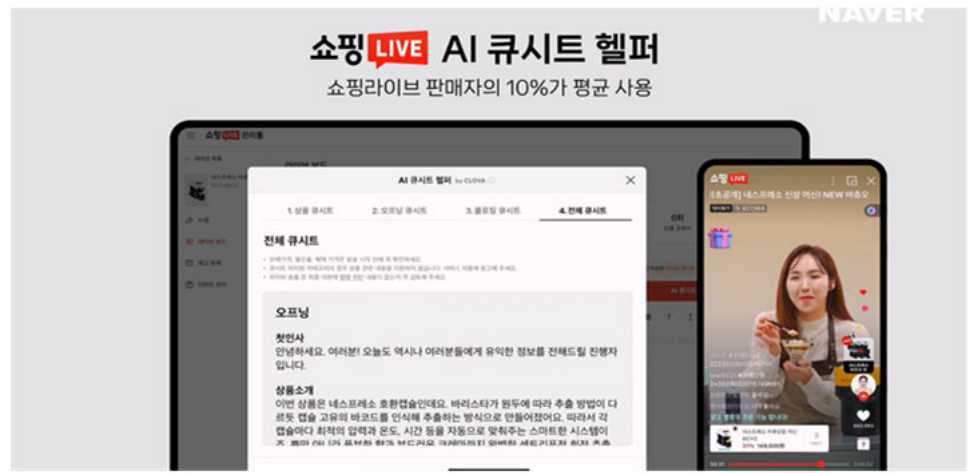
그림 119. NAVER 쇼핑라이브 알림 기능



자료: 네이버쇼핑, 미래에셋증권 리서치센터

생성 AI의 도입은 라이브커머스를 위한 대본 준비의 부담도 완화시켜 판매자 측면에서의 확장도 가속화되고 있다. 기존 홈쇼핑에서는 대본 작성에 대한 부담, 실시간 의견 취합 등의 이슈로 브랜드 규모가 있는 곳이 주로 이용했지만 생성 AI의 도입을 통해 라이브커머스의 기회가 중소상공인에게도 열리고 있다. NAVER의 AI 큐시트 헬퍼의 경우 상품 특징, 고객 리뷰 등을 바탕으로 전문 쇼호스트나 인플루언서의 화법이 반영된 대본을 작성해준다. 지난 6월부터 도입되어 시간이 많이 지나지 않았음에도 AI 큐시트 헬퍼 침투율은 쇼핑라이브 판매자의 10%까지 올라왔다.

그림 120. 네이버 쇼핑라이브 AI 큐시트 헬퍼



자료: NAVER, 미래에셋증권 리서치센터

생성 AI 도입으로 AI 휴먼 보급도 보편화되면 직접 출연에 대한 부담 또는 쇼호스트 고용 비용 부담이 완화돼 라이브커머스 판매자 증가가 가속화될 것이다. AI/메타버스 기업 이모션웨이브는 9월 싱가포르 라이브커머스 기업 팬텀리소스와 MOU를 체결하기도 했다. 라이브커머스가 활성화된 중국에서도 무인 전자상거래를 위한 AI 휴먼에 대한 개발에 뛰어들어 따라 AI 휴먼 호스트 시대의 도래도 머지않은 것으로 판단한다.

그림 121. 생성 AI 기반 디지털 휴먼 방송 솔루션 AIMPLE Studio



자료: 이모션웨이브, 미래에셋증권 리서치센터

생성 AI를 통해 급격히 올라오고 있는 실시간 통번역 능력은 다국어 라이브 스트리밍 서비스가 가능하게 만들 것이다. 이전 스트리밍 서비스는 국가간 경계가 확실했지만 빠른 속도로 허물어질 전망이다. 글로벌 라이브커머스 솔루션 업체 라라스테이션은 실시간 번역 서비스를 적용해 다국어 라이브 스트리밍 서비스를 보여주고 있기도 하다.

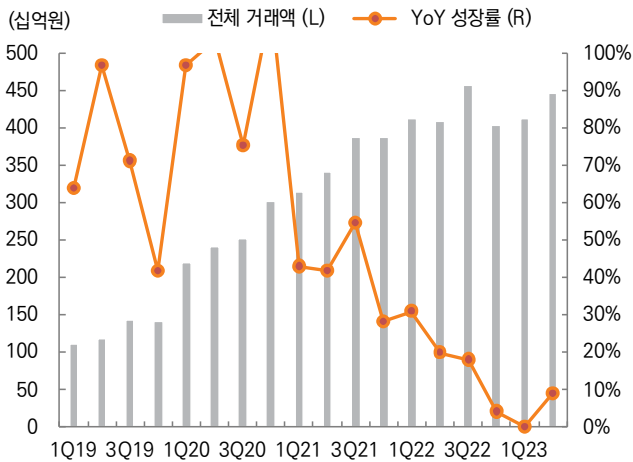
성장 둔화에 직면한 웹툰·웹소설 산업

3. 콘텐츠: 생성 AI 도입을 통한 공급난 해소 기대

생성형 AI의 도입은 현재 웹툰·웹소설 산업이 직면하고 있는 콘텐츠 공급난을 완화시켜 줄 것이다. 글로벌 웹툰·웹소설 산업은 22년 이후 성장세가 급격하게 둔화되고 있다. 글로벌 최대 웹툰 플랫폼인 NAVER와 카카오조차도 2H22부터 콘텐츠 거래액의 YoY 역성장이 관찰되고 있는 상황이다.

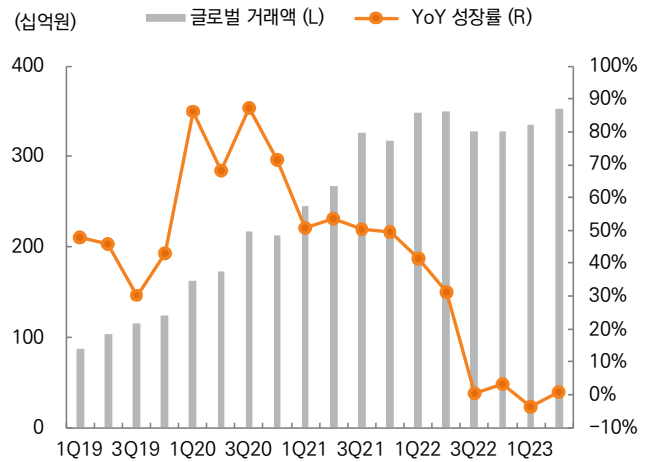
성장세의 급격한 둔화는 콘텐츠 공급 부족에 기인한다. 19~21년 웹툰 작가가 각광받으면서 신규 웹툰 작가들이 대거 유입되고 콘텐츠 공급량이 급격히 확대되었으나 현재는 신규 작가 유입 증가세가 일단락된 상태다. 현재 상황에서 콘텐츠 공급 확대를 위해 기대할 수 있는 부분은 생산성 증가 부분이다.

그림 122. 네이버웹툰 글로벌 분기 거래액



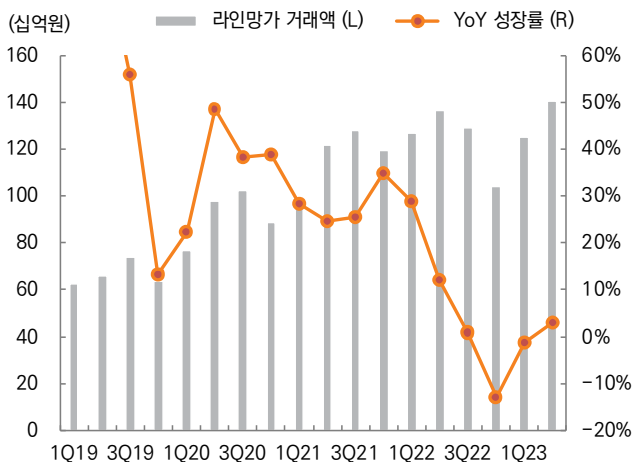
자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 123. 카카오 웹툰 글로벌 분기 거래액



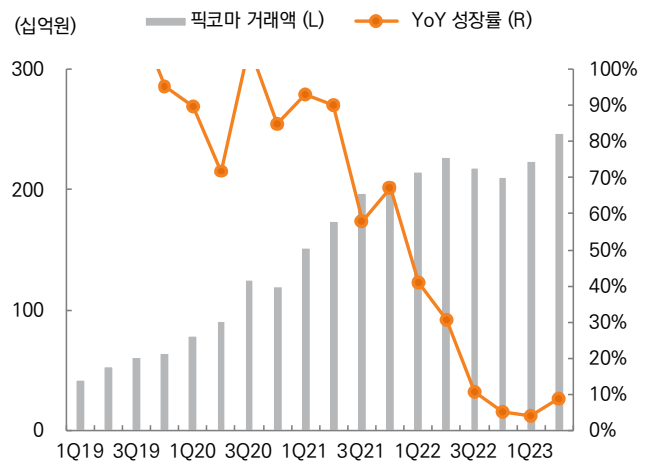
자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 124. 라인망가 분기 거래액



자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 125. 픽코마 분기 거래액

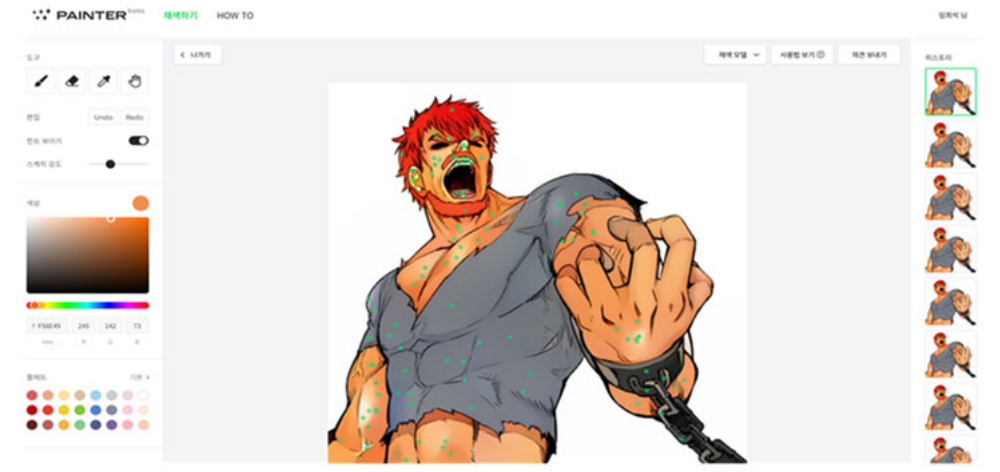


자료: 센서타워, 미래에셋증권 리서치센터 추정

생성 AI 도입이 불러올 생산성 향상

생성형 AI와 AI툴의 도입이 양질의 웹툰/웹소설 콘텐츠 증가로 이어질 전망이다. 이미 웹툰/웹소설 시나리오를 생성하는 플랫폼들이 등장하고 있으며 클릭 한 번으로 시가 채색을 완성시켜주는 툴도 고도화되고 있다. 웹툰은 작가가 하루 평균 그림을 그리는 시간이 10시간 반(콘진원 2022년 웹툰 작가 실태 조사)에 달할 정도로 채색 노동 부담이 높다. AI 툴의 적용으로 채색 부담이 크게 완화될 것으로 기대된다.

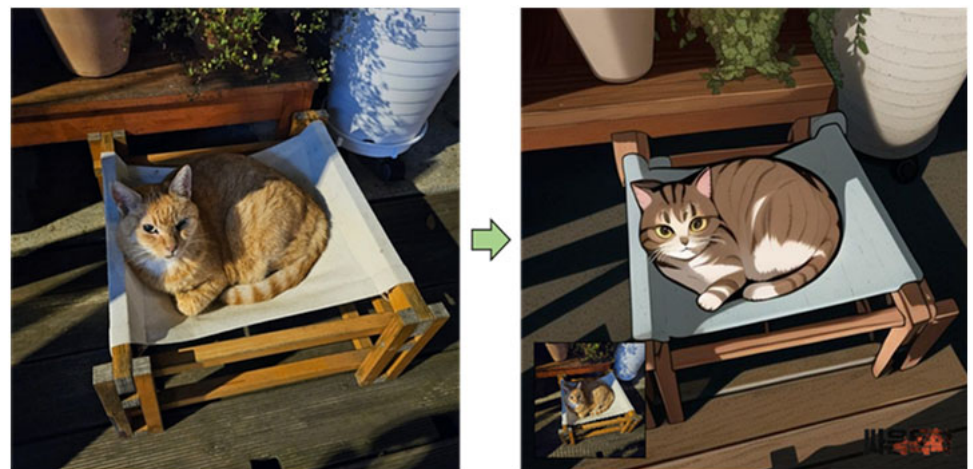
그림 126. AI 페인터: 클릭 몇 번만으로 손쉽게 채색이 가능하다.



자료: 본인 화면, AI 페인터, 미래에셋증권 리서치센터

글로벌 최대 웹툰 플랫폼 네이버웹툰은 AI 기술을 웹툰에 발빠르게 도입하고 있다. AI 페인터를 통해서 자동으로 채색할 수 있도록 도와주며 튠필터의 도입을 통해 사진이나 영상을 웹툰으로 변환해주기도 한다. NAVER의 AI 채색 툴 AI 페인터는 그림 쪽으로는 문외한인 필자조차도 10분 정도면 한 장을 완성할 정도까지 고도화되었다.

그림 127. 튠필터 이용 예시: 집주변 길냥이를 싸움독학 웹툰 그림체로 변환해 보았다.



자료: 본인 화면, 튠필터, 미래에셋증권 리서치센터

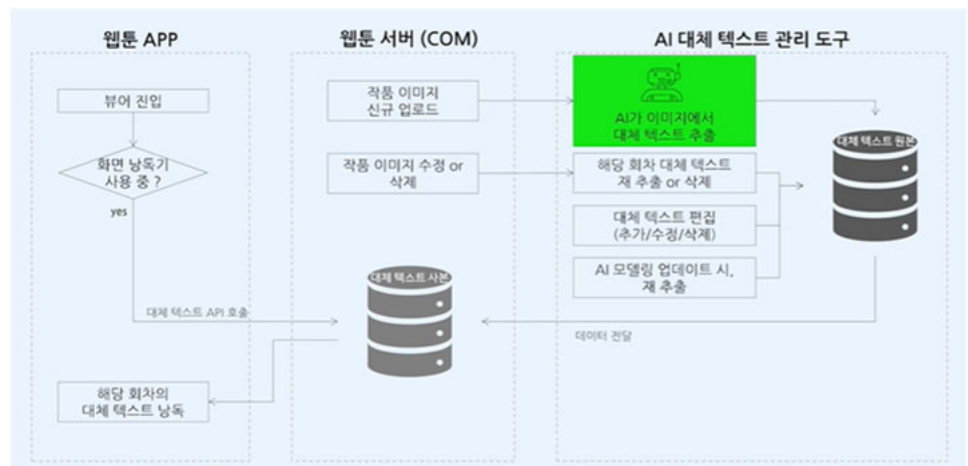
채색뿐만 아니라 이미지의 대체 텍스트화 측면에서도 AI는 혁신을 일으키고 있다. 모든 웹툰 이미지에 대해 사람이 일일이 텍스트 정보를 입력하는 데는 한계가 존재했으나 AI는 무리 없이 가능하기 때문이다. AI 기술을 통해 웹툰 이미지를 텍스트로 변환하고 이를 다시 청각 정보로 변환해 시각 약자도 즐길 수 있는 배리어프리 웹툰을 제공할 수 있을 것으로 기대된다. 시각 약자도 웹툰 콘텐츠를 즐길 수 있게 되면서 이용자 측면에서의 확장이 추가적으로 일어날 전망이다.

그림 128. AI 기술 적용을 통한 이미지의 청각 정보 변환



자료: 네이버웹툰, 미래에셋증권 리서치센터

그림 129. 네이버웹툰 AI 대체 텍스트 서비스 구조도



자료: 네이버웹툰, 미래에셋증권 리서치센터

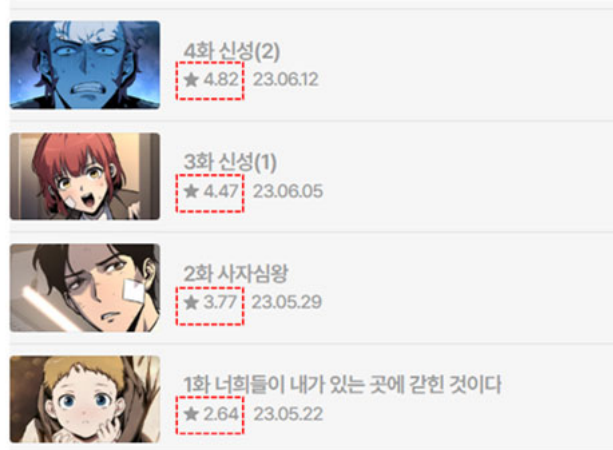
AI 툴 보편화에는 시간이 필요할
전망

역사적으로 콘텐츠 생산의 진입장벽 하락은 늘 거대 시장을 몰고 왔다. 게임 제작 진입 장벽을 낮춘 앱스토어의 등장은 모바일 게임 시대를 열며 게임 산업의 퀀텀 점프를 이끌었다. 영상 콘텐츠 제작 진입장벽을 낮춘 유튜브의 등장은 수십조 단위의 스트리밍 광고 시장을 만들었다. 스토리를 만드는 생성형 AI와 이를 웹툰화 하는 AI 툴의 등장은 새로운 거대 시장 개화의 도화선이 될 수 있다. 미래에는 생성형 AI 구상 시나리오에 AI 채색이 엮어진 수준급 콘텐츠들에 이용자들이 지갑을 여는 시기가 도래할 것이다.

그림 130. 네이버웹툰 도전만화 작가 보이콧 사태(6월) 그림 131. AI 논란으로 별점테러 받은 '신과함께 돌아온 기사왕님'



자료: 도전만화, 미래에셋증권 리서치센터



자료: 네이버웹툰, 미래에셋증권 리서치센터

다만 생성 AI 툴의 전면적인 도입과 이를 통한 콘텐츠 공급의 확장에는 시간이 좀 더 필요할 전망이다. 1) 생성 AI의 도입에 대한 작가 및 독자의 거부감, 2) 이익 분배 구조에 대한 기준 미확립, 3) 생성 AI의 무분별 학습에 대한 검증 능력 부재 이슈가 존재하기 때문이다. 해당 이슈들의 해결에는 1년 이상이 소요될 것으로 판단하며 신규 콘텐츠 공급 확대에 따른 웹툰 산업 고성장세 회복은 25년 이후에 이어질 것으로 판단한다.

네이버웹툰의 경우 지난 5월 생성 AI 작화 적용 논란에 휩싸인 '신과함께 돌아온 기사왕님'에 대한 독자들의 별점테러가 이어지기도 했으며, 6월에는 AI 웹툰에 대한 비판으로 작가들의 네이버웹툰 도전만화 보이콧 사태가 일어나기도 했다. 네이버웹툰은 독자들의 반발에 따른 조치로 작가 지망생들의 등용문으로 여겨지는 지상최대공모전에서 생성형 AI를 금지하기도 했다. 네이버웹툰은 9월 작가 개인용 AI를 개발할 예정이라고 발표했다. 특정 작가의 이미지를 학습시키고 그 작가만 사용하게 하는 방식이다. 22년 말에서 23년 초부터 작가들과 협업이 시작될 예정이다.

4. 게임: 생산성 증대와 비용 감소에 따른 공급 확대 전망

생성 AI 게임 내 적용 본격화

NPC 적용

생성 AI의 게임 내 NPC 적용이 본격화되고 있다. 6월 중국에서 출시된 역수한(넷이즈 개발)은 생성 AI를 NPC에 전격적으로 도입했다. 생성 AI가 도입된 지능형 NPC의 경우 고유의 성격, 행동 논리, 기억력을 갖고 있으며 플레이어의 행동과 환경 변화에 능동적으로 대응한다. 캐릭터 커스터마이징에도 생성 AI가 도입되었는데 플레이어가 원하는 외형을 문장으로 입력하면 묘사에 부합하는 캐릭터의 얼굴이 생성된다. 이 외에도 플레이어는 게임을 탐험하면서 생성 AI를 활용해 시를 작성할 수도 있다.

생성 AI의 도입으로 게임 몰입감을 상승시킨 역수한의 돌풍은 매출로도 이미 증명되고 있다. 6월 출시 후 2~3위 수준의 매출 최상위권을 유지하고 있으며 iOS 매출로만 일 40억원 이상이 발생하고 있다.

그림 132. 생성 AI를 통해 텍스트 기반 캐릭터 커스터마이징도 가능한 역수한



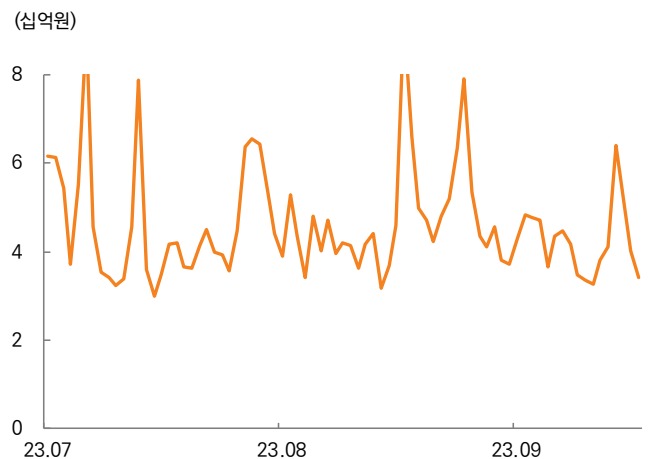
자료: 역수한, 미래에셋증권 리서치센터

그림 133. 왕자영요에 이어 iOS 매출 2위를 유지 중인 역수한



자료: 센서타워, 미래에셋증권 리서치센터

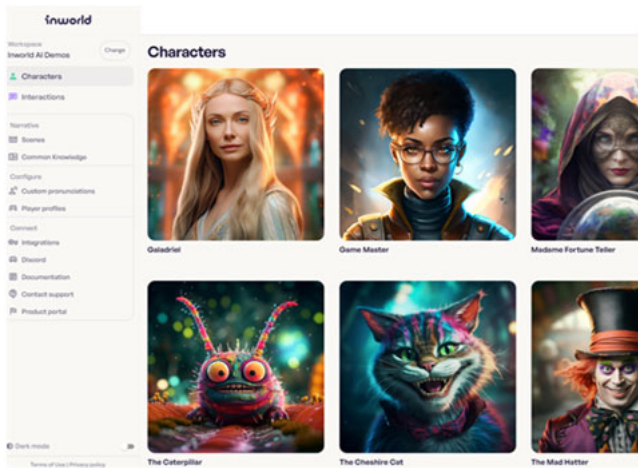
그림 134. 역수한 iOS 매출 추이: 40억원 전후 일매출 유지



자료: 미래에셋증권 리서치센터

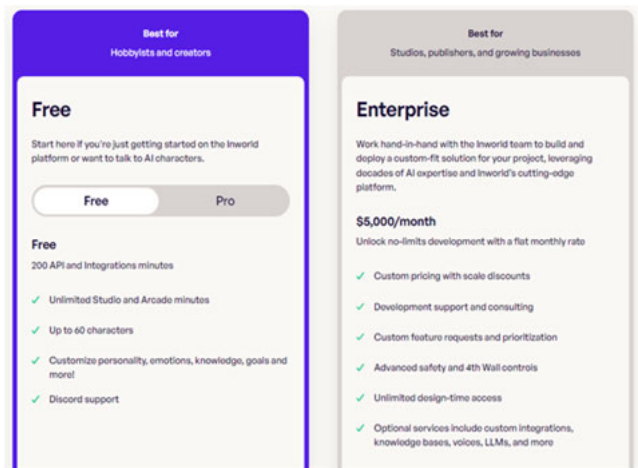
NPC에 대한 생성 AI 적용은 중국뿐만 아니라 글로벌에서 빠른 속도로 보편화될 전망이다. 딥마인드 출신 인력이 설립한 스타트업 인월드 AI는 게임 속 NPC를 AI로 생성하는 툴을 제공한다. 인월드 AI의 NPC는 자율적으로 목표를 정하고 행동을 수행한다. 인월드 AI 플랫폼 전반을 무제한으로 이용할 수 있는 월 5천달러 요금제가 존재한다. 인월드 AI는 지난 8월 MS, 삼성, LG 등으로부터 5천만 달러 규모의 투자를 유치하기도 했다.

그림 135. 인월드 AI 캐릭터



자료: 인월드 AI, 미래에셋증권 리서치센터

그림 136. 인월드 AI 요금제



자료: 인월드 AI, 미래에셋증권 리서치센터

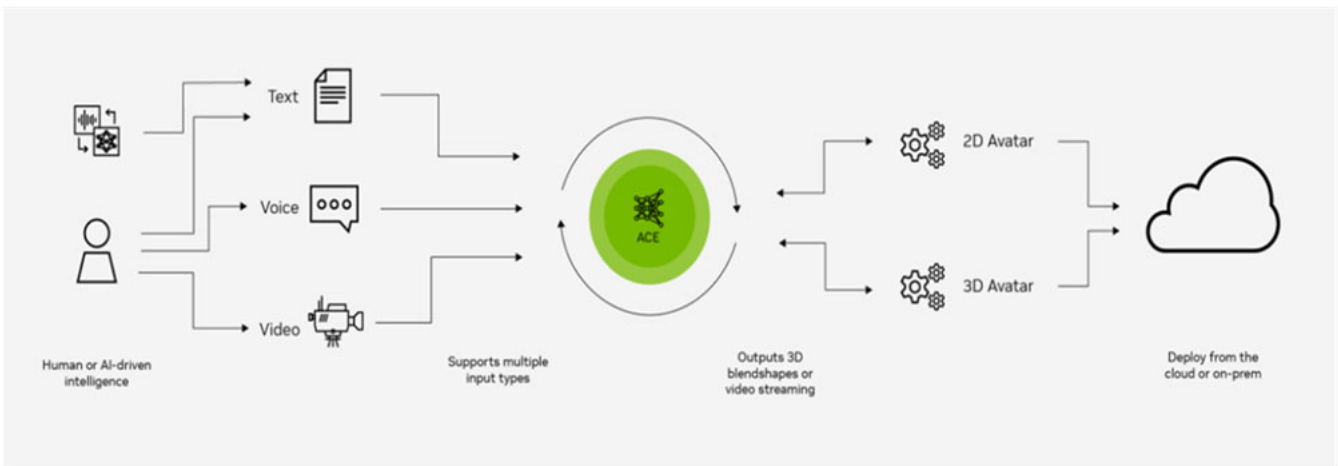
엔비디아가 지난 5월 공개한 ACE(Avatar Cloud Engine)는 AI NPC가 게임에 보편적으로 도입되는 시기가 머지않았음을 보여준다. 이번 데모에서 공개한 사이버펑크2077 세계관을 기반으로 만든 미래형 라면가게 주인 JIN은 유저와 상호작용이 가능한 모습을 보여주었다. AI NPC가 플레이어와의 자연스러운 대화를 통해 게임 스토리 전개를 돕거나 퀘스트를 자율적으로 주는 등의 행동을 통해 플레이어의 몰입감을 향상시킬 전망이다.

그림 137. 엔비디아가 공개한 사이버펑크2077 세계관 기반 AI NPC 'JIN'



자료: 엔비디아, 미래에셋증권 리서치센터

그림 138. 게임 캐릭터 개발의 End-to-End까지 제공하는 ACE 클라우드 API

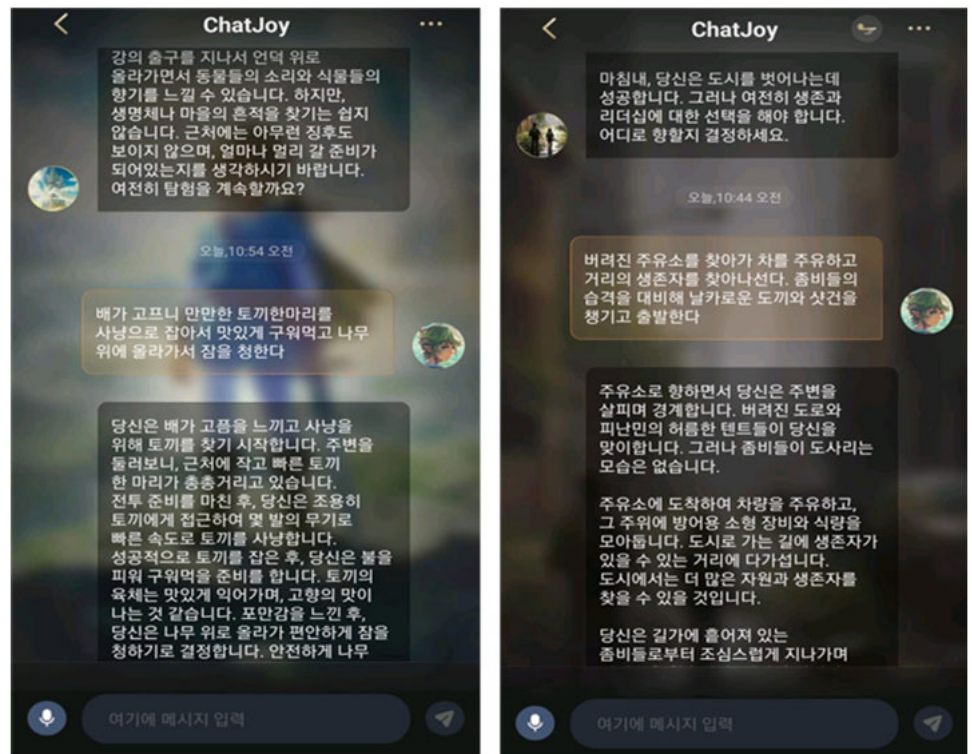


자료: 엔비디아, 미래에셋증권 리서치센터

자유로운 세계관

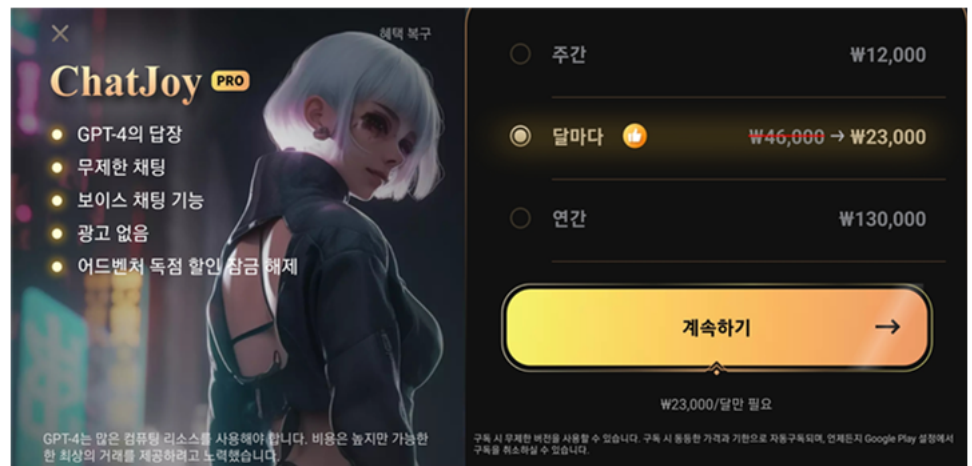
정해진 세계관 내에서 NPC에 부분적으로 생성 AI를 적용하는 것에서 나아가 세계관 자체를 플레이어가 함께 만들고 모험을 진행하는 게임도 등장하고 있다. 챗조이라는 챗 GPT4를 활용한 AI 채팅 텍스트 기반 RPG가 대표적이다. 게임 내 여러가지 세계관이 존재하며 세계관을 고른 후에는 텍스트 기반으로 대화하며 스토리를 진행해 나가는 구조다. 기존에 없었던 자율성을 갖는 것이 특징이다. 구독 요금제를 주요 BM으로 하고 있으며 월간 23,000원, 연간 130,000원의 요금제를 갖고 있다.

그림 139. 챗조이 플레이 예시



자료: 본인 화면, 미래에셋증권 리서치센터

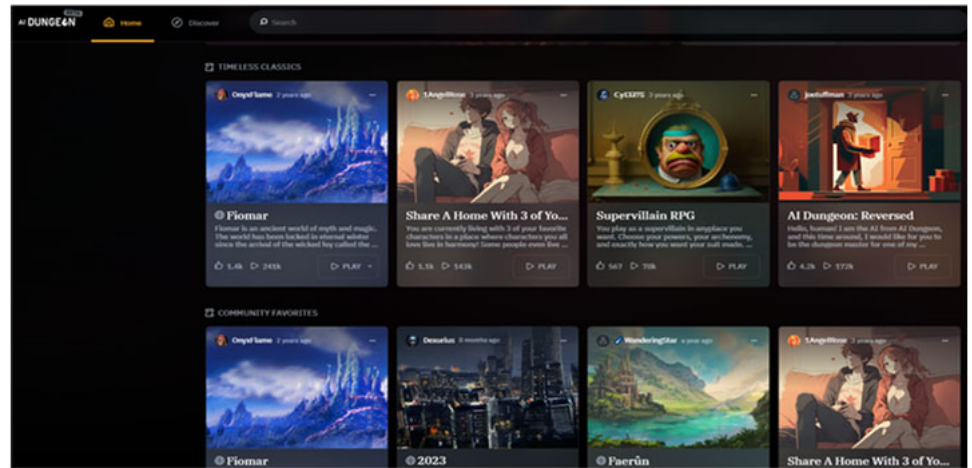
그림 140. 챗조이 요금제



자료: 본인 화면, 미래에셋증권 리서치센터

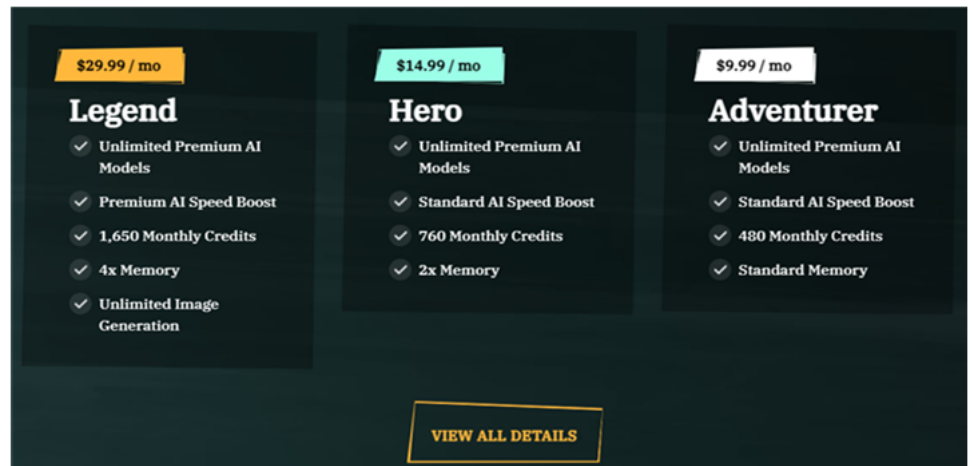
챗GPT를 활용한 AI던전(2019년 출시)도 대표적인 사례다. AI던전도 유저가 텍스트를 입력하고 시가 다음에 올 스토리를 창작하면서 이야기가 진행된다. 출시 초반만 하더라도 스토리 전개에 미흡한 부분이 많아 혹평 일색이었으나 유저들의 사용 데이터를 학습하면서 AI 수준이 급격하게 올라와 인기를 얻게 되었다. 출시 이후 앱 다운로드 수는 300만 건을 넘어섰으며, 월 1~2억 수준의 인앱 결제 매출이 발생 중인 것으로 추정된다.

그림 141. 유저가 함께 세계관을 만들어나가는 AI던전



자료: AI던전, 미래에셋증권 리서치센터

그림 142. AI던전 요금제

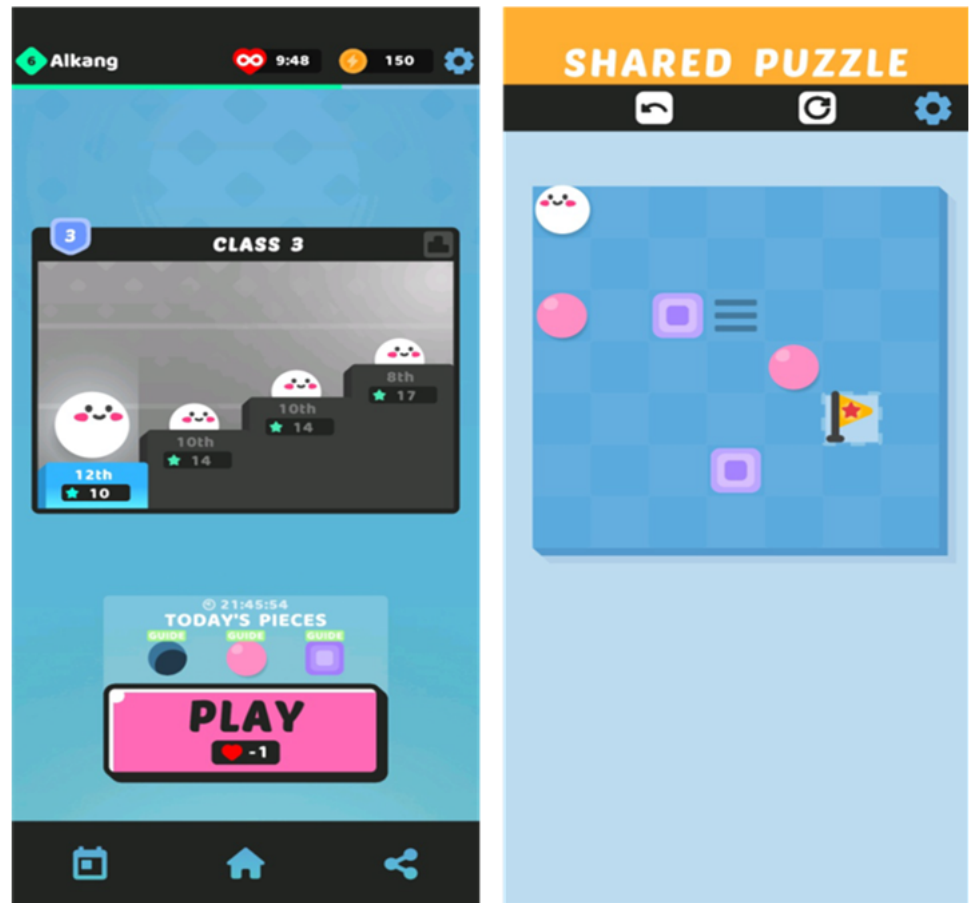


자료: AI던전, 미래에셋증권 리서치센터

스테이지형 게임의 진화

생성 AI의 도입은 스테이지형 게임의 진화를 이끌 것으로 기대된다. 퍼즐 게임에는 이미 변화가 일어나고 있다. 크래프톤의 독립 스튜디오 렐루게임즈가 출시한 '푼다: AI퍼즐'은 AI가 스테이지를 생성해서 플레이어에게 제공한다. 개별 플레이어의 수준에 맞게 스테이지가 생성되기 때문에 플레이어들은 초개인화된 퍼즐 경험을 할 수 있게 된다. 특정 스테이지에 대한 공유도 가능해 친구들과 함께 즐길 수 있는 장점도 존재한다.

그림 143. 푼다: AI 퍼즐

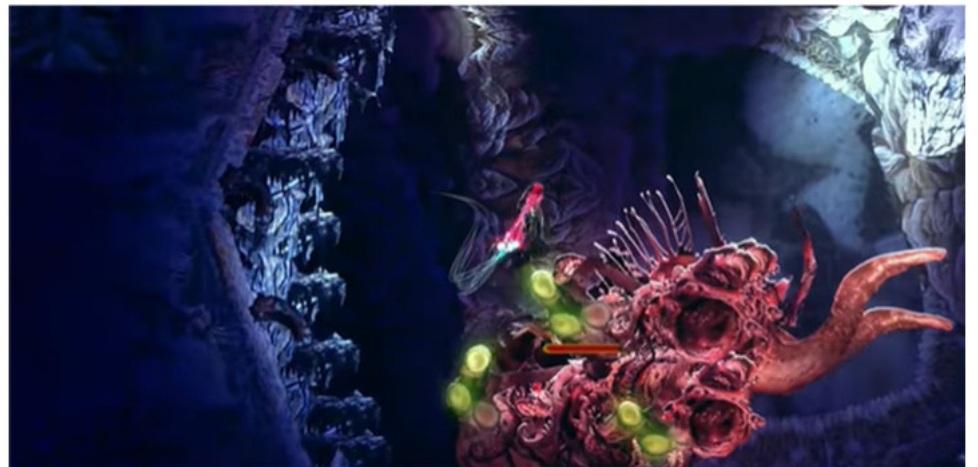


자료: 본인 화면, 미래에셋증권 리서치센터

생성AI의 도입은 로그라이크 게임의 진화로 이어질 가능성도 존재한다. 로그라이크 게임은 스테이지 또는 클리어 과정에서의 선택지들이 랜덤하게 생성되며 죽었을 때 모든 것이 초기화된다는 특징을 지니고 있다. 다양한 성공 루트의 존재와 높은 난이도로 높은 인기를 이어오고 있지만 랜덤 요소들이 과도한 스트레스로 작용해 유저 이탈로 이어지는 부분도 크다.

생성 AI의 도입을 통해 특정 시점의 선택지에 대한 클리어 기뻐감 추론이 고도화될 것이다. 생성 AI 도입으로 '너무 비합리적이다'라고 유저들이 느낄 만한 선택지들의 연속적인 제시를 막아 스트레스성 이탈을 완화시킬 수 있을 것으로 기대된다.

그림 144. 소스 오브 매드니스: AI가 플레이어 수준에 맞춰 몬스터를 지속 생성

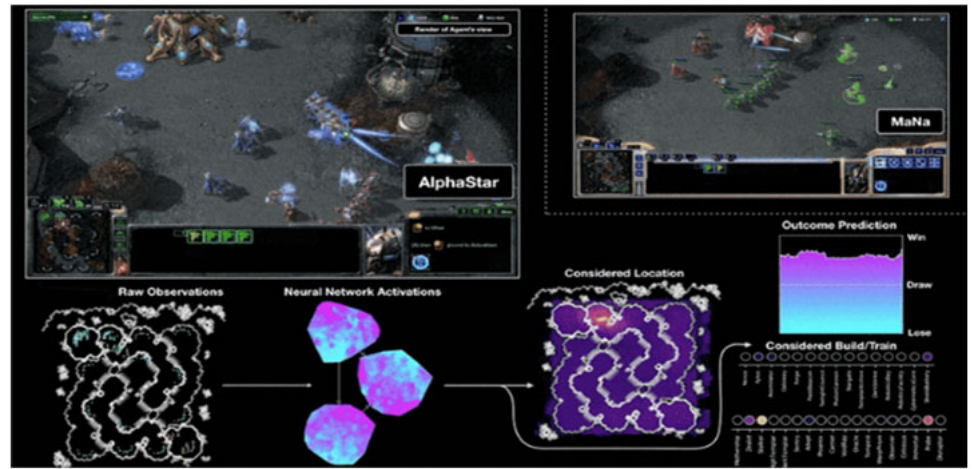


자료: 스팀, 미래셋증권 리서치센터

멀티플레이어 게임에서의 가능성

생성 AI의 고도화로 멀티플레이어 게임의 재미가 극대화될 것으로 판단한다. 스타크래프트, 오버워치, 배틀그라운드 등 팀을 이루어 플레이하는 멀티플레이어 게임은 팀원에 대한 이슈가 발생하는 경우가 많다. 본인이 제 몫을 다하며 준수한 플레이를 펼치더라도 팀원이 트롤링을 시작하면서 속수무책으로 게임을 지게 되는 경우가 심심찮게 발생한다. 또한 특정 시간대에 비슷한 실력대의 사람이 부족한 경우 대기 시간이 과도하게 길어지는 문제도 존재한다. 이러한 문제들로 인해 멀티플레이어 게임에 대한 이탈로 이어지는 경우가 많다.

그림 145. 딥마인드의 AI '알파스타': 19년부터 스타2 프로게이머 상대로 이기기 시작



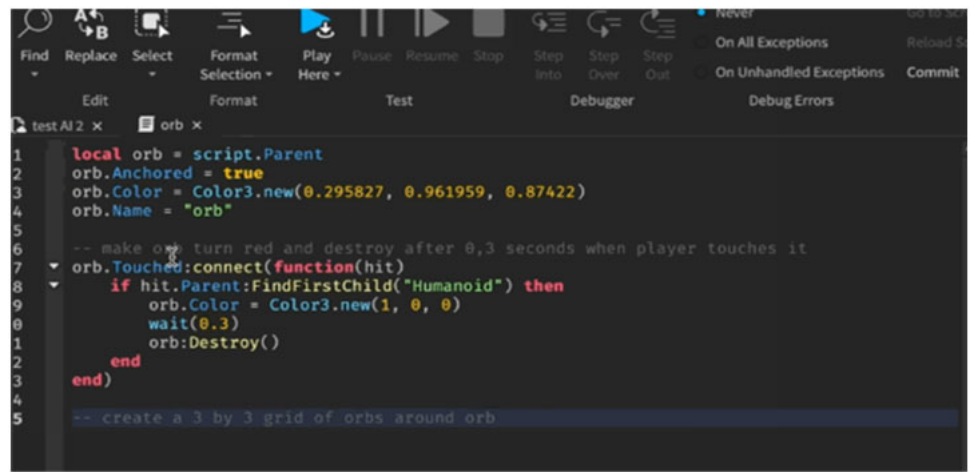
자료: 딥마인드, 미래에셋증권 리서치센터

AI 플레이어 도입의 보편화는 팀원간 과도한 실력 편차, 고의적 트롤링, 과도한 대기 시간 문제를 완화시켜줄 수 있다. 스타크래프트처럼 20년 전에도 AI 플레이어 설정은 여러 게임에 존재했으나 수준 미달의 플레이로 인해 유명무실한 경우가 대다수였다. 현재는 AI의 전략 게임 플레이 수준이 이미 일반인을 크게 상회하고 있는 상황이다. 시야, APM 등 피지컬 수준을 사람 정도로 제한해서도 비슷한 플레이를 보일 수 있게 된다면 멀티플레이어 게임 내 AI 팀원의 활용이 크게 증가할 것으로 보인다.

생산성 증대에 따른 콘텐츠 공급 증가

생성 AI의 적용은 게임 생산성을 극대화해 코로나 이후 둔화된 게임 업종 내 콘텐츠 공급을 가속화할 전망이다. 게임은 기술집약적 산업임과 동시에 스토리, 그래픽, 사운드, 프로그래밍 등 작업 전반을 사람이 수작업해야 하는 노동집약적 산업이었다. 앞으로는 생성 AI가 스토리를 생성하고, 그래픽과 사운드를 입히고, 프로그래밍을 보조하는 시대가 될 것이다. 개발자들이 리소스를 수급하는데 사용한 시간과 비용의 상당 부분이 콘텐츠 개발로 전환될 것이다.

그림 146. 로블록스 코드어시스트 예시



자료: 로블록스, 미래에셋증권 리서치센터

노코드 개발의 보편화와 개발자 저변 확대

생성 AI의 도입으로 노코드 개발 툴이 보편화되고 개발자 저변이 확대될 전망이다. 로블록스는 지난 3월 코드어시스트라는 생성형 AI 기반 게임 제작 툴을 공개했다. 코드어시스트는 원하는 것을 입력하면 코드를 자동으로 생성해준다. 전문적인 코딩 지식이 없는 사람도 텍스트 입력만을 통해서 아이템 제작이 가능하다.

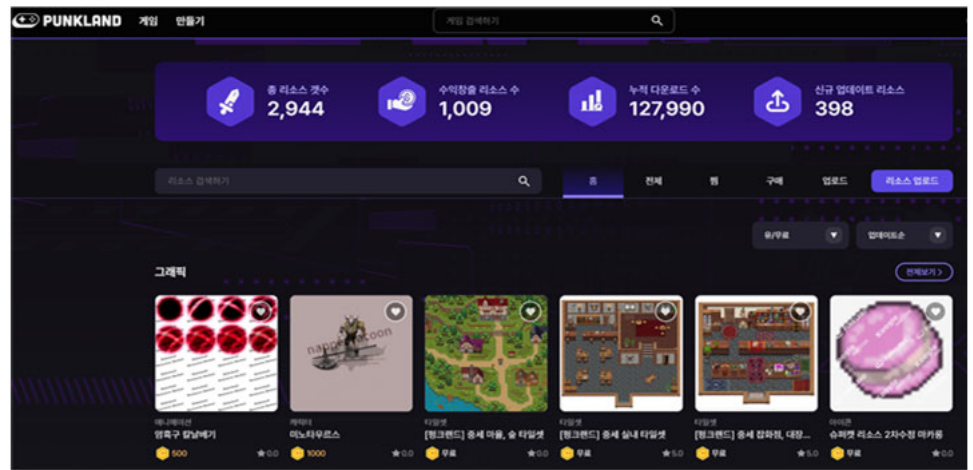
코딩에 대한 진입장벽 완화로 개발자 연령층도 확대될 전망이다. 노코드 개발 툴을 도입하고 있는 로블록스는 플레이어였던 이용자가 게임을 즐기고 성장하면서 자연스럽게 개발자가 되는 경우가 많다. 그러나 이용자의 43%가 13세 미만이지만 게임 제작 이용자는 주로 18세 이상이다. 생성 AI로 게임 개발 진입장벽이 낮아지면서 게임 개발자 연령층이 보다 넓어질 것으로 판단한다.

1인, 소규모 단위 게임 개발 증가
예상

생성 AI 기반의 리소스 보급은 개발자들이 리소스를 수급하는데 기존에 사용한 시간과 비용을 콘텐츠 개발에 사용할 수 있도록 만든다. 이는 아트 리소스 수급에 큰 부담을 가진 1인 또는 소규모 개발자들의 개발 편의성을 대폭 향상시켜 콘텐츠 공급 증가로 이어질 전망이다.

슈퍼켓의 노코드 게임 개발 플랫폼 핑크랜드는 AI 이미지 생성기를 무료로 도입해 아트 리소스 수급을 즉각적으로 가능하게 만들었다. 1인 게임 개발자 '33게임즈'가 선보인 '닥사RPG: 파밍용사키우기'는 서비스 50일만에 매출 5억원을 달성하기도 했다. 게임 개발 난이도의 하락으로 신규 게임 개발자가 증가하고 플랫폼 내 리소스가 확대되어 게임 퀄리티 증가로 이어지는 선순환 구조가 지속될 것으로 기대된다.

그림 147. 슈퍼켓의 노코드 게임 개발 플랫폼 핑크랜드



자료: 핑크랜드, 미래에셋증권 리서치센터

아트 비용의 감소

생성 AI의 도입으로 게임 개발 생산성이 좋아짐과 동시에 제작 비용이 드라마틱하게 감소할 것으로 전망한다. 트리플A급 또는 대작 게임을 만드는 데 있어서 들어가는 비용은 적게는 수백억에서 많게는 수천억원에 달한다. 그리고 이러한 게임 제작 비용에서 개발 비용과 함께 가장 많은 비중을 차지하는 영역이 아트 비용이다. 생성 AI 툴의 도입으로 2~3년 내 아트 비용이 급감하기 시작할 것으로 판단한다.

생성 AI 기반 이미지 활용이 본격화되면서 기존 일러스트 관련 비용이 급감할 것이다. 일러스트레이터가 수작업으로 그린 게임 일러스트 가격은 장당 수십만원을 호가하지만 생성 AI가 그린 일러스트 가격은 장당 수십원 수준에 불과하다. 물론 생성 AI 툴로 그린 그림에 대한 리터칭 필요성이 있으므로 비용은 일부 추가될 것이나 기존의 순수 수작업과 비교 시 10분의 1의 비용도 들지 않게 될 것이다.

그림 148. 미드저니 요금제

	Basic Plan	Standard Plan	Pro Plan	Mega Plan
Monthly Subscription Cost	\$10	\$30	\$60	\$120
Annual Subscription Cost	\$96 (\$8 / month)	\$288 (\$24 / month)	\$576 (\$48 / month)	\$1152 (\$96 / month)
Fast GPU Time	3.3 hr/month	15 hr/month	30 hr/month	60 hr/month
Relax GPU Time	-	Unlimited	Unlimited	Unlimited
Purchase Extra GPU Time	\$4/hr	\$4/hr	\$4/hr	\$4/hr
Work Solo In Your Direct Messages	✓	✓	✓	✓
Stealth Mode	-	-	✓	✓

자료: 미드저니, 미래에셋증권 리서치센터

21년 5월 오픈AI의 달리(DALL-E), 22년 7월 미드저니의 미드저니(Midjourney), 22년 8월 스테빌리티AI의 스테이블디퓨전(Stable Diffusion)이 출시되면서 이미지 생성 AI의 성능이 급격하게 올라오기 시작했다. 22년 9월에는 미드저니가 생성한 그림이 미국 주립 박람회 미술대회에서 우승하며 이미 AI 생성 이미지 수준이 사람 수준 이상으로 올라왔음을 증명하기도 했다.

그림 149. 스테이블디퓨전 예시1



자료: 스테이블디퓨전, 미래에셋증권 리서치센터

그림 150. 스테이블디퓨전 예시2



자료: 스테이블디퓨전, 미래에셋증권 리서치센터

웨이보 등 중국 SNS 플랫폼에서는 게임 업체들이 스테이블디퓨전, 미드저니를 도입하기 시작하면서 해고된 일러스트레이터들이 사례가 넘쳐나고 있다. 생성 AI 도입 이전에는 캐릭터 외주 비용이 8,000위안 수준이었지만 현재는 2,000위안 수준까지 감소한 것으로 파악된다. 현재는 게임회사에서 AI 시안을 먼저 제공하고 구체화 및 리터칭을 외주가 담당하는 일이 증가하고 있다. 작업에 1~2주 정도 걸린 정교한 일러스트도 생성 AI 도입 이후 3~4일이면 충분한 수준까지 올라왔다.

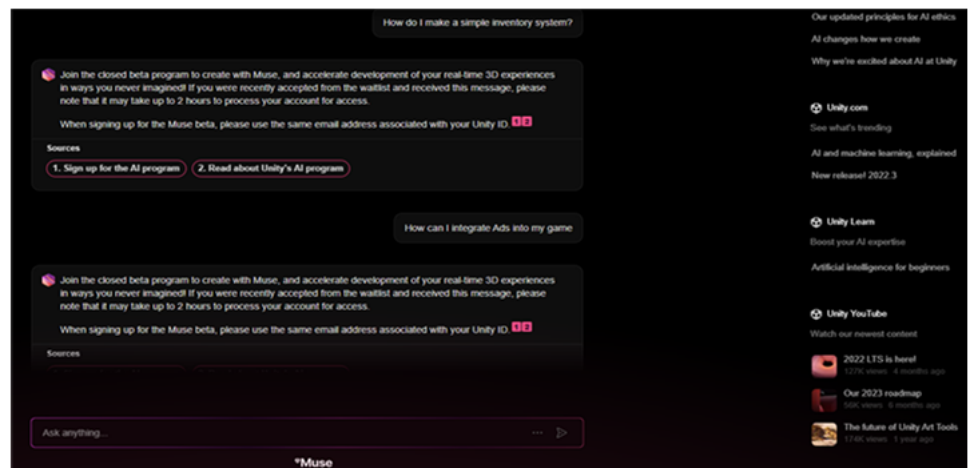
중국 게임사 XD네트워크(心动网络)의 회장 황이명에 따르면 게임 개발팀들에서 기존에 외주를 주던 일러스트와 번역 업무를 더이상 주지 않는 경우가 늘어나고 있다. 중국 대형 게임사들은 이미 초기 디자인과 세부 특성화 작업에 AI를 도입하고 있다. 넷이즈가 대표적인 경우인데 넷이즈는 현재 게임 제작 전반에 생성 AI를 적극적으로 활용하고 있다. 중소 게임사들도 더이상 외주 일러스트를 활용하지 않고 직접 AI툴을 이용하는 양상이 많아지고 있다.

게임 엔진사 수혜를 예상

생성 AI 도입에 따른 게임 산업 내 콘텐츠 공급 확대로 게임 엔진사가 수혜를 볼 것이다. 생성 AI 도입으로 1) 프로그래밍 생산성 향상, 2) 1인, 소규모 개발 활성화, 3) 노코드 툴 도입에 따른 개발자 저변 확대, 4) 아트 리소스 재분배 효과가 이어져 콘텐츠 공급이 확대될 것이다. 1H23부터 게임사들의 생성 AI 테스트가 시작되고 있는 점을 고려할 때 2H24 또는 25년부터는 생성 AI 도입 게임들의 출시가 본격화될 전망이다.

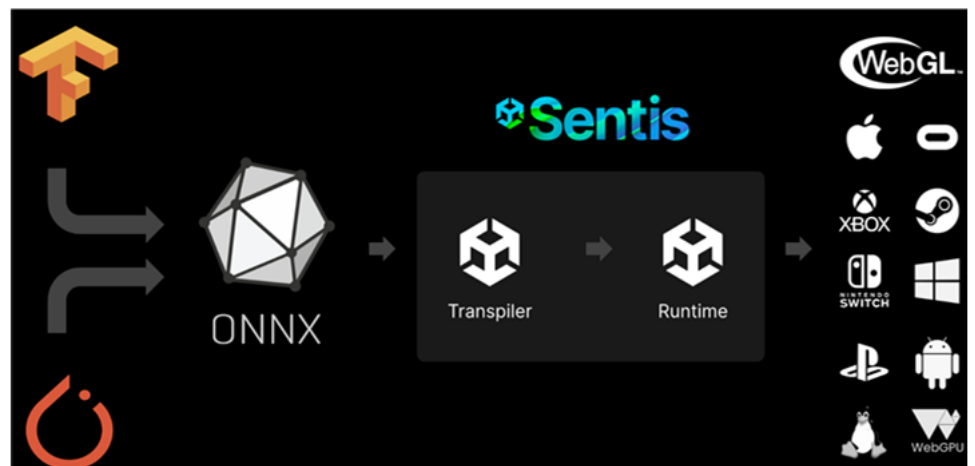
생성 AI 도입을 통한 게임 산업 회복에 따른 수혜는 시장을 10년 넘게 양분하고 있는 유니티와 언리얼엔진에 집중될 가능성이 높다. 게임 엔진은 각 플랫폼별로 언어가 다르기에 락인효과가 큰 편이며 생성 AI를 엔진사들 대부분이 도입한다고 해도 그 수준의 차이는 그간 쌓인 에셋의 양과 질에서 올 것이기 때문이다. 양사 모두 에셋의 양과 질 측면에서 타사들을 압도하고 있다.

그림 151. 유니티 신규 AI 툴 뮤즈



자료: 유니티, 미래에셋증권 리서치센터

그림 152. 유니티 신규 AI 툴 센티스



자료: 유니티, 미래에셋증권 리서치센터

VI. 금융

1. 금융 AI 시장은 26년 562억달러까지 성장할 전망

21년 세계 금융분야 AI 시장 규모는 113억달러다. 26년까지 연평균 38% 성장하여 562억달러의 시장을 형성할 전망이다. 전체 AI 시장 내 금융분야 비중이 21년 19.4%에서 26년 18.2%로 소폭 하락할 것이나, 여전히 가장 큰 비중을 차지할 것이다.

국내도 마찬가지다. 금융분야가 가장 큰 비중인 18.8%이며, 21년 시장 규모는 5.2억달러를 기록했다. 26년까지 연평균 38% 성장하여 28억달러의 시장을 형성할 것이다.

전체 AI 시장 내 금융권 비중이 높은 이유는 그만큼 시가 금융권 내 강한 임팩트를 줄 수 있는 기술이기 때문이다.

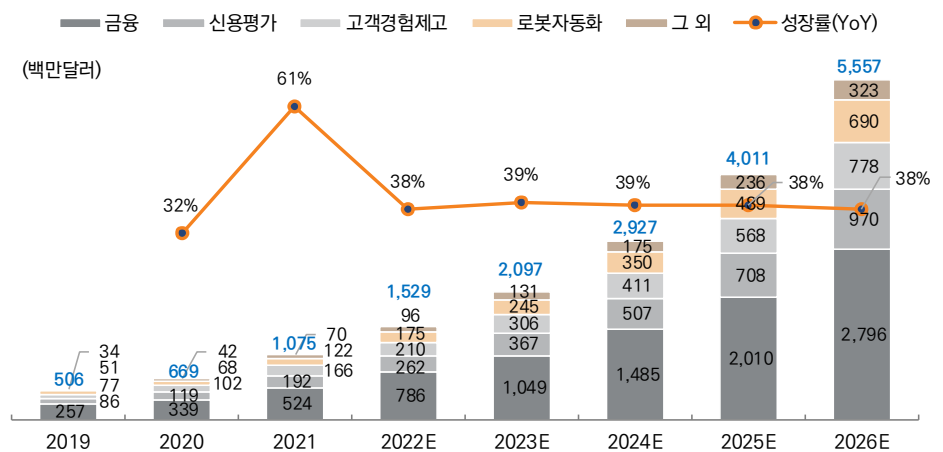
표 20. 세계 AI 시장 규모 전망

(억달러)

구분	2019	2020	2021	2022F	2023F	2024F	2025F	2026F	CAGR (21~26)
시장규모	289	377	581	869	1,252	1,742	2,337	3,095	39.7%
금융	57	74	113	167	238	327	433	562	37.8%
IT	48	62	96	142	203	281	375	494	38.8%
소매 및 전자상거래	42	55	86	130	190	268	364	488	41.5%
헬스케어	40	53	82	125	182	257	350	475	42.1%
운수 및 물류	24	32	50	76	110	155	210	282	41.3%
국방 및 공공	30	39	59	87	124	170	225	294	37.9%
제조	22	29	45	68	100	140	190	254	41.4%
에너지	19	25	38	57	82	115	154	204	39.9%
그 외	7	8	12	17	23	29	36	42	28.5%
성장률(YoY)	-	30.4%	54.1%	49.6%	44.1%	39.1%	34.2%	32.4%	-

자료: MarketandMarkets(2021), 미래에셋증권 리서치센터

그림 153. 국내 금융분야 AI 시장규모



자료: MarketandMarkets(2021), 미래에셋증권 리서치센터

2. 금융권에서의 생성형 AI의 역할




(국제금융센터의 ‘글로벌 은행권의 ChatGPT 활용 가능성 및 과제(23.2.28 발간)’를 인용)

금융권은 다양한 업무에서 생성형 AI를 통해 1) 생산성 및 효율성을 향상시키고, 2) 고객경험을 개선하고, 3) 보안 및 리스크 관리 강화를 모색할 수 있다.

1) 생성형 AI를 통한 업무 자동화는 생산성 제고는 물론, 작업 부하를 줄이고 작업 완료 속도를 높여 조직의 효율성을 끌어올린다. 가령 對고객서비스(질문 답변, 제품/서비스 정보 제공 등), 데이터 입력 및 일상적 마케팅 등의 반복적 작업을 자동화할 수 있다. 이를 통해 해당업무의 담당자는 조금 더 부가가치가 높은 작업에 집중할 수 있다. 게다가 법률계약서 작성 및 분석, 재무보고 등의 고부가가치 업무를 수행함에 있어서도 시간을 절약해주고 속도와 정확성을 높여줄 것으로 기대된다. 결국 이는 인건비, 교육비 절감에 따른 금융회사의 비용 효율성 제고로도 직결된다.

2) 고객경험을 개선시킬 수 있다. 생성형 AI를 활용해 연중무휴 신속한 고객 지원이 가능해진다. 또한 고객 데이터 분석을 통한 맞춤형 금융서비스 제공이 용이 해진다. 가령 고객이 신규 신용카드를 신청 시, AI 가상비서가 지출습관, 재정상태 등을 분석해 가장 적합한 상품을 추천할 수 있다. 이는 고객경험을 향상시켜 충성도 강화로 이어진다.

3) 보안 및 리스크 관리를 강화할 수 있다. 금융거래의 이상 패턴을 파악하도록 AI 알고리즘을 훈련시킬 수 있으며, 이를 통해 사기거래를 감지하고 예방할 수 있다. 대형언어모델(LLM)과 신경망 데이터 등의 기술은 고객 행동패턴 및 거래 데이터 등에 대한 분석을 지원해, 은행의 비정상 활동들을 탐지하는 데 기여할 수 있다. 또한 방대한 데이터에 대한 실시간 분석을 통해 시장동향 및 잠재 리스크 식별과 관리에 도움이 된다.

그림 154. 금융권의 생성형 AI 활용 가능성		
생산성 및 효율성 향상	고객경험 개선	보안 및 리스크 관리 강화
		
<ul style="list-style-type: none"> - 對고객 서비스 및 반복 작업 자동화 → 고부가가치 업무 집중 - 법률계약서 및 재무보고 등의 작성 속도와 정확도 향상 - 인건비 및 교육비 절감 등 비용 효율성 제고 	<ul style="list-style-type: none"> - 고객 대기시간 단축 및 문의 신속 정확 대응 - 고객 데이터/선호도/주 사용시간 등 분석 → 맞춤형 금융서비스 제공 & 효과적 타겟 마케팅 전략 수립 - 인터넷 및 모바일 앱 내 유익한 콘텐츠 생성 	<ul style="list-style-type: none"> - 이상 금융거래 등 비정상 활동 사전 감지 및 예방 - 시장 동향 & 잠재적 리스크 평가예측 속도와 정확성 향상

자료: 국제금융센터, KCIF, 미래에셋증권 리서치센터

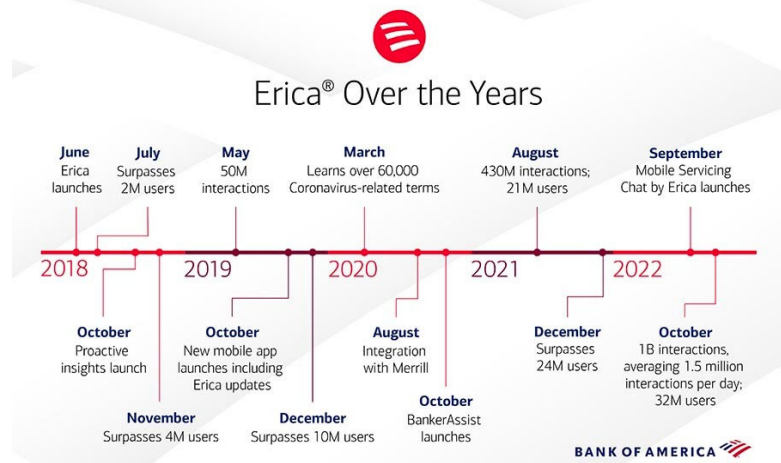
3. 해외 금융업계의 생성형 AI 도입 사례

뱅크오브아메리카(BofA)의 Erica 서비스

Erica는 BofA가 2018년에 출시한 AI 금융 비서다. 고객들의 재정 및 업무 관리를 도와주며, 계좌 정보 안내, 소비/대출/자산 관리, 결제 및 송금, 거래 내역 조회 등을 포함한 다양한 서비스를 제공한다. Erica는 전세계 금융권 내 생성형 AI를 banking에 적용한 가장 성공적인 사례로 꼽힌다.

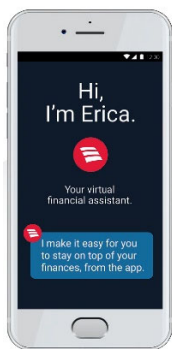
Erica는 고급 머신 러닝 및 자연어 처리 기술을 사용하여 고객 문의를 이해하고 관련 답변을 제공한다. 또한 고객의 피드백과 행동을 통해 학습하고, BofA 내 다양한 데이터 소스와 시스템에 접근해 상황에 맞는 개인화된 정보와 솔루션을 제공할 수 있다.

그림 155. Erica 연혁



자료: BofA, 미래에셋증권 리서치센터

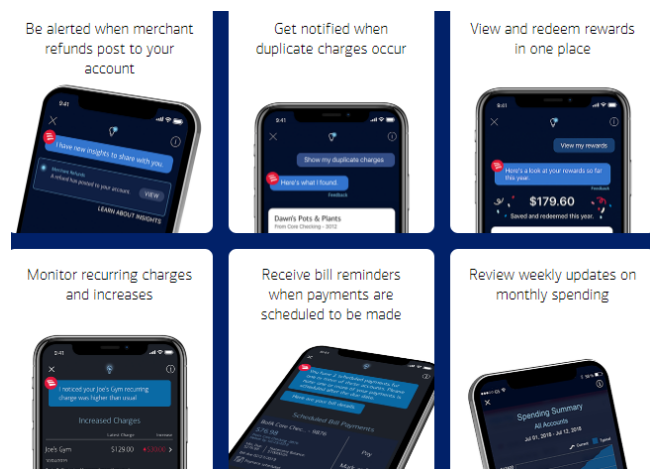
그림 156. Erica 주요 서비스



- ✓ 계좌 정보 안내(계좌번호, 잔액 등)
- ✓ 소비/대출/자산 관리
- ✓ 청구 예정 사항 공지
- ✓ 카드 분실 시 잠금
- ✓ 거래 내역 조회
- ✓ 결제 및 송금
- ✓ 신용 점수 조회
- ✓ 정기 결제 안내
- ✓ 리워드 혜택 정보
- ✓ 환불 입금 알림

자료: BofA, 미래에셋증권 리서치센터

그림 157. Erica 사용 사례



자료: BofA, 미래에셋증권 리서치센터

Erica는 BofA 모바일 앱이나 모바일 웹사이트 등을 통해 접속 가능하다. 고객은 음성, 문자 또는 제스처를 통해 Erica와 상호작용한다. "Hey Erica"라고 말하거나 Erica 아이콘을 누르고 대화를 시작해 [그림158]처럼 banking 관련 질문이나 요청을 할 수 있다. 이에 Erica는 적절한 답변이나 조치로 응답하거나 필요시 추가 정보를 요청한다. 또한 [그림159]처럼 고객의 재무 상황과 목표에 따라 사전 안내 및 제안을 제공하기도 한다.

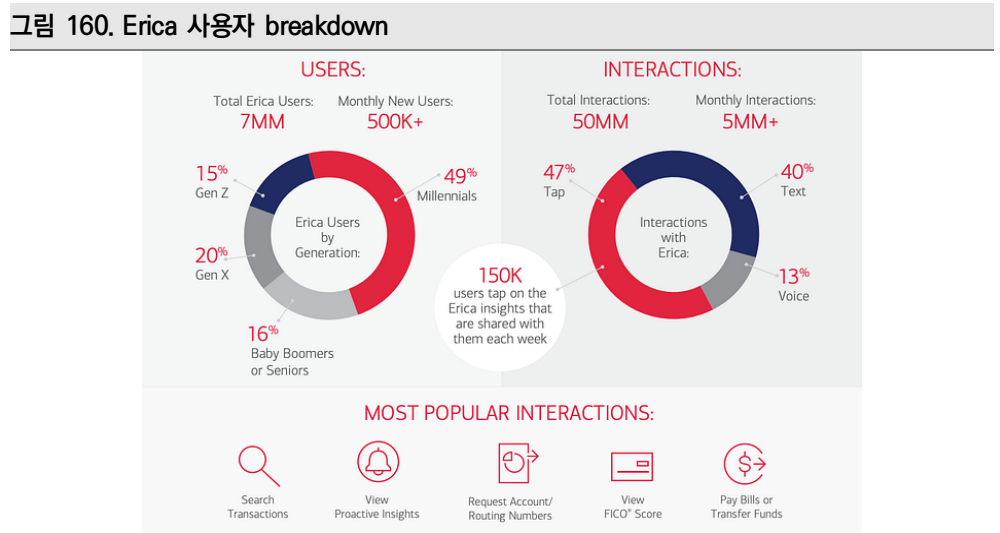
그림 158. banking 관련 질문 및 요청 예시(고객 → Erica)		그림 159. 사전 안내 및 제안 예시(Erica → 고객)	
업무 구분	질문	업무 구분	사전 안내 및 제안
계좌 정보 안내	"내 계좌에 얼마가 남아있어?"	계좌 정보 안내	"잔액 부족 알람이 있습니다. 저축 계좌에서 돈을 이체해 드릴까요?"
송금	"내 예금 계좌에서 신용 카드로 100달러를 송금해줘"	정기 결제 안내	"곧 신용카드 결제 기한이 도래합니다. 지금 결제하시겠습니까, 아니면 미리 알림을 설정하시겠습니까?"
결제	"내일까지 전기세를 내줘."	소비 관리	"이번 달 외식비가 평소보다 많이 지출되었습니다. 예산을 세우는 데 도움을 드릴까요?"
자산 관리	"어떻게 하면 더 많은 돈을 절약할 수 있을지 보여줘"	리워드 혜택 정보	"기프트 카드로 교환할 수 있는 리워드 포인트가 충분히 적립되었습니다. 몇 가지 옵션을 보여드릴까요?"
신용점수 조회	"내 신용 점수는 몇 점이야?"	대출 관리	"주택담보대출 이자율이 높습니다. 재융자를 도와드릴까요?"

자료: BofA, 미래에셋증권 리서치센터

자료: BofA, 미래에셋증권 리서치센터

Erica 도입 효과

Erica는 BofA의 사업 전반에 걸쳐 직간접적으로 긍정적 영향을 끼치고 있다. 무엇보다 고객 만족도 향상에 큰 기여를 했다. 편리하고 쉬운 방법으로 고객의 banking 니즈를 지원해 BofA에 대한 만족도와 신뢰도를 높였다. BofA에 따르면, Erica를 사용하는 고객은 그렇지 않은 고객보다 전반적인 banking 경험을 더 높게 평가했다고 한다. 디지털 전환에 취약한 베이비부머들도 Erica 사용자의 16%나 차지하는 점은 주목할 만하다[그림 160].



자료: BofA, 미래에셋증권 리서치센터

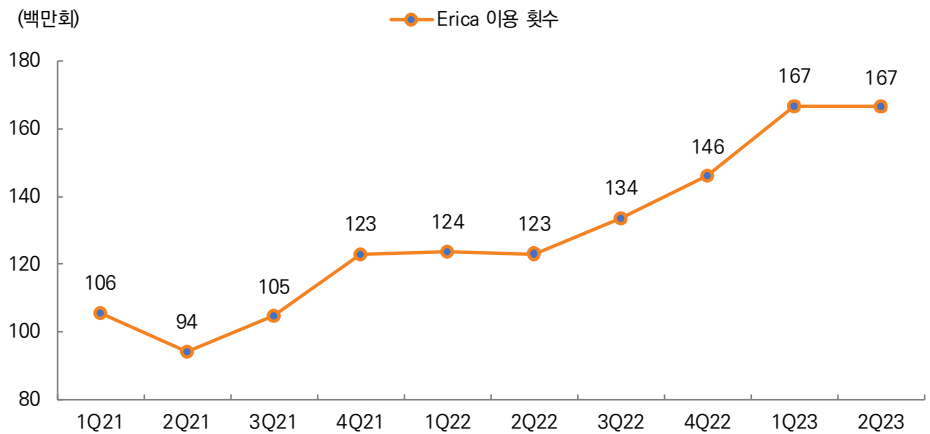
고객 참여도를 증가시켰다. Erica는 고객이 BofA의 상품/서비스와 더 자주, 더 깊이 상호 작용하도록 장려한다. BofA에 따르면, Erica를 사용하는 고객은 그렇지 않은 고객보다 모바일 앱에 두 배 더 자주 로그인한다고 한다.

Erica의 분기별 이용 횟수는 견조하게 우상향 중인데, 2Q23 기준 1.67억회를 기록했다[그림9]. 출시 4년만인 22년 10월에 누적 이용 횟수 10억회를 돌파했고, 9개월 이후인 23년 7월에는 15억회를 돌파했다.

비용 절감 효과가 있다. Erica는 일상적인 작업과 문의를 자동화하여 운영 비용을 절감시켰다. BofA에 따르면 Erica는 출시이후 100만 시간 이상의 직원 시간을 절약했다.

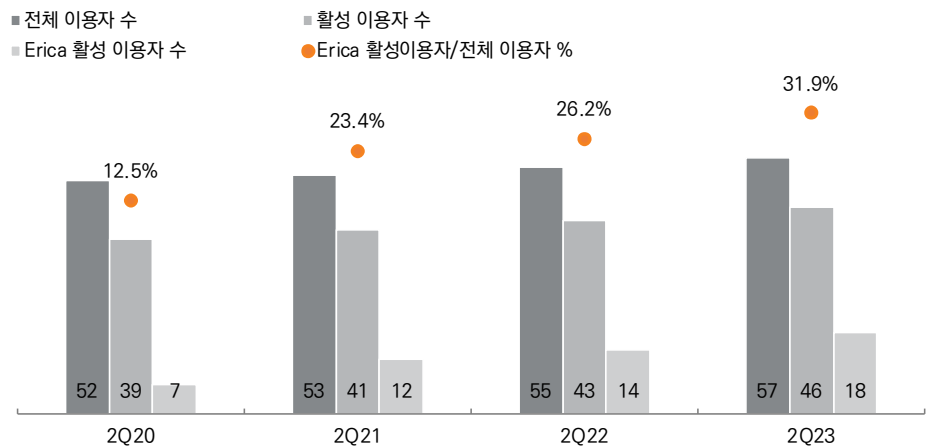
효율성도 향상시킨다. Erica는 고객이나 직원에게 복잡하거나 번거로울 수 있는 프로세스를 간소화한다. 예를 들어, 고객이 대출을 신청하거나 계좌를 개설하는데 몇 시간 또는 며칠이 아닌, 몇 분 만에 처리할 수 있도록 지원한다.

그림 161. Erica 이용 횟수 추이



주: Consumer digital 기준
 자료: BofA, 미래에셋증권 리서치센터

그림 162. Erica 이용자 증가 추이(백만명)



주: 활성이용자=90이내 활동 기록 존재하는 이용자
 자료: BofA, 미래에셋증권 리서치센터

캐나다왕립은행(RBC)의 NOMI 서비스

NOMI는 RBC가 2017년 출시한 AI 금융 비서이다. 론칭 초기에는 Insights와 Find & Save 서비스만 있었지만, Budgets(2019년), Ask NOMI(2020년), Forecast(2021년)를 추가로 제공하면서 더욱 고도화된 서비스를 선보였다. RBC는 NOMI를 기반으로 자국 내 초개인화 banking 영역에서 가장 선두적인 입지를 구축했다.

Erica와 NOMI, 모두 자사의 데이터소스와 시스템에 접근해 개인별 맞춤 분석을 제공하며 사용자 친화적인 UI로 대화 형식 챗봇으로 구현한다는 점에서 비슷하다. 차이점은 음성 인식 여부다. 음성 소통이 가능한 Erica와 달리 NOMI는 현재까지 채팅을 통한 소통만 지원하고 있다.

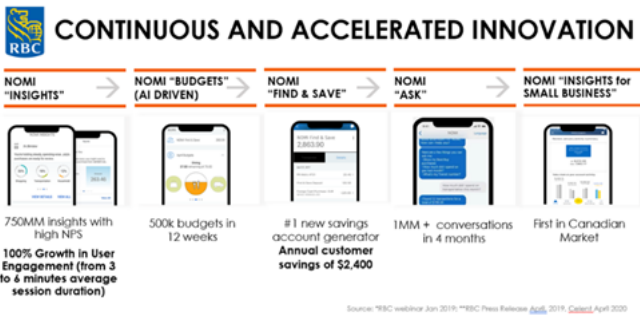
그림 163. NOMI 주요 서비스

서비스	상세 설명
NOMI Insights	개인 재무 현황 분석 및 주요 알림 제공
NOMI Find & Save	저축 여력 분석 및 자동 이체 서비스 제공
NOMI Budgets	예산을 수립하고 점검
Ask NOMI	13개 언어를 인식해 모바일 banking 업무를 처리하는 챗봇
NOMI Forecast	과거 거래를 토대로 향후 7일 동안의 지출을 예상

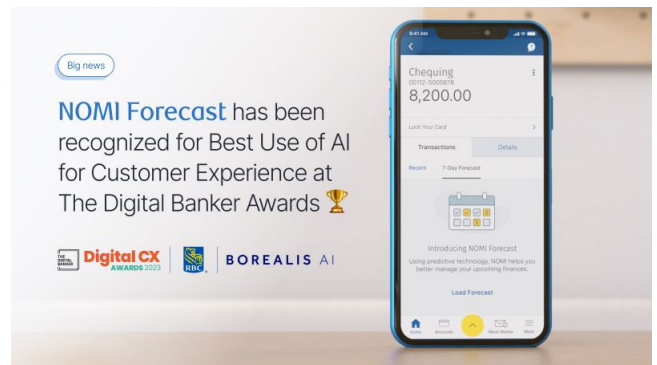
자료: RBC, 미래에셋증권 리서치센터

그림 164. NOMI를 통해 초개인화 banking에서 선두적 입지 구축

그림 165. AI를 통한 고객 경험 우수 사례: NOMI Forecast



자료: RBC, 미래에셋증권 리서치센터



자료: RBC, 미래에셋증권 리서치센터

캐피탈원(Capital One)의 Eno 서비스

Eno는 Capital One이 2017년 출시한 SMS 기반의 AI 가상 비서이다. Erica와 NOMI 처럼 머신 러닝과 자연어 처리 기술을 기반해 인간의 대화 뉘앙스를 이해해 실제 인간 처럼 대답할 수 있다. Eno는 모바일을 비롯한 인터넷 웹상에서도 지원 가능하고, 월 구독료 4.99달러(23년부터)를 받는다는 점에서 타 서비스와 다르다.

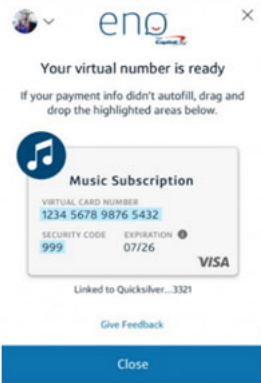
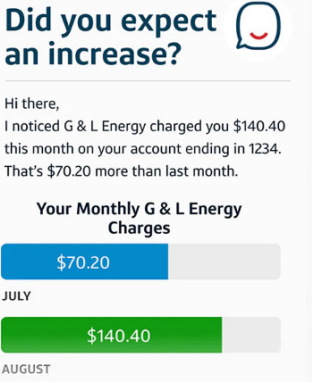
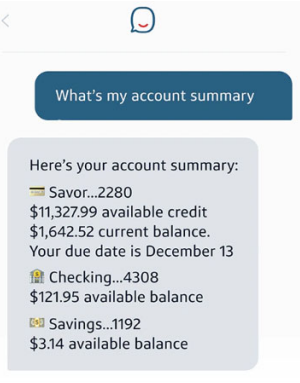
Eno의 서비스는 크게 세 가지 카테고리로 구분되는데, 1) 카드 보안(Card Security), 2) 소비 분석(Spending Insights), 3) 질의응답(Ask Eno)이다.

1) 카드 보안 서비스로는 부정 탐지, 가상 카드번호 생성 및 제공, 상시 모니터링 등이 있다. 부정 의심 징후를 실시간으로 탐지하며, 의심되는 거래나 비정상적인 계정 활동에 대해 문자 또는 이메일 알림을 보낸다. 또한 온라인 쇼핑을 하는 고객에게 가상 카드번호를 제공해 사기/부정사건 등으로부터 안전하게 카드번호를 유지하게 도와준다.

2) 소비 분석 서비스로는 무료체험 만료 안내, 결제액 관련 특이사항 모니터링, 자동결제 추이 분석, 환입 안내 등이 있다. 전체적인 자금 흐름에 대해 모니터링해주고, 고객이 필요로 하는 솔루션을 제공하고, 고객의 수고와 노력을 덜어준다.

3) 질의응답 서비스는 고객의 니즈나 궁금증을 해결해준다. 잔액, 청구기한 등 고객의 금융활동과 관련한 질의는 물론 자유대화(농담)도 가능하다. 대화의 사실감을 더하기 위해 고객의 실제 Chat 로그를 분석해 만들어졌으며, 특정 질의유형에 약 2,200가지의 표현을 이해하고 지원할 수 있다.

그림 166. Eno 주요 서비스

카드 보안	소비 분석	질의응답
		
의심 거래 포착	무료 체험 만료 안내	계좌 번호 및 잔액 확인
온라인결제 (보안용)가상카드번호 제공	결제액 증가 특이사항 확인	결제 신청
중복 거래 확인	자동결제 추이 분석	카드 잠금/해지
	환입 안내	자유대화(농담)
	카드거절 사유분석 및 지원	

자료: Capital One, 미래에셋증권 리서치센터

4. 국내 금융업계의 생성형 AI 도입 현황

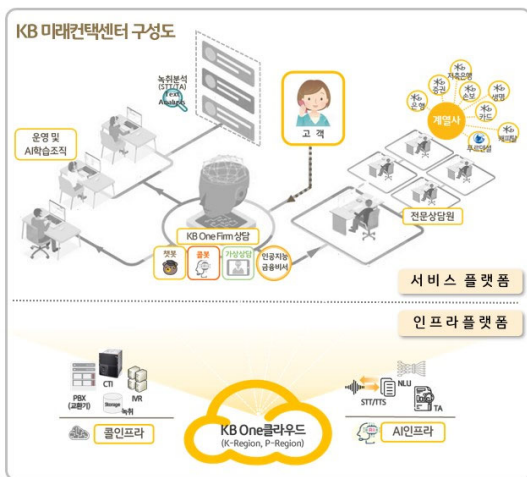
고객 접점 관리 고도화

국내 금융업계도 글로벌 흐름에 맞춰 챗봇 도입과 고도화에 힘쓰고 있다. 2017년 이후 4대 은행 모두 모바일 앱 전용 챗봇을 개발하여 운영하고 있다. 초기 챗봇은 키워드를 인식하고 시나리오에 의해 정보를 안내하는 ‘룰 기반’ 방식으로 작동되었던 반면, 현재는 지속적인 업데이트를 통해 머신러닝 방식으로 전환 중이다.

KB국민은행은 FCC(Future Contact Center) 챗봇으로 전면 개편했다. 은행 업무와 금융스케줄을 분석해 대출상환일자 등을 안내하고, 세대별 맞춤형 상품·이벤트를 추천해주는 개인화 서비스를 강화했다. 하루 평균 약 5만건의 상담을 처리하고 지속적인 학습관리를 통해 90% 수준의 답변 정확도를 기록했다. 그룹차원에서 ‘KB FCC 구축 프로젝트’를 추진해 계열사별 업무상담을 한번에 가능하게 한다는 계획을 갖고 있다.

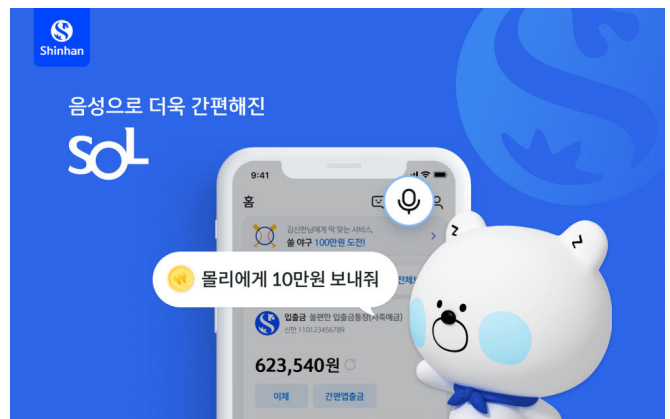
신한은행은 업계 최초로 AI 음성봇인 솔리(SOLi)에 이어, 최근에는 AI 음성뱅킹 서비스를 도입했다. 업무 내용 음성 지시를 통해 거래내역 조회, 계좌 이체 등 450여개 업무를 처리한다. 챗봇, 전화 상담 2천만건 중 유효 데이터 8만건을 분리하고 고객 사용 단어들의 분석 데이터를 탑재한 ‘자연어 이해 시스템’이 기반이 됐다. 대화형 AI 서비스로 구현하고, 모든 금융 업무를 처리할 수 있는 인비지블 은행의 핵심 채널로 활용한다는 계획이다.

그림 167. KB금융의 미래컨택센터(FCC) 구성도



자료: KB금융, 미래에셋증권 리서치센터

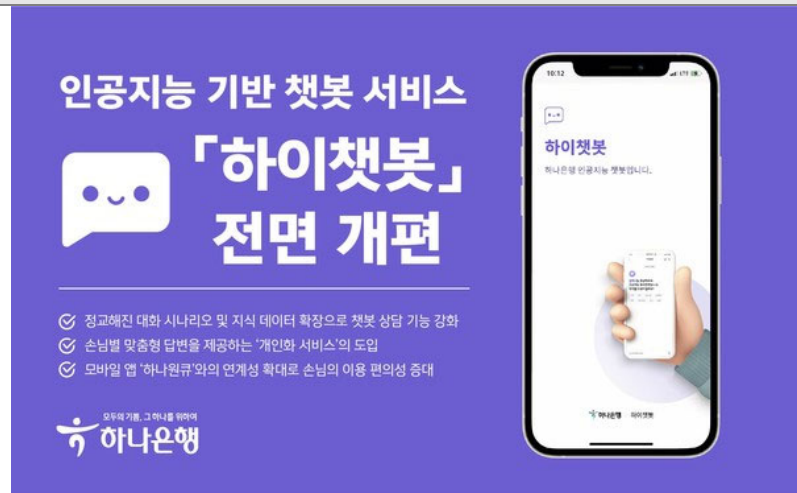
그림 168. 신한은행의 AI 음성뱅킹 서비스



자료: 신한은행, 미래에셋증권 리서치센터

하나은행은 AI 챗봇 서비스인 하이챗봇을 전면 개편했다. 대화 시나리오와 지식 데이터 확장을 위해 자체개발한 NLP(자연어처리기술) 엔진을 적용하여 챗봇의 상담 기능을 강화했다. 하이챗봇은 AI성능 고도화, 대화 시나리오 및 데이터 확장, 다른 채널로의 연계 확대 등을 지속적으로 추진해 나갈 예정이다.

그림 169. 하나은행의 AI 챗봇 서비스인 하이챗봇



자료: 하나은행, 미래에셋증권 리서치센터

표 21. 주요 국내 금융지주의 고객 접점 관리 관련 AI 도입 현황

금융사	서비스	사례
KB	챗봇	(22) KB증권, AI 기반 음성 상담 서비스 'AI 상담 콜봇' 도입 (22) 계열사 통합 미래컨택센터 'FCC(Future Contact Center) 챗봇' 도입 시작 (22) KB국민은행 기존 챗봇 서비스를 FCC 챗봇으로 전면 개편 (23) KB국민은행 FCC STT.TA 시스템 오픈: 고객센터 상담 내용을 텍스트로 변환하여 AI 및 데이터 분석 기술을 활용해 분석 (23) KB증권, KB손해보험 FCC 챗봇 서비스 출시
	로보 어드바이저	(18) 'Kbot SAM(케이봇샘)': AI와 전문가가 고객 최적 포트폴리오 제공 (21~) KB증권: 파운트, 이루다투자일임, 쿼트, 업라이즈투자자문 등과 제휴, 로보어드바이저 투자일임 서비스 제공
신한	챗봇	(18) AI 챗봇 '오로라(Orora)' 도입 (19) AI 음성봇 '솔리(SOLi)' 도입 (22) 챗봇 오로라 고도화 - 고객 맞춤형 서비스 제공 (23) 신한 솔 'AI 음성뱅크' 서비스 도입 - 음성지시 가능
	로보 어드바이저	(18) 디지털 자산관리서비스 'SOL Rich' 출시 (22) 디셈버(로보어드바이저 전문기업, 16년부터 신한은행 연계)와 5년만에 결별, 솔리치 고도화 집중
하나	챗봇	(22) AI 챗봇 서비스 하이챗봇 전면 개편: 상담 기능 강화, 개인화 서비스 도입
	로보 어드바이저	(17) 딥러닝 인공지능 로보어드바이저 하이 로보(HAI Robo) 출시, 금융상품 포트폴리오 추천 (22) 크래프트테크놀로지스와 AI기반 금융서비스 개발 업무협약 체결
우리	챗봇	(22) AI 상담(음성: AI 상담봇 & 문자: 챗봇 서비스) 고도화
	로보 어드바이저	(21) 퇴직연금 상품에 로보어드바이저 업체인 파운트 '블루웨이' 도입 (22) '우리로보-알파' 앱 서비스(17년 시작) 종료, 원업 전략 가속

자료: Deloitte, 언론자료, 미래에셋증권 리서치센터

업무 자동화

고객 접점 관리 고도화 못지않게 업무 자동화 및 규제준수에도 국내 금융권들은 시를 적극적으로 도입하고 있다.

로봇프로세스 자동화(RPA, Robotic Process Automation)는 소프트웨어 프로그램이 사람을 대체해 반복 업무를 자동 처리하는 기술로, 정확성과 효율성을 제고한다.

신한은행은 '알파봇(RPA bot)'이라는 이름으로 전 영업점에서의 RPA 사용 확대해 고객 상담 및 업무처리에서 활용 가능하다. 22년부터는 전 직원 대상 'R 비서'를 구축하여 은행원 개인에 맞춤형 업무 솔루션 제공까지 꾀하고 있다. KB 국민은행은 21년 31개 업무에 RPA 를 적용했으며 우리은행은 19년 RPA 도입 이후 59개 업무에 적용하는 등 자동화 확대하고 있다. 하나은행 또한 RPA 도입을 통해 연간 약 150만 시간의 직원 업무 경감 효과를 창출했다고 보도했다.

표 22. 주요 국내은행의 업무 자동화 관련 AI 도입 현황

은행	서비스	사례
KB	RPA	(17) 본부 업무 RPA(Robotic Process Automation) 자동화 구축 시작 (22) RPA 도입 240개 업무로 확장, 이를 통한 290만 시간 절약 발표
	신용평가	(22) '기업여신 자동심사 지원시스템(Bics: Big data CSS)' 도입
신한	RPA	(17) 신한은행, 은행권 최초 여신업무에 RPA 프로그램 도입 (21) 신한은행, 영업점 업무 자동화 프로세스 '알파봇(RPA bot)' 도입 (22) 'R비서' 구축 및 개발 시범 도입 → 23년 보급
	신용평가	(19) 기업의 정형, 비정형 데이터를 점수화해 기업여신 심사하는 기업CSS 도입 (22) 인공지능 플랫폼 SACP(Shinhan AI Core Platform)에 머신러닝 자기학습 프로세스 이행, 신용평가모형 정교화
하나	RPA	(23) 로봇 자동화 프로그램 개발 플랫폼을 구축, '1부서 1봇' 프로젝트 시행
	신용평가	(17) 기업여신 자동심사시스템 첫 개시 (21) 하나은행, 머신러닝 적용 대출한도모형 기반 'AI대출' 출시
우리	RPA	(19) RPA 도입 (21) RPA 3단계 확대 시행
	신용평가	(18) 빅데이터 활용 기업진단시스템 '빅아이(Big Eye)' 도입 (19) '기업여신 자동심사시스템' 도입

자료: Deloitte, 언론자료, 미래에셋증권 리서치센터

이상거래탐지(FDS) 및 자금세탁방지(AML)

규제준수와 리스크 탐지 관련 기술은 '레그테크'에 속한다. 국내외 금융사들은 AI 및 빅데이터 기술을 활용해 업무를 보다 빠르고 정확하게 진행할 수 있도록 보조하고자 한다. 이를 통해 금융사고를 방지할 수 있으며, 규제 준수 비용이 절감 가능하다.

이상거래탐지시스템은 불법이체 등 의심거래를 실시간 분석하고 탐지하는 시스템이다. KB국민은행은 'AI보이스피싱 모니터링 시스템' 고도화로 이상거래 탐지율을 34.4%까지 높였으며, 하나은행은 AI기반 신 이상금융거래탐지시스템을 통해 22년 보이스피싱 위험으로부터 1,814억원의 고객자산을 보호했다고 밝혔다.

자금세탁방지는 불법자금세탁을 적발하고 예방하기 위한 시스템이다. 현재 AML에의 AI 활용은 시나리오 기반 작동에 머물러 상대적으로 제한적이다. 그러나 기술 개발의 가속화로 광학문자인식(OCR)과 자연어처리기술(NLP) 수준이 올라가며 비정형 문서에도 유의미한 정보를 선별해 자금세탁 의심 사례 전수조사가 가능해지고 있다.

표 23. 주요 국내은행의 규제 준수 관련 AI 도입 현황

은행	서비스	사례
KB	이상거래 탐지시스템	(22) 'AI 보이스피싱 모니터링 시스템' 고도화
	자금세탁방지	(21) 스마트시재관리기개시 (22) 뉴 KB 글로벌(NEW KB Global) AML 시스템 구축
신한	이상거래 탐지시스템	(22) 'AI 이상행동탐지 ATM' 도입
	자금세탁방지	(20) 자금세탁방지 고도화 프로젝트 완료: 머신러닝을 활용해 자금세탁 위험도 측정 모델 개발, 탐지 정확도 향상
하나	이상거래 탐지시스템	(18) 전기통신금융사기(대포통장, 보이스피싱) 사고패턴 AI 학습분석, 'AI 기반 신 이상금융거래탐지시스템(FDS)' 도입 (23) 모바일 앱에 보이스피싱 탐지기능 신규 탑재
	자금세탁방지	(23) '자금세탁방지'(AML, Anti-Money Laundering)를 위한 머신러닝(Machine Learning)모형 자체 개발
우리	이상거래 탐지시스템	(19) 인공지능기반 이상거래탐지시스템(AI-FDS) 은행권 최초 상용화
	자금세탁방지	(20) 글로벌통합AML시스템 구축

자료: Deloitte, 언론자료, 미래에셋증권 리서치센터

5. 금융기관의 생성형 AI 기술 채택 증가 전망

(국제금융센터의 ‘글로벌 은행권의 ChatGPT 활용 가능성 및 과제(23.2.28 발간)’를 인용)

향후 생성형·대화형 AI가 금융권에 거대한 임팩트를 끼칠 것이라는 사실은 자명하다. 최근 다양한 기관에서 진행한 서베이들의 결과들이 이를 증명하고 있다.

- ① 금융서비스업 종사자 가운데 58%가 AI가 회사의 미래 성공을 위해 중요하다고 생각 (Bank Automation News)
- ② 22년 영국 내 금융기관 중 72%(19년 67%)가 머신러닝 애플리케이션을 사용하거나 개발 중이라고 보고 (Bank of England)
- ③ 트레이더 중 53%가 향후 3년간 AI가 자신이 맡은 업무에 가장 큰 영향을 미칠 것이라고 생각 (JPMorgan)

일각에서는 금융업 등 대부분의 기업들이 ChatGPT 등의 생성형 AI에 완벽하게 준비되지 않은 상황이라 지적하기도 하고, 그것을 중심으로 운영되기에는 너무 이르다고 평가하는 시기상조론도 제기한다.

하지만 다수의 전문가들은 ChatGPT가 공개된 22년을 생성형 AI 기술의 전환점으로 보고, 향후 수년 동안 금융기관의 동 기술 채택이 크게 증가할 것으로 내다보고 있다. 향후 생성형 AI가 지금보다 강력하고 다재다능하게 발전하면서 금융회사의 다양한 업무 및 애플리케이션 등에 널리 사용될 것으로 전망한다.

특히 ChatGPT의 등장으로 단순 Q&A 기능에 머물렀던 기존의 금융 챗봇이, 고객과 포괄적 대화를 나누고 요구를 신속히 해결하는 AI 금융 비서로 진화하고 있다. 우리는 앞서 BofA, RBC, Capital One의 사례를 통해 이러한 AI 금융 비서가 해당 기업의 생산성 및 효율성뿐만 아니라 고객경험 개선에 도움이 된 것을 살펴봤다.

그림 170. 챗봇 문의 사안별 해결 비율(Gartner)

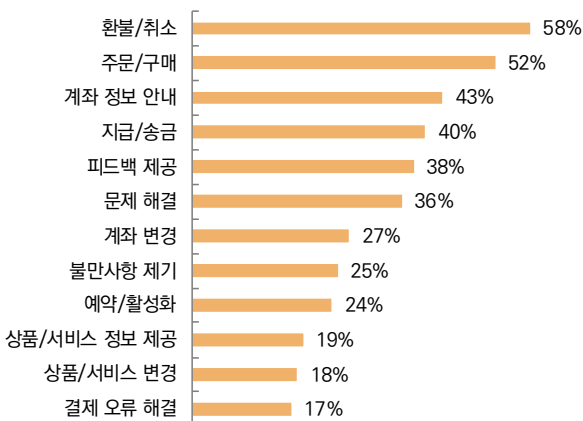
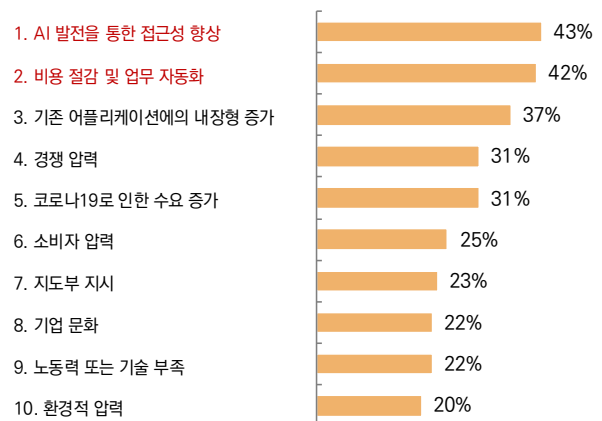


그림 171. AI 채택을 주도하는 상위 10가지 요인(IBM)



자료: 가트너(23.06), 미래에셋증권 리서치센터

자료: IBM(2022), 미래에셋증권 리서치센터

6. 생성형 AI 도입을 위해 풀어나가야 할 과제

(국제금융센터의 ‘글로벌 은행권의 ChatGPT 활용 가능성 및 과제(23.2.28 발간)’를 인용)

생성형 AI의 발전 잠재력에도 불구하고 넘어서야 할 몇 가지 과제들은 잔존한다. 1) 정보 신뢰성 한계 2) 고객 데이터 보호 우려 3) 도입비용 부담 4) 규제 도입이 대표적이다.

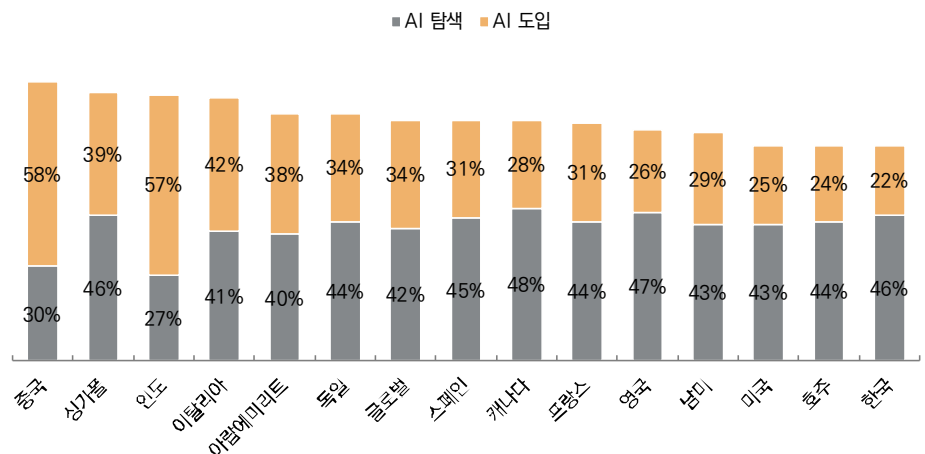
1) 정보 신뢰성 한계는 생성형 AI가 기본적으로 기계학습에 기반하기 때문에 발생한다. AI 프로그램은 자의식이 없는 만큼 훈련 데이터의 품질이 좋지 않은 경우 잘못된 정보를 제공할 수 있다. 따라서 정보를 선별하거나 진실을 확인할 때는 인간이 개입해야 된다.

또한 2) 고객 데이터 보호는 AI가 방대한 데이터 수집을 통해 훈련되는 만큼, 개인의 프라이버시 침해에 대한 우려가 존재한다. 특히나 개인들의 금융 정보는 매우 민감하다. 생성형 AI 엔지니어들의 무단 접근이 심각한 결과를 초래할 우려도 존재하며, 외부 해킹 세력들의 생성형 AI로의 침입 시도 가능성도 배제할 수 없다.

3) 도입비용 부담이 존재한다. 생성형 AI의 도입을 위한 상당한 컴퓨팅 자원은 물론 사내 통합, 교육 등 운영비와 규제비와 같은 부수적 비용까지 발생한다. 또한 향후 AI가 고도화됨에 따라 프로그램 가격 인상 가능성도 존재한다. 실제 22년에 Microsoft가 보안 강화와 AI 기능 추가 등을 이유로 구독 요금을 약 10% 인상한 바 있음을 참고할 필요 있다. 이는 소형 금융사들에겐 부담으로 작용한다.

4) 마지막 과제로 관련 규제가 정비될 필요 있다. AI 발전 및 활용에 속도가 붙으며 2~3년여전부터 주요국들을 중심으로 관련 규제를 검토하기 시작했다. 캐나다는 22년 6월 AI 법을 시행했으며 EU의 경우 23년 3월 AI 법 초안을 공개한 후 25년 중 발효가 예상된다. 미국은 22년 10월 백악관과학기술정책국이 AI 윤리지침을 공개하는 등 관련 규제 마련에 관심을 보이고 있다. 금융권은 앞으로 규제 정비에 미리 대비하면서 또한 초기 정립된 규제를 준수하기 위한 노력이 필요할 것이다.

그림 172. 국가별 AI 채택 지수(IBM)



자료: IBM(2022), 미래에셋증권 리서치센터

Global Company Analysis

(Analyst) 김영건 younggun.kim.a@miraeeasset.com



 미래에셋증권

NVDA US · 글로벌 테크 · 미국

엔비디아

빅테크들의 경쟁은 동사에 유리

Not Rated

Refinitiv 평균목표주가

USD 520.85

상승여력

19.7%

현재주가(2023/10/3)

USD 435.17

나스닥종합주가(p)	13,059.47	시가총액(십억USD)	1,074.87	시가총액(조원)	1,462.90	유통주식수(백만주)	2,470.00
------------	-----------	-------------	----------	----------	----------	------------	----------

Report summary

데이터센터 중심의 GPU 강세

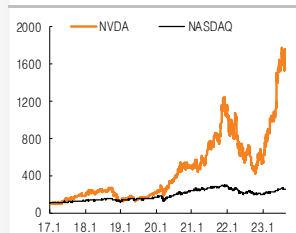
엔비디아의 23F 매출액은 \$27B(YoY +0.3%) 성장할 것으로 전망한다. AI 수요 증가로 인해 하이퍼스케일 러향 GPU 공급이 빠르게 증가하고 있다. 기업체들의 비용 부담에 따른 클라우드 성장률이 일시적 둔화기를 겪고 있는 상황임에도 불구하고 AI 선제적 투자의 중요성을 반증한다. 동사의 최근 데이터센터향 GPU 매출액은 대부분 클라우드 업체들이나, 매크로 상황 개선과 함께 개별기업(Enterprise)향 매출액도 회복될 것으로 기대한다.

데이터센터 중심의 GPU 강세

AI 추론시장에서도 CUDA 기반의 동사의 시장 점유가 지속될 가능성이 높다고 판단하며, 긍정적인 전망을 유지한다. 비록 빅테크 업체들의 자체 프레임워크와 라이브러리를 구축할 의지와 역량이 있지만, 헤게모니 확보 경쟁 가운데 개발 소스를 상호 오픈하기는 구조적으로 어려워 보인다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	6.3	99.1	174.2
상대주가	8.7	80.1	155.2

Earnings and valuation metrics

결산기 (1월)	1/22	1/23	1/24F	1/25F	1/26F
매출액 (십억USD)	26.9	27.0	44.0	61.4	74.2
영업이익 (십억USD)	10.0	4.2	22.8	33.5	40.6
영업이익률 (%)	37.31	15.66	51.78	54.58	54.65
순이익 (십억USD)	9.75	4.37	17.06	25.41	25.40
EPS (USD)	3.9	1.8	7.1	10.6	12.1
ROE (%)	44.8	17.9	59.1	57.8	50.1
P/E (배)	59.0	83.3	57.8	38.9	31.6
P/B (배)	21.5	22.7	32.1	19.6	13.0

주: GAAP 기준, 순이익은 지배주주 귀속 순이익
 자료: 엔비디아, Bloomberg, 미래에셋증권 리서치센터

추론시장에서도 지속되는 경쟁력

동사가 GPU 높은 점유율을 유지할 수 있는 이유는 프로세서 자체의 경쟁력 뿐 아니라 자사 GPU 기반의 CUDA(쿠다) 소프트웨어 생태계의 요인이 크다. CUDA를 활용한 cuDNN(CUDA Deep Neural Network)이 AI 라이브러리를 선점함에 따라 이후 대부분의 AI 개발자는 이 라이브러리 위에 AI 모델을 구축하기 시작했다.

AMD와 Intel이 각각 ROCm와 원API를 출시하며 CUDA 기반 GPU 소프트웨어 생태계를 대체하겠다고 나서고 있으나, 동사가 십수년간 쌓은 영역을 후발주자가 단기간에 따라잡기는 쉽지 않은 상황이다. 동사 또한 DNN 라이브러리 뿐 만 아니라 다수의 CUDA 기반 소프트웨어 생태계를 구축하는데 대규모 자본을 투입하는 중이다.

추론 인프라는 동사의 GPU 대체 여지가 일부 있다. 그러나, 하드웨어 제조사마다 칩에서 지원하는 기능이 동일할 수 없으므로 학습 단계부터 추론 인프라를 고려해 AI 모델을 설계하는 수밖에 없다. 이런 설계역량을 갖춘 주체는 극소수다. Meta와 Google, MSFT, AWS 등은 자체적인 AI 스택을 보유했거나 구축중인 것으로 알려져 있다.

그러나, 상기 빅테크 업체들이 추론용 자체 AI 소프트웨어 스택을 상호 공개할 가능성은 점차 줄고 있다. AI 주도권을 확보하기 위한 경쟁이 심화되고 있기 때문이다. NPU 프로세서 개발사들 또한 하드웨어 설계와 함께 소프트웨어 스택 개발에도 나선 사례가 있으나, GPU에 대한 완전한 독립은 불투명하다.

당사는 AI 추론시장에서도 NVIDIA의 GPU 및 생태계 시장 점유가 지속될 가능성이 높다고 판단하며, NVIDIA에 대한 긍정적인 전망을 유지한다. 비록 빅테크 업체들의 자체 프레임워크와 라이브러리를 구축할 의지와 역량이 있지만, 헤게모니 확보 경쟁 가운데 개발 소스를 상호 오픈하기는 구조적으로 어려워 보인다. 빅테크사의 자체 추론 모델에의 NVIDIA 의존도는 떨어질 수 있겠으나, 그 외의 대부분의 추론 모델 운영주체에게는 CUDA외의 대안이 마땅치 않은 것으로 판단한다.

표 24. 데이터센터 시장 및 엔비디아 데이터센터 성장률 추정

데이터센터용 GPU 수요 전망	
Data Center CAGR Assumption	8%
Data Center Value in 2027 (US\$mn)	180,778
GPU 2016-2022 CAGR	55.2%
Bear Case	
GPU % of total value	20%
Expect GPU value (US\$mn) (Bear Case)	36,156
Expect CAGR in 5 years	25%
Consensus CAGR in 5 years	21%
Base Case	
GPU % of total value	30%
Expect GPU value (US\$mn) (Base Case)	54,233
Expect CAGR in 5 years	36%
Consensus CAGR in 5 years	21%
Bull Case	
GPU % of total value	40%
Expect GPU value (US\$mn) (Bull Case)	72,311
Expect CAGR in 5 years	44%
Consensus CAGR in 5 years	21%

자료: IDC, 미래에셋증권 리서치센터 추정 (생성AI: 제2의 기계시대)

표 25. Nvidia 실적 전망

(US\$mn)

	FY2023				FY2024F				FY22	FY23	FY24F
	1Q	2Q	3Q	4Q	1Q	2Q	3QF	4QF			
Revenue (Total)	8,288	6,704	5,931	6,051	7,192	13,507	16,231	17,405	26,914	26,974	54,335
Platform Segments											
Gaming	3,620	2,042	1,574	1,831	2,240	2,486	2,610	2,715	12,462	9,067	10,051
Professional Visualization	622	496	200	226	295	379	387	414	2,111	1,544	1,474
Datacenter	3,750	3,806	3,833	3,616	4,284	10,323	12,904	13,936	10,613	15,005	41,447
Automotive	138	220	251	294	296	253	261	274	566	903	1,083
OEM & IP	158	140	73	84	77	66	69	67	1,162	455	280
Platform Segments (% of sales)											
Gaming	44%	30%	27%	30%	31%	18%	16%	16%	46%	34%	18%
Professional Visualization	8%	7%	3%	4%	4%	3%	2%	2%	8%	6%	3%
Datacenter	45%	57%	65%	60%	60%	76%	80%	80%	39%	56%	76%
Automotive	2%	3%	4%	5%	4%	2%	2%	2%	2%	3%	2%
OEM & IP	2%	2%	1%	1%	1%	0%	0%	0%	4%	2%	1%
Platform Segments (QoQ)											
Revenue (Total)	8.4%	-19.1%	-11.5%	2.0%	18.9%	87.8%	20.2%	7.2%			
Gaming	5.8%	-43.6%	-22.9%	16.3%	22.3%	11.0%	5.0%	4.0%			
Professional Visualization	-3.3%	-20.3%	-59.7%	13.0%	30.5%	28.5%	2.0%	7.0%			
Datacenter	14.9%	1.5%	0.7%	-5.7%	18.5%	141.0%	25.0%	8.0%			
Automotive	10.4%	59.4%	14.1%	17.1%	0.7%	-14.5%	3.0%	5.0%			
OEM & IP	-17.7%	-11.4%	-47.9%	15.1%	-8.3%	-14.3%	5.0%	-3.0%			
Platform Segments (YoY)											
Revenue (Total)	46.4%	3.0%	-16.5%	-20.8%	-13.2%	101.5%	173.7%	187.6%	61.4%	0.2%	101.4%
Gaming	31.2%	-33.3%	-51.1%	-46.5%	-38.1%	21.7%	65.8%	48.3%	61%	-27%	11%
Professional Visualization	67.2%	-4.4%	-65.3%	-64.9%	-52.6%	-23.6%	93.3%	83.0%	100%	-27%	-5%
Datacenter	83.1%	60.9%	30.6%	10.8%	14.2%	171.2%	236.6%	285.4%	58%	41%	176%
Automotive	-10.4%	44.7%	85.9%	135.2%	114.5%	15.0%	3.8%	-6.9%	6%	60%	20%
OEM & IP	-51.7%	-65.8%	-68.8%	-56.3%	-51.3%	-52.9%	-5.1%	-20.0%	84%	-61%	-39%
GPM(%)	66%	46%	56%	63%	65%	70%	72%	72%	66%	58%	70%
Operating Income	3,955	1,325	1,536	2,224	3,052	6,800	8,034	8,877	11,269	10,580	26,763
Net income	3,443	1,292	1,456	2,174	2,713	6,188	7,036	7,765	10,403	5,780	23,702

자료: Nvidia, 미래에셋증권 리서치센터 추정

표 26. Global hyperscaler capex 및 processor 매출액 전망

Hyperscaler Capex(US\$bn)	2017	2018	2019	2020	2021	2022	2023F	2024F	2025F
Microsoft	11.4	15.8	18.0	20.6	27.5	28.4	43.5	50.8	58.7
Amazon	12.0	13.4	16.9	40.1	61.1	63.6	50.7		
AWS	9.2	9.8	13.1	16.5	22.0	27.8	30.0	34.3	41.2
Alphabet	13.2	25.1	23.5	22.3	24.6	31.5	32.2	37.2	45.8
Meta	6.7	14.0	15.7	15.7	19.2	32.0	28.4	33.2	37.4
Apple	12.0	13.7	9.2	8.7	10.4	11.7	13.0	14.4	16.1
Baidu	0.7	1.3	0.9	0.8	1.7	1.2	1.6	1.7	1.8
Alibaba	3.6	4.9	3.8	5.1	6.0	5.5	4.7	5.2	6.2
Sum	56.9	84.5	84.2	89.8	111.6	138.1	153.4	176.8	207.2
Capex(N)/Internet sales(N-1) portion(%)									
Microsoft	44.4	54.0	50.8	47.1	51.7	42.0	53.3	55.0	55.0
AWS	75.2	56.0	50.9	47.2	48.6	44.6	37.5	37.0	37.0
Alphabet				92.6	75.1	65.5	58.0	60.0	60.0
Meta	24.4	34.4	28.0	22.2	22.4	27.2	24.4	25.0	25.0
Apple	47.1	43.8	22.3	18.1	18.3	16.2	16.4	17.0	17.0
Baidu	7.2	9.8	6.2	5.0	10.4	6.1	8.7	9.0	9.0
Alibaba	443.1	278.1	117.1	101.3	72.3	49.0	40.9	42.0	42.0
Average	55.8	63.4	47.7	37.1	37.3	34.6	34.7	35.6	36.1
Processor sales(US\$bn)									
Intel	62.8	70.8	72.0	77.9	79.0	63.1	52.5	58.7	63.9
AMD	5.3	6.5	6.7	9.8	16.4	23.6	22.8	27.5	31.8
NVIDIA	9.7	11.7	10.9	16.7	26.9	27.0	53.7	78.1	90.9
Sum	77.7	89.0	89.6	104.3	122.4	113.6	129.0	164.3	186.5
Processor DC portion(%)									
Intel	30.4	32.5	32.6	33.5	28.7	30.4	32.3	32.0	34.0
AMD					22.5	25.6	27.3	35.0	35.0
NVIDIA	19.9	25.0	27.3	40.2	39.4	55.6	77.4	85.0	85.0
Sum	27.0	29.1	29.5	31.4	30.2	35.4	50.2	57.7	59.0
Processor sales for DC(US\$bn)									
Intel	19.1	23.0	23.5	26.1	22.7	19.2	16.9	18.8	21.7
AMD					3.7	6.0	6.2	9.6	11.1
NVIDIA	1.9	2.9	3.0	6.7	10.6	15.0	41.6	66.4	77.2
Sum	21.0	25.9	26.5	32.8	37.0	40.2	64.7	94.8	110.1
Processor sales/IDC Capex(%)	36.9	30.7	31.4	36.5	33.2	29.1	42.2	53.6	53.1

자료: 각 사, Refinitiv, 미래에셋증권 리서치센터 추정

(Analyst) 김수진 soojin.kim@miraeasset.com



MSFT US · 글로벌 소프트웨어 · 미국

마이크로소프트

Microsoft Copilot 공식 발표 – MS 365 Copilot 11월 출시

(유지) 매수	목표주가 USD 433.4	상승여력 38.3%	현재주가(23/10/3) USD 313.39
나스닥 종합(p)	13,223.99	시가총액(십억USD)	2,345.95
		시가총액(조원)	3,190.67
		유통주식수(백만주)	7,441

Report summary

MS365부터 윈도우, 검색엔진 Bing, 웹브라우저 Edge까지 전방위 생성AI 도입 계획 발표

11월 1일 Microsoft365 Copilot을 비롯해 본격적으로 Copilot 제품군을 런칭하는 일정을 발표했다. 앞선 9월 25일(현지시각)에는 Windows11 업데이트로 생성AI 기능이 대거 적용된다. 추가로 검색엔진 Bing과 웹브라우저 Edge에도 Copilot이 적용된다. Bing에는 이미지 생성AI DALL-E 3 적용도 시작된다. 따라서 올해 말부터 MS365 Copilot과 Windows11 전환이 동시에 진행되며 매출 반등이 가능할 것으로 기대한다. MS의 생성AI 기능들은 '마이크로소프트 코파일럿'이란 명칭으로 통합된다.

다만 제품이 출시되었지만, 실제로 솔루션이 구축되는 데에 시간이 소요되기 때문에 매출이 당장 발생하지는 않는다. 따라서 본격적인 매출 인식은 내년부터 일 것으로 전망하므로, 본격적인 모멘텀은 CY3Q23 실적 발표 이후 CY24년 가이던스가 구체화되는 시점부터 발생할 것으로 예상된다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	6.8	41.5	39.3
상대주가	1.2	25.8	19.6

Earnings and valuation metrics

결산기 (06월)	2020	2021	2022	2023	2024F
매출액 (백만USD)	14,3015	168,088	198,270	211,915	234,772
영업이익 (십억USD)	52,959	69,916	83,383	88,523	98,761
영업이익률 (%)	37.0	41.6	42.1	34.1	42.1
순이익 (십억USD)	44,281	61,271	72,738	72,361	82,562
EPS (USD)	5.82	8.12	9.70	9.72	12.7
ROE (%)	40.1	47.1	47.2	47.1	32.3
P/E (배)	35.1	34.2	28.0	35.0	27.7
P/B (배)	13.0	14.3	11.5	11.5	7.8

주: GAAP 기준, 순이익은 지배주주 귀속 순이익

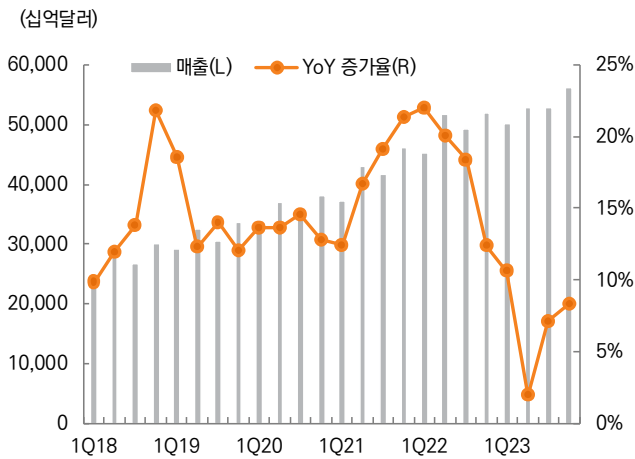
자료: 마이크로소프트, Refinitiv, 미래에셋증권 리서치센터

MS, Microsoft Copilot 공식 발표 – MS365 Copilot은 11월 1일 런칭

OpenAI에 Azure 인프라를 투자한 뒤 ChatGPT를 출시하며 주도권을 잡은 마이크로소프트가 9월 드디어 Microsoft Copilot 출시를 본격화했다. MS의 생성AI 이름을 모두 ‘마이크로소프트 코파일럿’으로 통합했다. MS는 빅테크 기업들 가운데 생성AI 어플리케이션으로 가장 먼저 어플리케이션을 통해 직접적으로 수익화를 실현할 수 있을 것으로 기대한다. 다만 그 시장 규모나 파급력을 아직까지 예측하기 어려운 상황이다. 생성AI 솔루션의 매출 인식은 올해 말부터 시작되기 때문에 내년부터 반영될 것으로 전망한다.

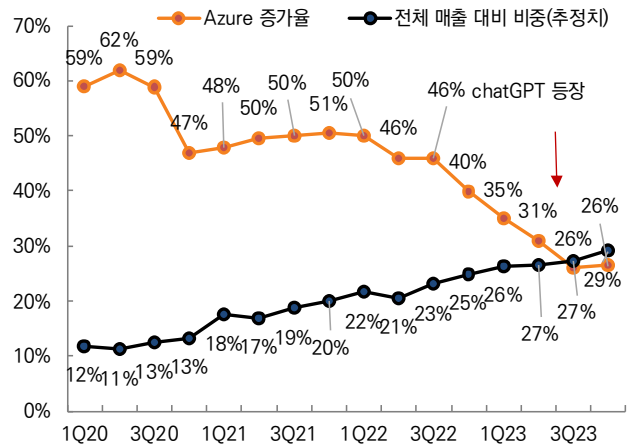
현재 Copilot에 집중하고 있는 관계로 클라우드 컴퓨팅 매출 성장세가 소폭 둔화되고 있다. 따라서 CY3Q23 실적을 통해 MS가 클라우드 비즈니스에 얼마나 집중하고 있는지를 확인할 수 있을 것으로 예상된다.

그림 173. 마이크로소프트 매출



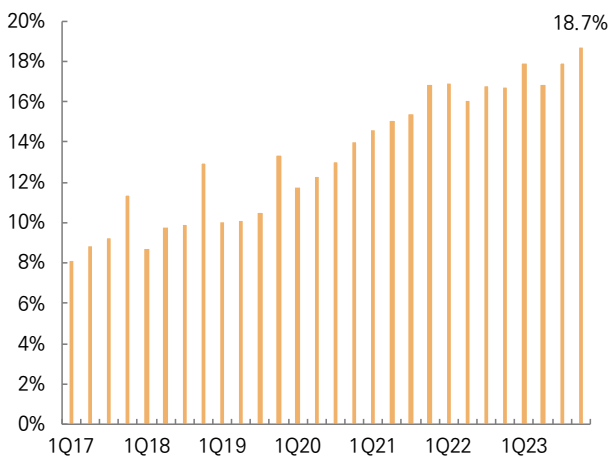
자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 174. Azure 매출 증가율



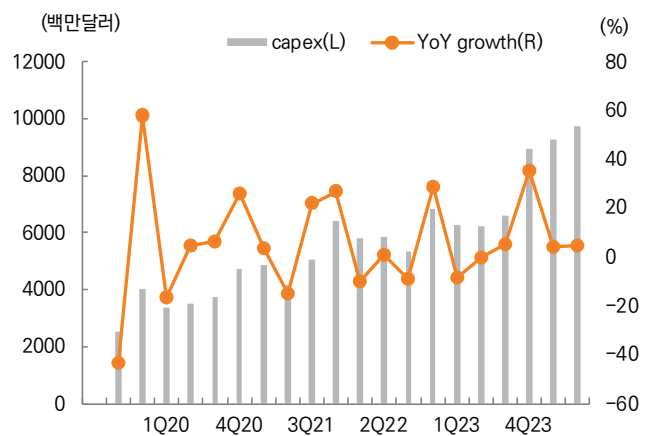
자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 175. Intelligent cloud opm



자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 176. CapEx



자료: 마이크로소프트, 미래에셋증권 리서치센터

① **Microsoft 365 Copilot**

9월 21일 'MS Copilot(코파일럿)'의 정식 런칭을 발표했다. Copilot은 AI 비서로 생산성 솔루션에서 사용된다. 가장 처음 Copilot이 쓰인 서비스는 2022년 개발 도구 Github Copilot이었으며, 올해 3월에는 오피스에서 사용될 'Microsoft 365 Copilot'과 보안에서 사용될 'Microsoft Security Copilot'도 공개되었다. 5월에는 'Windows Copilot'에 대한 언급됐고, 실제 제품 출시일이 9월에 발표된 것이다. Copilot은 윈도우11, 마이크로소프트 365, 엣지 및 Bing 등에 적용될 예정이다.

우선 가장 기대되는 제품은 Microsoft 365 Copilot이다. 런칭 예정일은 11월 1일이다. 워드부터 사용할 수 있다고 설명했다. 한국어 버전은 내년 하반기로 예상한다. 가격은 1인당 월 30달러로 책정됐다. 현재 Microsoft 365 E3 가격이 36달러인데(E5는 57달러), 비교해서 추가로 30달러가 더해져 66달러로 83% 인상되는 것이다. 따라서 시나리오 분석에 따르면 이용자가 10% 증가할 경우, 매출 증가율이 12% 수준을 기록할 것으로 예상한다.

② **Microsot Security Copilot**

보안에서도 챗봇 형태의 Copilot 이 적용된다. 이 제품이 필요한 이유는, 사이버공격이 늘어나는 상황에서 보안 전문가를 챗봇이 대체할 수 있기 때문이다. 또 문제가 발생했을 때, 문제 발생 이유와 해결 방법을 챗봇에게 물어 해결책을 찾아낼 수 있다.

③ **Windows Copilot**

9 월 25 일 윈도우 업데이트를 통해 150 개 이상의 신규 기능이 제공되는데, 이 때 코파일럿 연동, AI 기능이 추가된 그림판, 노트패드, 자연스러운 음성을 지원하는 내레이터 등 AI가 추가된 기능이 대거 등장했다.

④ **Bing과 Edge 브라우저**

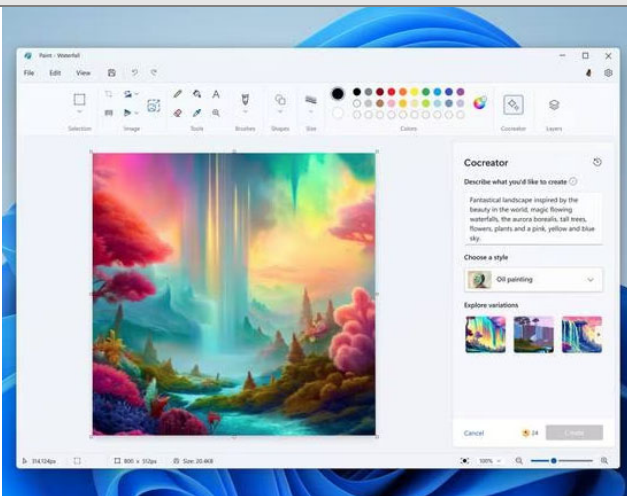
이미 3 월부터 gpt 를 Bing 에서 사용할 수 있었으나, 이제 이미지 AI 인 Dall.E3 도 지원된다. 텍스트 검색을 통해 이미지를 생성할 수 있는 것이다. 또 모든 이미지에는 생성된 시간과 날짜를 포함해 암호화된 디지털 워터마크가 추가된다. 이를 콘텐츠 자격증명(content credentials)라 한다. 마이크로소프트 엣지 모바일 앱에서도 AI 기능을 사용 가능하며, 빙챗 엔터프라이즈에서는 멀티모달 비주얼 서치와 이미지 크리에이터 기능이 추가된다.

표 27. 마이크로소프트 사이버보안 제품

제품명	기능	대표 peer	대형 보안 기업 출시 제품
Microsoft Security Copilot	생성AI로 위협 발견, 조사 및 대응	클라우드스트라이크(CRWD US) 외 등장 중	
Microsoft Sentinel	SIEM	Splunk(SPLK US), Elastic, Rapid7	아마존, 포티넷, IBM, 포티넷, 팔로알토네트웍스, 오라클,
Microsoft Defender	XDR	클라우드스트라이크(CRWD US), 센티널원(S US), 트린드마이크로(4704 JP)	포티넷, 시만텍
Microsoft Intune	모바일 디바이스 관리 보안	Jamf(JAMF US)	VM웨어, 시스코, Sophos, 시트릭스, AWS, IBM
Microsoft Entra	ID 보안	OKTA(OKTA US)	포티넷, 아마존, 구글, IBM
Microsoft Purview	데이터/거버넌스 보안	informatica(INFA US)	IBM, SAP

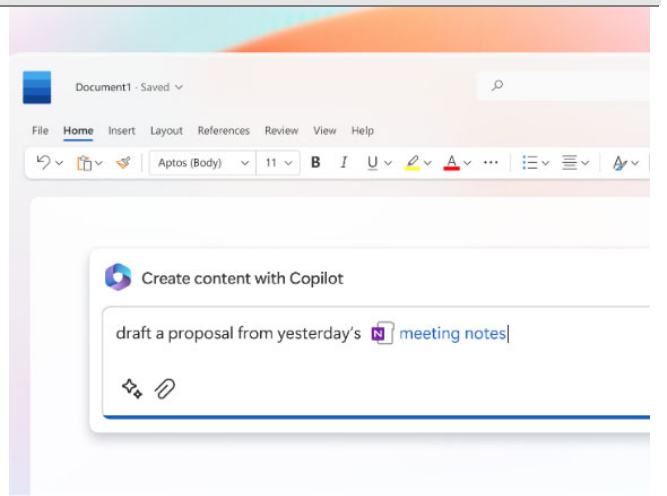
자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 177. 윈도우 그림판에 AI 기능 추가



자료: 마이크로소프트, 미래에셋증권 리서치센터

그림 178. Microsoft Copilot



자료: 마이크로소프트, 미래에셋증권 리서치센터

표 28. Microsoft 365 Copilot 매출 시뮬레이션

(백만달러, 백만명)

	1Q24E	2Q24E	3Q24E	4Q24E
커머셜 구독자(백만명)	262.4621	265.0868	273.8504	274.7171
월 30달러 Copilot 이용 가정시 추가 발생 가능한 매출/전체 커머셜 구독자 중 비중 가정				
5%	394	398	411	412
10%	1,181	1,193	1,232	1,236
15%	5,315	5,368	5,545	5,563
오피스 매출 예상치	14,629	14,928	15,573	16,102
YoY growth(%)	9.9%	12.0%	13.0%	12.5%

자료: 마이크로소프트, 미래에셋증권 리서치센터

표 29. 실적추정표

(백만달러)

	FY2022				FY2023				FY2024F				FY2022 22/06	FY2023 23/06	FY2024F 23/07
	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23	4Q23	1Q24F	2Q24F	3Q24F	4Q24F			
	21/09	21/12	22/03	22/06	22/09	22/12	23/03	23/06	23/09	23/12	23/03	23/06			
총매출	45,317	51,728	49,360	51,865	50,122	52,747	52,857	56,189	53,868	58,105	58,656	64,143	198,268	211,915	234,758
YoY growth(%)	22.0%	20.1%	18.4%	12.4%	10.6%	2.0%	7.1%	8.3%	7.5%	10.2%	11.0%	14.2%	18.0%	7.6%	10.0%
PBP	15,039	15,936	15,789	16,600	16,465	17,002	17,516	18,291	18,102	18,960	19,563	20,426	63,099	69,511	77,052
YoY growth(%)	22.1%	19.3%	16.5%	13.0%	9.5%	6.7%	10.9%	10.2%	9.9%	11.5%	11.7%	11.7%	17.0%	10.2%	10.8%
매출 대비 비중(%)	33.2%	30.8%	32.0%	32.0%	32.8%	32.2%	33.1%	32.6%	33.6%	32.6%	33.4%	31.8%	31.8%	32.6%	32.8%
IC	16,964	18,327	19,051	20,909	20,325	21,508	22,081	23,993	23,389	24,756	25,109	28,337	75,251	87,835	101,591
YoY growth(%)	30.6%	25.5%	26.0%	20.3%	19.8%	17.4%	15.9%	14.7%	15.1%	15.1%	13.7%	18.1%	25.3%	16.7%	15.7%
매출 대비 비중(%)	37.4%	35.4%	38.6%	40.3%	40.6%	40.8%	41.8%	42.7%	43.4%	42.6%	42.8%	44.2%	38.0%	41.2%	43.3%
MPC	13,314	17,465	14,520	14,356	13,332	14,237	13,260	13,905	12,377	14,388	13,983	15,380	59,918	56,070	56,129
YoY growth(%)	12.4%	15.5%	11.4%	1.9%	0.1%	-18.5%	-8.7%	-3.1%	-7.2%	1.1%	5.5%	10.6%	10.8%	-6.4%	0.1%
매출 대비 비중(%)	29.4%	33.8%	29.4%	27.7%	26.6%	27.0%	25.1%	24.7%	23.0%	24.8%	23.8%	24.0%	30.2%	26.3%	23.9%
Gross Profit	31,671	34,768	33,745	35,445	34,670	35,259	36,729	39,394	37,070	38,370	39,348	42,991	135,629	146,052	157,779
총이익률(%)	69.9%	67.2%	68.4%	68.3%	69.2%	66.8%	69.5%	70.1%	68.8%	66.0%	67.1%	67.0%	68.4%	68.4%	67.2%
영업비용	11,433	12,521	13,381	14,902	13,152	14,860	14,377	15,140	13,939	14,038	14,696	16,345	52,237	57,529	59,018
매출 대비 비중(%)	25.2%	24.2%	27.1%	28.7%	26.2%	28.2%	27.2%	26.9%	25.9%	24.2%	25.1%	25.5%	26.3%	27.0%	25.1%
R&D	5,599	5,758	6,306	6,849	6,628	6,844	6,984	6,739	6,166	6,291	6,493	7,106	24,512	27,195	26,055
매출 대비 비중(%)	12.4%	11.1%	12.8%	13.2%	13.2%	13.0%	13.2%	12.0%	11.4%	10.8%	11.1%	11.1%	12.4%	12.7%	11.1%
세일즈/마케팅	4,547	5,379	5,595	6,304	5,126	5,679	5,750	6,204	6,222	6,033	6,461	7,074	21,825	22,759	25,789
매출 대비 비중(%)	10.0%	10.4%	11.3%	12.2%	10.2%	10.8%	10.9%	11.0%	11.5%	10.4%	11.0%	11.0%	11.0%	10.7%	11.0%
G&A	1,287	1,384	1,480	1,749	1,398	2,337	1,643	2,197	1,552	1,714	1,743	2,165	5,900	7,575	7,173
매출 대비 비중(%)	2.8%	2.7%	3.0%	3.4%	2.8%	4.4%	3.1%	3.5%	2.9%	3.0%	3.0%	3.4%	3.0%	3.5%	3.1%
D&A	3,212	3,496	3,773	3,979	2,790	3,648	3,549	3,874	3,153	2,939	3,104	4,042	14,460	13,861	13,238
매출 대비 비중(%)	7.1%	6.8%	7.6%	7.7%	5.6%	6.9%	6.7%	6.9%	5.9%	5.1%	5.3%	6.3%	7.3%	6.5%	5.6%
영업이익	20,238	22,247	20,364	20,543	21,518	21,570	22,352	24,254	23,131	24,332	24,651	26,647	83,392	89,694	98,761
영업이익률(%)	44.7%	43.0%	41.3%	39.6%	42.9%	40.9%	42.3%	43.2%	42.9%	41.9%	42.0%	41.5%	42.1%	42.0%	42.1%
세전이익	20,524	22,515	20,190	20,496	21,572	21,510	22,673	24,727	23,454	24,681	25,003	27,031	83,725	90,482	100,170
순이익	20,505	18,765	16,728	16,749	17,556	17,596	18,299	20,081	19,414	20,323	20,604	22,221	72,747	73,532	82,562
순이익률(%)	45.2%	36.3%	33.9%	32.3%	35.0%	33.4%	34.6%	35.7%	36.0%	35.0%	35.1%	34.6%	36.7%	34.5%	35.2%
Non-GAAP EPS	2.71	2.48	2.22	2.23	2.35	2.35	2.45	2.69	2.60	2.72	2.76	2.98	9.69	9.85	11.06
YoY growth(%)	49.0%	22.3%	13.7%	2.8%	-13.4%	-5.2%	10.4%	20.5%	10.9%	15.6%	12.6%	10.7%	21.6%	1.6%	12.3%

자료: 마이크로소프트, 미래에셋증권 리서치센터

마이크로소프트 (MSFT US)

예상 포괄손익계산서 (요약)

(백만USD)	2021	2022	2023	2024F
매출액	168,088	198,279	211,915	234,772
매출원가	52,232	62,650	65,863	76,993
매출총이익	115,856	135,629	146,052	157,779
판매비와관리비	5,107	5,900	7,575	7,173
영업이익	69,916	83,392	89,694	98,761
비영업손익	-	-	-	-
금융손익	-	-	-	-
기타이익추가	1186	333	788	1409
법인세차감전순이익	71,102	83,725	90,482	100,170
법인세	10,451	10,978	16,950	17,608
당기순이익	60,651	72,747	73,532	82,562
지배주주		72,747	73,532	82,562
비지배주주	0	0	0	0

Growth & Margins (%)	2021	2022	2023	2024F
매출액증가율	17.5%	18.0%	7.6%	10.0%
매출총이익증가율	19.5%	17.1%	7.7%	8.0%
영업이익증가율	32.0%	19.3%	7.6%	10.1%
순이익증가율	37.0%	19.9%	1.1%	12.3%
EPS증가율	38.3%	21.6%	1.6%	12.3%
매출총이익률	68.9%	68.4%	68.4%	0.0%
영업이익률	41.6%	42.1%	42.0%	42.1%
당기순이익률	36.1%	36.7%	34.5%	35.2%

예상 현금흐름표 (요약)

(백만USD)	2021	2022	2023	2024F
영업활동 현금흐름	76,740	89,044	88,753	89,488
당기순이익	60,651	72,747	73,532	82,562
감가상각비	6,118	7,502	9,611	7,854
기타	11,686	14,460	13,861	13,238
투자활동 현금흐름	-2,335	-5,665	-8,251	-14,166
- 자본적 지출(CAPEX)	-27,577	-30,311	-22,680	-30,448
기타	-20,622	-23,886	-28,107	-27,236
재무활동 현금흐름	-6,955	39,566	-16,197	-25,000
배당금	-48,486	-58,876	-43,935	-55,471
자본 증가	-16521	-18135	-19800	-21029
장단기금융부채의 증가(감소)	1693	1841	1866	1304
기타	-33,658	-38,644	-32,916	-39,818
현금의 증감	648	-284	21,944	3,428
기초현금	13,576	14,224	13,940	35,884
기말현금	14,224	13,940	35,884	39,312

자료: 마이크로소프트, 미래에셋증권 리서치센터

예상 재무상태표 (요약)

(백만USD)	2021	2022	2023	2024F
유동자산	184,406	169,684	184,257	158,584
현금성자산	14,224	13,931	35,884	39,312
매출채권	38,043	44,261	48,688	40,419
재고자산	2,636	3,742	2,500	3,664
기타유동자산	129,503	107,750	98,365	96,479
비유동자산	149,373	195,156	227,719	241,911
투자자산	59,715	74,398	95,641	113,844
무형자산	89,658	120,758	132,078	128,067
기타비유동자산	333,779	364,840	411,976	400,495
자산총계	88,657	95,082	104,149	99,789
유동부채	15,163	19,000	18,095	17,848
매입채무	-	-	-	-
단기차입금	73,494	76,082	86,054	81,941
기타유동부채	103,134	103,216	101,604	83,171
비유동부채	50,074	47,032	41,990	41,965
장기금융부채	53,060	56,184	59,614	41,206
기타비유동부채	191,791	198,298	205,753	182,960
부채총계	141,988	166,542	206,223	139,134
지배주주지분(연결)	83,111	86,939	93,718	99,633
자본금	57,055	84,281	118,848	45,145
자본잉여금	1,822	-4,678	-6,343	-5,644
이익잉여금	-	-	-	1
기타	333,779	364,840	411,976	322,094
비지배주주지분(연결)	0	0	0	0
자본총계	1,023	1,183	1,420	1,665

예상 주당가치 및 valuation (요약)

	2021	2022	2023	2024F
P/E (배)	34.2	28.0	34.7	27.7
P/B (배)	14.3	11.5	9.6	7.8
EV/EBITDA (배)	23.8	18.8	19.4	19.4
EPS (USD)	8.1	9.7	9.7	12.7
BPS (USD)	42.2	36.7	29.6	44.7
DPS (USD)	2.2	2.5	2.7	3.1
배당성향 (%)	27.5	25.5	0.3	0.2
배당수익률 (%)	0.9	0.8	0.9	0.8
매출채권회전율 (회)	4.8	4.8	5.9	5.9
재고자산회전율 (회)	23.1	19.6	21.2	21.2
매입채무 회전율 (회)	3.8	3.7	4.1	4.1
ROA (%)	19.3	20.8	19.3	19.6
ROE (%)	47.1	47.2	33.9	32.3
ROIC (%)	33.6	36.6	32.3	32.3
부채비율 (배)	2.3	2.3	2.3	2.3
유동비율 (배)	2.1	1.8	1.8	1.8
순차입금/자기자본 (배)	-33.8	-15.8	-15.8	-15.8

(Analyst) 김수진 soojin.kim@miraeasset.com



ADBE US · 글로벌 소프트웨어 · 미국

어도비

생성시 수익화 첫 단추도 어도비가 채운다.

Not Rated

Refinitiv 평균목표주가

USD 604.07

상승여력

19.1%

현재주가(23/10/3)

USD 507.03

나스닥 종합(p)	13,059.47	시가총액(십억USD)	230.85	시가총액(조원)	310.45	유통주식수(백만주)	455.3
-----------	-----------	-------------	--------	----------	--------	------------	-------

Report summary

기존 제품 가격 인상 + 9월부터 Firefly 이미지 생성 월 4.99달러로 Monetize 본격화

생성시 수익화 계획을 구체화했다. Firefly 구독 요금제가 9월부터 실시된다. 월 4.99달러에 100크레딧을 받는다. 이 크레딧으로 Adobe Stock에 접근할 수 있고, 워터마크가 없는 Firefly 생성 이미지를 사용할 수 있게 된다. 또 미국과 일부 유럽 국가부터 요금 인상 계획도 발표했다. 가격 변경은 4Q부터 진행된다. Firefly의 베타 어플리케이션은 300만건 이상 다운로드 됐고, 20억개가 넘는 이미지가 생성됐다. Firefly는 포토샵, 일러스트레이터, 어도비 익스프레스 등에 통합됐다. 추후에는 API 형태로도 제공되어 더욱 본격적인 생성시 비즈니스를 확대할 것으로 보인다. 생성시 시대 새로운 모멘텀이 4Q부터 가시화되기 시작한다.

3Q 매출 48.9억달러(+10% YoY), EPS 4.09달러(+20% YoY)로 예상 상회했다. 다만 4Q23 가이던스가 예상 부합 수준에 그쳤다. 가격 인상 효과는 3년에 걸쳐 반영될 것이므로 전망치가 보수적이었다. 생성시 투자에도 영업이익률이 34.7%(+1%pt QoQ)로 개선된 점은 긍정적이다. 생성시 수혜 대표주로 추천한다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	42.2	45.0	34.5
상대주가	32.1	26.2	11.4

Earnings and valuation metrics

결산기 (11월)	2021	2022	2023F	2024F	2025F
매출액 (백만USD)	15785.0	15785.0	19379.1	21781.3	24495.2
영업이익 (백만USD)	5802.0	5802.0	8764.5	9769.1	10853.5
영업이익률 (%)	36.8	36.8	45.2	44.9	44.3
순이익 (백만USD)	4822.00	4822.00	5385.32	6119.37	7014.08
EPS (USD)	10.10	10.10	11.91	13.43	15.33
ROE (%)	34.4	34.4	41.7	35.6	32.5
P/E (배)	61.7	61.7	32.2	28.5	25.0
P/B (배)	19.8	19.8	13.9	10.9	8.7

주: GAAP 기준, 순이익은 지배주주 귀속 순이익
자료: 어도비, Refinitiv, 미래에셋증권 리서치센터

어도비의 생성AI 솔루션 - 1) Adobe Firefly

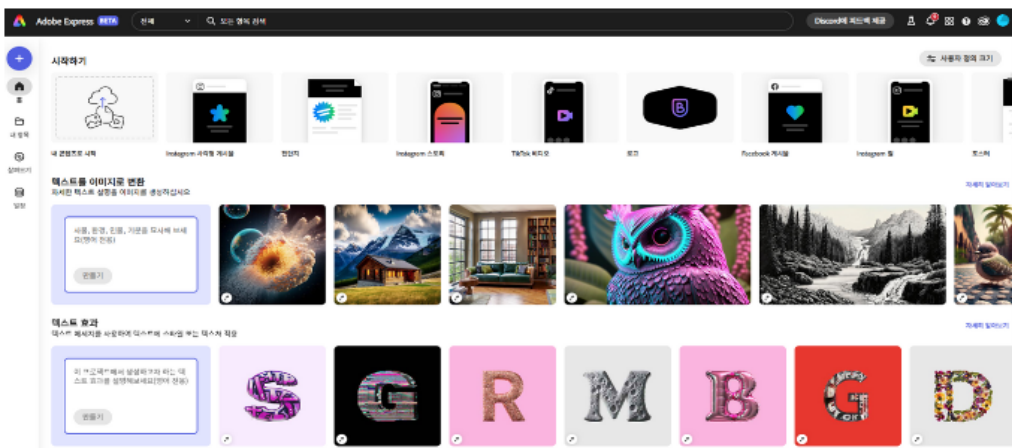
Adobe Firefly는 어도비가 출시한 생성형AI 기반의 웹 애플리케이션이다. 텍스트 프롬프트를 작성하면 원하는 이미지가 생성되고, 또 브러시/텍스처/벡터 등 구성요소에도 활용할 수 있다. 추후에는 이미지/영상/3D 등 다양한 미디어를 생성할 것이다.

현재는 1) 텍스트를 이미지로 생성, 2) 프롬프트로 개체 추가, 3) 텍스트 스타일이나 텍스처 생성, 4) 색 다시 칠하기, 5) 3D를 이미지로 생성, 6) 이미지 종횡비 변경 등의 기능을 제공한다. 이 기능들은 어도비의 대표 제품인 포토샵, 일러스트레이터, Express에 통합되어 사용할 수 있다.

3월 첫 출시 이후 베타버전이 무료로 제공되었으나, 9월부터 유료 요금제가 출시됐다. 월 4.99달러에 100 크레딧을 받고, 이 크레딧으로 Adobe Stock에 접근할 수 있고, 워터마크가 없는 Firefly 생성 이미지를 사용할 수 있게 된다. 무료로는 월 25개의 크레딧을 받게된다. Firefly는 공개 라이선스 작품, 저작권이 만료된 공개 도메인 콘텐츠, Adobe Stock의 데이터 세트를 학습한다.

어도비는 무엇보다 이미지 데이터의 신뢰도를 강조해 차별점을 두고자 한다. 어도비는 콘텐츠 자격증명(Content Credentials)이란 기능을 준비 중이다. 이는 이미지에 정보를 변조 방지 메타데이터 형태로 저장해 생성AI가 만든 이미지인지, 혹은 특정 개인의 이미지인지를 구분할 수 있게 하는 기능이다. 그리고 제작자는 이 메타데이터에 생성AI가 학습하기를 원하는지 아닌지 설정할 수도 있다.

그림 179. Adobe Express에 Firefly 탑재



자료: 어도비, 미래에셋증권 리서치센터

표 30. Firefly 생성 크레딧 요금표

기능	생성 크레딧 사용 요금
색 채우기, 이미지 확장, 텍스트로 이미지로, 다시 색칠하기	1 크레딧
텍스트 효과	23년 11월 1일 이전: 0 크레딧 23년 11월 1일부터: 1 크레딧

자료: 어도비, 미래에셋증권 리서치센터

주: 최대 2000 X 2000 픽셀 표준 이미지 기준이며, 향후 고품상도 이미지, 애니메이션, 비디오 및 3D 생성형 AI 기능을 제공할 계획인데, 이 기능들은 크레딧수가 더 클 수 있음

어도비의 생성AI 솔루션 - 2) GenStudio 출시

9월 19일에는 GenStudio를 새롭게 발표했다. 이는 콘텐츠 아이디어 구상, 제작, 생산 및 활성화를 위한 최고의 기능을 결합한 플랫폼 형태의 콘텐츠 공급망 솔루션이다. 이는 포토샵 등 Creative Cloud 제품들과 마케팅/광고를 위한 Experience Cloud 플랫폼이 모두 통합되어 생성형AI 기능들을 기반으로 콘텐츠를 제작하도록 돕는 역할을 한다.

그림 180. 어도비의 GenStudio



자료: 어도비, 미래에셋증권 리서치센터

그림 181. GenStudio는 Content Supply Chain으로 분류되어 출시 - 여러 사람이 함께 콘텐츠를 제작하는 과정을 돕는 플랫폼



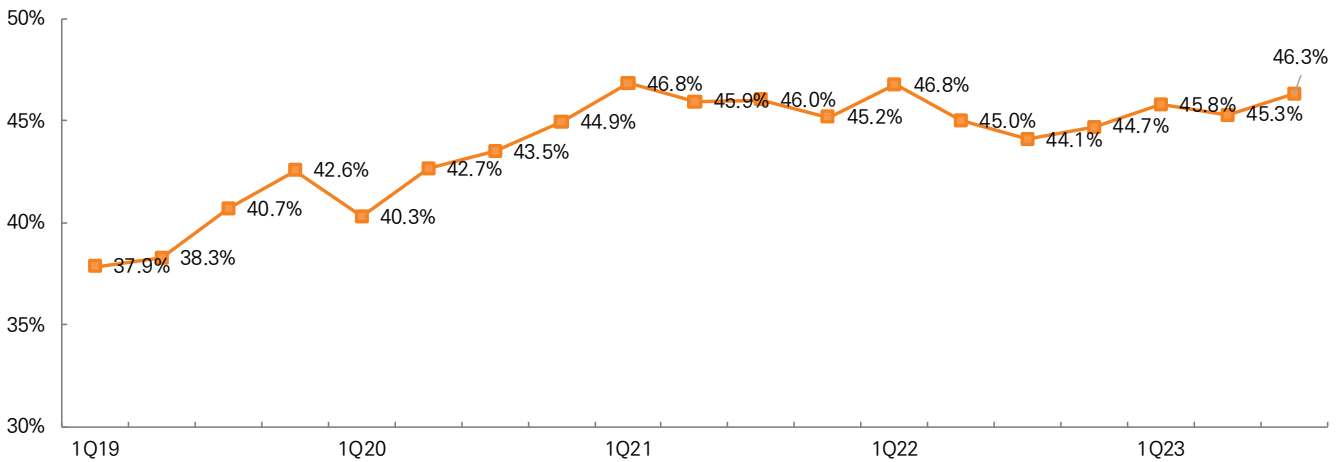
자료: 어도비, 미래에셋증권 리서치센터

표 31. 사업부문별 매출

(십억달러/%)	1Q21	2Q21	3Q21	4Q21	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23
총매출	3,905	3,835	3,935	4,110	4,262	4,386	4,433	4,525.0	4,655	4,816	4,890
Digital Media	2,859	2,787	2,865	3,009	3,110	3,200	3,232	3,300.0	3,395	3,511	3,594
Creative	2,379	2,320	2,372	2,477	2,548	2,605	2,625	2,681.0	2,761	2,852	2,909
Document Cloud	480	469	493	532	562	595	607	619.0	634	659	685
Digital Experience	934	938	985	1,010	1,057	1,095	1,120	1,150.0	1,176	1,222	1,229
Experience Cloud 구독	812	817	864	886	932	961	981	1,006.0	1,042	1,070	1,096
YoY 증가율											
Total Revenue	26.3%	22.5%	22.0%	20.0%	9.1%	14.4%	12.7%	10.1%	9.2%	9.8%	10.3%
Digital Media	31.8%	25.0%	22.6%	20.6%	8.8%	14.8%	12.8%	9.7%	9.2%	9.7%	11.2%
Creative	30.9%	24.1%	20.9%	19.1%	7.1%	12.3%	10.7%	8.2%	8.4%	9.5%	10.8%
Document Cloud	36.8%	30.3%	31.5%	29.4%	17.1%	26.9%	23.1%	16.4%	12.8%	10.8%	12.9%
Digital Experience	8.9%	13.6%	17.5%	8.7%	13.2%	16.7%	13.7%	13.9%	11.3%	11.6%	9.7%
Experience Cloud 구독	9.9%	15.6%	18.5%	8.2%	14.8%	17.6%	13.5%	13.5%	11.8%	11.3%	11.7%
매출 대비 비중											
Digital Media	73.2%	72.7%	72.8%	73.2%	73.0%	73.0%	72.9%	72.9%	72.9%	72.9%	73.5%
Creative	60.9%	60.5%	60.3%	60.3%	59.8%	59.4%	59.2%	59.2%	59.3%	59.2%	59.5%
Document Cloud	12.3%	12.2%	12.5%	12.9%	13.2%	13.6%	13.7%	13.7%	13.6%	13.7%	14.0%
Digital Experience	23.9%	24.5%	25.0%	24.6%	24.8%	25.0%	25.3%	25.4%	25.3%	25.4%	25.1%
Experience Cloud 구독	20.8%	21.3%	22.0%	21.6%	21.9%	21.9%	22.1%	22.2%	22.4%	22.2%	22.4%

자료: 어도비, 미래에셋증권 리서치센터

그림 182. 영업이익률 추이



자료: 어도비, 미래에셋증권 리서치센터

(Analyst) 임희석 heeseok.lim@miraeeasset.com



 미래에셋증권

U US · 글로벌 게임 · 미국

유니티

생성 AI 도입이 가져올 변화

Not Rated

Refinitiv 평균목표주가

USD 44.07

상승여력

50.3%

현재주가(23/10/3)

USD 29.32

다우존스 산업평균(p)	4,229.45	시가총액(십억USD)	11.24	시가총액(조원)	15.12	유통주식수(백만주)	261.20
--------------	----------	-------------	-------	----------	-------	------------	--------

Report summary

게임 엔진사에 대한 확실한 수혜 예상

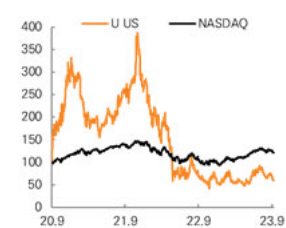
생성 AI 도입은 2010년대 초중반 앱스토어/플레이스토어 보편화에 따른 콘텐츠 공급 증가기와 비슷한 양상을 보여줄 것이다. 기존 레거시 게임사들은 새로운 경쟁자들의 출현으로 큰 위기를 겪기도 했으나 과점 플랫폼과 엔진사들은 콘텐츠 급증의 수혜를 온전히 누릴 수 있었다. 생성 AI를 통한 게임 공급 증가 초입에서는 모바일 및 캐주얼 장르 게임의 급증이 발생할 것이며 그 수혜를 동사가 누리게 될 것이다.

중장기적 관점에서 큰 그림을 그려볼 시기

23F P/S 5.5배로 상장 이후 밸류에이션 최하단에서 거래 중이다. 반등을 위해서는 매출액 및 고객수 성장세 회복이 선행될 필요가 있다. 논란이 지속되고 있는 신규 요금제의 경우 도입이 된다면 바텀라인의 본격적인 개선이 기대되나 지연될 가능성이 높다고 판단한다. 생성 AI 접목에 따른 게임 산업 내 콘텐츠 공급 급증과 그에 따른 동사의 수혜는 2H24부터 본격화될 것으로 전망한다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	-20.9	-8.9	-13.6
상대주가	-15.6	-11.1	-24.9

Earnings and valuation metrics

계산기 (12월)	2018	2019	2020	2021	2022
매출액 (백만USD)	381	542	772	1,111	1,391
영업이익 (백만USD)	-130	-151	-275	-532	-882
영업이익률 (%)	-34.1	-27.9	-35.6	-47.9	-63.4
순이익 (백만USD)	-132	-163	-282	-533	-921
EPS (USD)	-0.55	-0.68	-1.03	-1.89	-2.97
ROE (%)	-83.3	-46.0	-23.2	-24.0	-31.1
P/E (배)	-	-	-	-	-
P/B (배)	-	-	20.6	17.5	3.0

주: GAAP 기준, 순이익은 지배주주 귀속 순이익
 자료: 유니티, Refinitiv, 미래에셋증권 리서치센터

생성 AI 도입이 가져올 게임 산업 반전

생성 AI 도입으로 1) 프로그래밍 생산성 향상, 2) 1인, 소규모 개발 활성화, 3) 노코드 툴 도입에 따른 개발자 저변 확대, 4) 아트 리소스 재분배 효과가 이어져 콘텐츠 공급이 확대될 것이다. 1H23부터 게임사들의 생성 AI 테스트가 시작되고 있는 점을 고려할 때 2H24 또는 25년부터 생성 AI 도입 게임들의 출시가 본격화될 전망이다. 생성 AI의 게임 내 도입을 통한 재미의 향상은 유저 ARPPU 상승으로도 이어질 것이다. AI NPC의 등장으로 게임 몰입감이 증대될 것이며 유저가 함께 세계관을 만들어 나가는 게임도 등장할 것이다. 스테이지형 게임은 AI 기반 스테이지 형성으로 업데이트 기간이 단축되고 멀티플레이어 게임은 고도화된 AI 팀원으로 팀전 재미가 향상될 것이다.

게임 엔진사에 대한 확실한 수혜 예상

생성 AI의 도입은 2010년대 초중반 앱스토어/플레이스토어 보편화에 따른 콘텐츠 공급 증가기와 비슷한 양상을 보여줄 것이다. 기존 유통으로 사용되던 리소스가 개발 쪽으로 전환됨에 따라 생산성이 올라갔고 중소규모 개발진이 개발한 신규 콘텐츠가 급증했다. 기존 레거시 게임사들은 새로운 경쟁자들의 출현으로 큰 위기를 겪기도 했으나 과점 플랫폼과 엔진사들은 콘텐츠 급증의 수혜를 온전히 누릴 수 있었다.

게임 산업 회복의 수혜는 시장을 양분하고 있는 유니티와 언리얼엔진에 집중될 것이다. 게임 엔진은 각 플랫폼별로 언어가 다르기에 락인효과가 크며 생성 AI를 엔진사들 대부분이 도입한다고 해도 그 수준의 차이는 그간 쌓인 에셋의 양과 질에서 올 것이기 때문이다. 생성 AI 접목을 통한 게임 공급 증가기의 초입에서는 모바일 및 캐주얼 장르 게임의 급증이 먼저 올 것으로 판단하며 그 수혜를 동사가 누리게 될 것이다.

중장기적 관점에서 큰 그림을 그려볼 시기

23F P/S 5.5배로 상장 이후 밸류에이션 최하단에서 거래 중이다. 반등을 위해서는 매출액 및 고객수 성장세 회복이 선행될 필요가 있다. 논란이 지속되고 있는 신규 요금제의 경우 도입이 된다면 바텀라인의 본격적인 개선이 기대되나 지연될 가능성이 높다고 판단한다. 생성 AI 접목에 따른 게임 산업 내 콘텐츠 공급 급증과 그에 따른 동사의 수혜는 2H24부터 본격화될 것으로 전망한다.

1년 이상의 기간을 두고 바라보는 이유는 게임사와 플랫폼 모두 생성 AI 도입에 있어 시간이 필요하기 때문이다. 게임사는 프로그래밍, 아트 등 개발 전반에 대한 AI 도입을 개발자에게 설득하고, 툴을 최적화하고, 파일럿 테스트할 시간이 필요하다. 플랫폼 역시 생성 AI 기반 게임을 온보딩 시키기 위해서는 온보딩에 대한 유저들의 찬성과 생성 AI 기반 에셋 사용에 대한 RS 기준 확립 등이 선행될 필요가 있다.

생성 AI 게임 내 적용 본격화

NPC 적용

생성 AI의 게임 내 NPC 적용이 본격화되고 있다. 6월 중국에서 출시된 역수한(넷이즈 개발)은 생성 AI를 NPC에 전격적으로 도입했다. 생성 AI가 도입된 지능형 NPC의 경우 고유의 성격, 행동 논리, 기억력을 갖고 있으며 플레이어의 행동과 환경 변화에 능동적으로 대응한다. 캐릭터 커스터마이징에도 생성 AI가 도입되었는데 플레이어가 원하는 외형을 문장으로 입력하면 묘사에 부합하는 캐릭터의 얼굴이 생성된다. 이 외에도 플레이어는 게임을 탐험하면서 생성 AI를 활용해 시를 작성할 수도 있다.

생성 AI의 도입으로 게임 몰입감을 상승시킨 역수한의 돌풍은 매출로도 이미 증명되고 있다. 6월 출시 후 2~3위 수준의 매출 최상위권을 유지하고 있으며 iOS 매출로만 일 40억원 이상이 발생하고 있다.

그림 183. 생성 AI를 통해 텍스트 기반 캐릭터 커스터마이징도 가능한 역수한



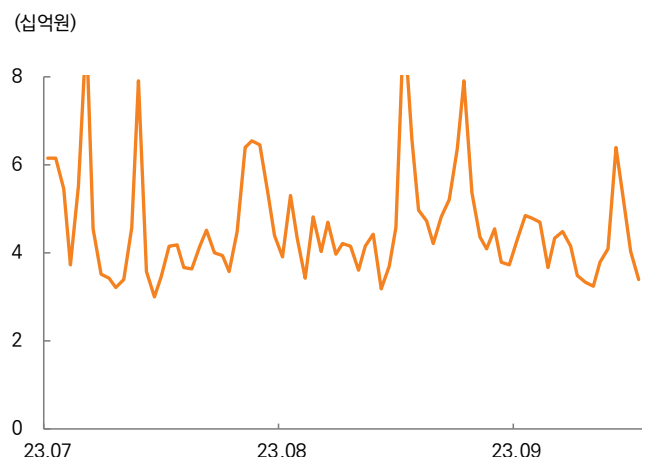
자료: 역수한, 미래에셋증권 리서치센터

그림 184. 왕자영요에 이어 iOS 매출 2위를 유지 중인 역수한



자료: 센서타워, 미래에셋증권 리서치센터

그림 185. 역수한 iOS 매출 추이: 40억원 전후 일매출 유지

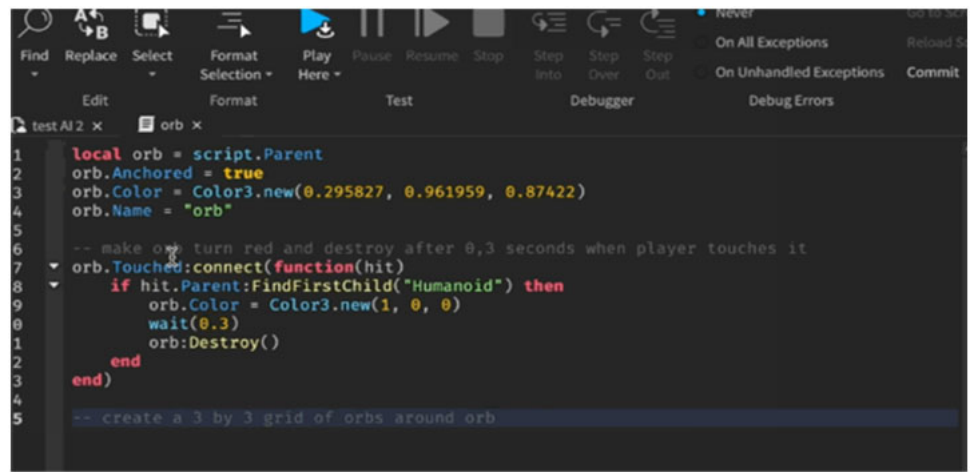


자료: 미래에셋증권 리서치센터

생산성 증대에 따른 콘텐츠 공급 증가

생성 AI의 적용은 게임 생산성을 극대화해 코로나 이후 둔화된 게임 업종 내 콘텐츠 공급을 가속화할 전망이다. 게임은 기술집약적 산업임과 동시에 스토리, 그래픽, 사운드, 프로그래밍 등 작업 전반을 사람이 수작업해야 하는 노동집약적 산업이었다. 앞으로는 생성 AI가 스토리를 생성하고, 그래픽과 사운드를 입히고, 프로그래밍을 보조하는 시대가 될 것이다. 개발자들이 리소스를 수급하는데 사용한 시간과 비용의 상당 부분이 콘텐츠 개발로 전환될 것이다.

그림 186. 로블록스 코드어시스트 예시



자료: 로블록스, 미래에셋증권 리서치센터

노코드 개발의 보편화와 개발자 저변 확대

생성 AI의 도입으로 노코드 개발 툴이 보편화되고 개발자 저변이 확대될 전망이다. 로블록스는 지난 3월 코드어시스트라는 생성형 AI 기반 게임 제작 툴을 공개했다. 코드어시스트는 원하는 것을 입력하면 코드를 자동으로 생성해준다. 전문적인 코딩 지식이 없는 사람도 텍스트 입력만을 통해서 아이템 제작이 가능하다.

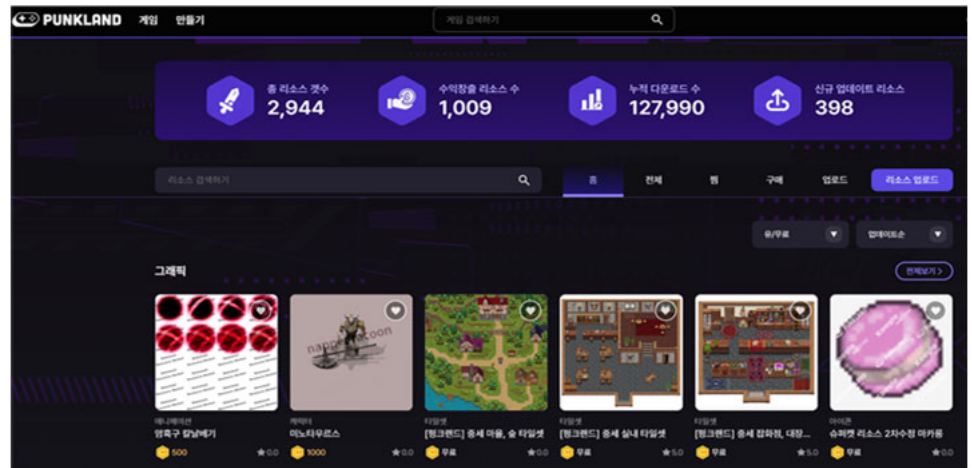
코딩에 대한 진입장벽 완화로 개발자 연령층도 확대될 전망이다. 노코드 개발 툴을 도입하고 있는 로블록스는 플레이어였던 이용자가 게임을 즐기고 성장하면서 자연스럽게 개발자가 되는 경우가 많다. 그러나 이용자의 43%가 13세 미만이지만 게임 제작 이용자는 주로 18세 이상이다. 생성 AI로 게임 개발 진입장벽이 낮아지면서 게임 개발자 연령층이 보다 넓어질 것으로 판단한다.

1인, 소규모 단위 게임 개발 증가 예상

생성 AI 기반의 리소스 보급은 개발자들이 리소스를 수급하는데 기존에 사용한 시간과 비용을 콘텐츠 개발에 사용할 수 있도록 만든다. 이는 아트 리소스 수급에 큰 부담을 가진 1인 또는 소규모 개발자들의 개발 편의성을 대폭 향상시켜 콘텐츠 공급 증가로 이어질 전망이다.

슈퍼캐트의 노코드 게임 개발 플랫폼 핑크랜드는 AI 이미지 생성기를 무료로 도입해 아트 리소스 수급을 즉각적으로 가능하게 만들었다. 1인 게임 개발자 '33게임즈'가 선보인 '닥사RPG: 파밍용사키우기'는 서비스 50일만에 매출 5억원을 달성하기도 했다. 게임 개발 난이도의 하락으로 신규 게임 개발자가 증가하고 플랫폼 내 리소스가 확대되어 게임 퀄리티 증가로 이어지는 선순환 구조가 지속될 것으로 기대된다.

그림 187. 슈퍼캐트의 노코드 게임 개발 플랫폼 핑크랜드



자료: 핑크랜드, 미래에셋증권 리서치센터

아트 비용의 감소

생성 AI의 도입으로 게임 개발 생산성이 좋아짐과 동시에 제작 비용이 드라마틱하게 감소할 것으로 전망한다. 트리플A급 또는 대작 게임을 만드는 데 있어서 들어가는 비용은 적게는 수백억에서 많게는 수천억원에 달한다. 그리고 이러한 게임 제작 비용에서 개발 비용과 함께 가장 많은 비중을 차지하는 영역이 아트 비용이다. 생성 AI 툴의 도입으로 2~3년 내 아트 비용이 급감하기 시작할 것으로 판단한다.

생성 AI 기반 이미지 활용이 본격화되면서 기존 일러스트 관련 비용이 급감할 것이다. 일러스트레이터가 수작업으로 그린 게임 일러스트 가격은 장당 수십만원을 호가하지만 생성 AI가 그린 일러스트 가격은 장당 수십원 수준에 불과하다. 물론 생성 AI 툴로 그린 그림에 대한 리터칭 필요성이 있으므로 비용은 일부 추가될 것이나 기존의 순수 수작업과 비교 시 10분의 1의 비용도 들지 않게 될 것이다.

그림 188. 미드저니 요금제

	Basic Plan	Standard Plan	Pro Plan	Mega Plan
Monthly Subscription Cost	\$10	\$30	\$60	\$120
Annual Subscription Cost	\$96 (\$8 / month)	\$288 (\$24 / month)	\$576 (\$48 / month)	\$1152 (\$96 / month)
Fast GPU Time	3.3 hr/month	15 hr/month	30 hr/month	60 hr/month
Relax GPU Time	-	Unlimited	Unlimited	Unlimited
Purchase Extra GPU Time	\$4/hr	\$4/hr	\$4/hr	\$4/hr
Work Solo In Your Direct Messages	✓	✓	✓	✓
Stealth Mode	-	-	✓	✓

자료: 미드저니, 미래에셋증권 리서치센터

21년 5월 오픈AI의 달리(DALL-E), 22년 7월 미드저니의 미드저니(Midjourney), 22년 8월 스테빌리티AI의 스테이블디퓨전(Stable Diffusion)이 출시되면서 이미지 생성 AI의 성능이 급격하게 올라오기 시작했다. 22년 9월에는 미드저니가 생성한 그림이 미국 주립 박람회 미술대회에서 우승하며 이미 AI 생성 이미지 수준이 사람 수준 이상으로 올라왔음을 증명하기도 했다.

그림 189. 스테이블디퓨전 예시1



자료: 스테이블디퓨전, 미래에셋증권 리서치센터

그림 190. 스테이블디퓨전 예시2



자료: 스테이블디퓨전, 미래에셋증권 리서치센터

웨이보 등 중국 SNS 플랫폼에서는 게임 업체들이 스테이블디퓨전, 미드저니를 도입하기 시작하면서 해고된 일러스트레이터들이 사례가 넘쳐나고 있다. 생성 AI 도입 이전에는 캐릭터 외주 비용이 8,000위안 수준이었지만 현재는 2,000위안 수준까지 감소한 것으로 파악된다. 현재는 게임회사에서 AI 시안을 먼저 제공하고 구체화 및 리터칭을 외주가 담당하는 일이 증가하고 있다. 작업에 1~2주 정도 걸린 정교한 일러스트도 생성 AI 도입 이후 3~4일이면 충분한 수준까지 올라왔다.

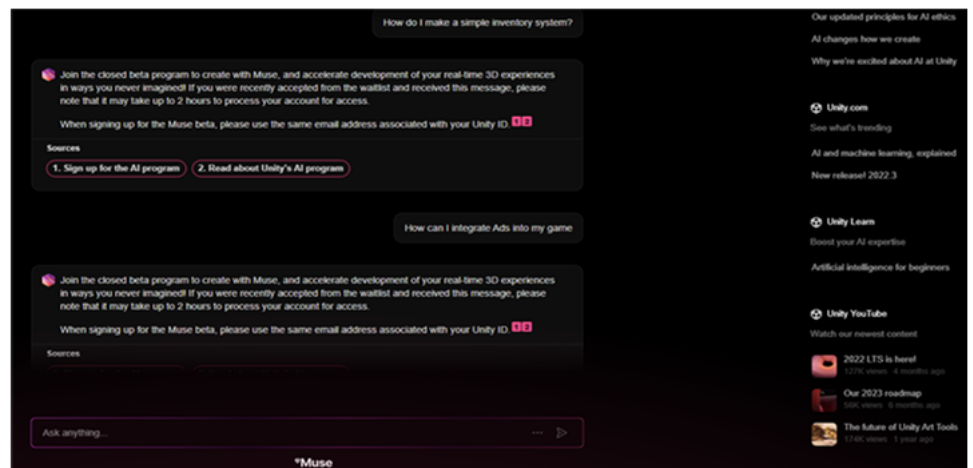
중국 게임사 XD네트워크(心动网络)의 회장 황이명에 따르면 게임 개발팀들에서 기존에 외주를 주던 일러스트와 번역 업무를 더이상 주지 않는 경우가 늘어나고 있다. 중국 대형 게임사들은 이미 초기 디자인과 세부 특성화 작업에 AI를 도입하고 있다. 넷이즈가 대표적인 경우인데 넷이즈는 현재 게임 제작 전반에 생성 AI를 적극적으로 활용하고 있다. 중소 게임사들도 더이상 외주 일러스트를 활용하지 않고 직접 AI툴을 이용하는 양상이 많아지고 있다.

게임 엔진사 수혜 예상

생성 AI 도입에 따른 게임 산업 내 콘텐츠 공급 확대로 게임 엔진사가 수혜를 볼 것이다. 생성 AI 도입으로 1) 프로그래밍 생산성 향상, 2) 1인, 소규모 개발 활성화, 3) 노코드 툴 도입에 따른 개발자 저변 확대, 4) 아트 리소스 재분배 효과가 이어져 콘텐츠 공급이 확대될 것이다. 1H23부터 게임사들의 생성 AI 테스트가 시작되고 있는 점을 고려할 때 2H24 또는 25년부터는 생성 AI 도입 게임들의 출시가 본격화될 전망이다.

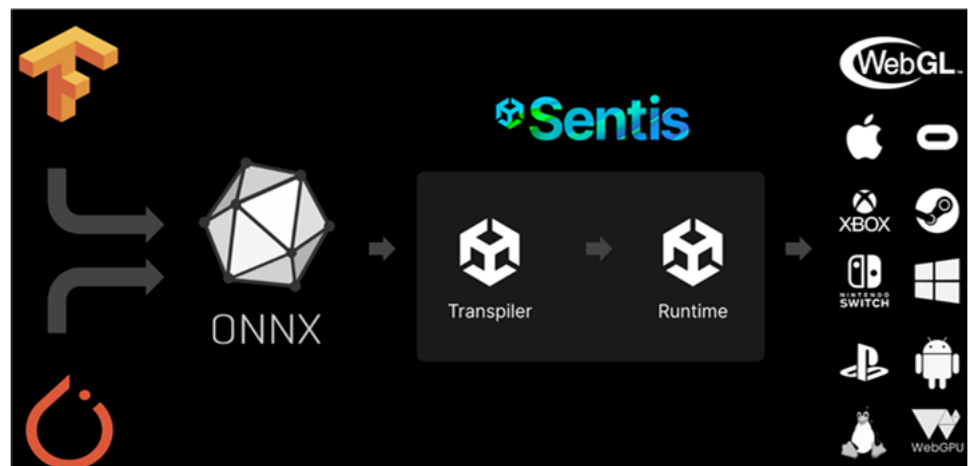
생성 AI 도입을 통한 게임 산업 회복에 따른 수혜는 시장을 10년 넘게 양분하고 있는 유니티와 언리얼엔진에 집중될 가능성이 높다. 게임 엔진은 각 플랫폼별로 언어가 다르기에 락인효과가 큰 편이며 생성 AI를 엔진사들 대부분이 도입한다고 해도 그 수준의 차이는 그간 쌓인 에셋의 양과 질에서 올 것이기 때문이다. 양사 모두 에셋의 양과 질 측면에서 타사들을 압도하고 있다.

그림 191. 유니티 신규 AI 툴 뮤즈



자료: 유니티, 미래에셋증권 리서치센터

그림 192. 유니티 신규 AI 툴 센티스



자료: 유니티, 미래에셋증권 리서치센터

표 32. 분기 및 연간 실적 전망

(백만USD)

	1Q21	2Q21	3Q21	4Q21	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	2019	2020	2021	2022
매출액	235	274	286	316	320	297	323	451	500	533	542	772	1,111	1,391
(% YoY)	41%	48%	43%	43%	36%	9%	13%	43%	56%	80%	42%	43%	44%	25%
Grow Solutions	164	201	203	216	204	176	194	253	313	340	373	541	784	716
(% YoY)	37%	56%	47%	41%	24%	-12%	-4%	17%	54%	93%	46%	45%	45%	-9%
Create Solutions	70	72	84	100	116	121	129	198	187	193	169	231	327	675
(% YoY)	51%	31%	34%	50%	65%	67%	54%	98%	61%	60%	34%	37%	41%	107%
고객수	837	888	973	1,052	1,083	1,085	1,075	1,340	1,322	1,330	600	793	1,052	1,340
매출원가	59	58	64	74	94	97	112	140	162	159	119	172	254	443
(% of Sales)	25%	21%	22%	23%	29%	33%	35%	31%	32%	30%	22%	22%	23%	32%
영업비용	287	365	350	387	397	398	451	585	593	566	574	875	1,389	1,831
(% of Sales)	122%	133%	122%	123%	124%	134%	140%	130%	119%	106%	106%	113%	125%	132%
R&D	154	154	178	209	221	216	248	274	280	268	256	404	696	959
마케팅	70	75	97	103	104	101	110	183	216	209	174	216	345	498
일반관리비	63	136	74	75	72	81	93	127	97	89	144	255	348	373
영업이익	-111	-149	-127	-145	-171	-198	-240	-274	-255	-191	-151	-275	-532	-882
영업이익률	-47%	-55%	-44%	-46%	-53%	-67%	-74%	-61%	-51%	-36%	-28%	-36%	-48%	-63%
지배주주순이익	-107	-148	-115	-162	-178	-204	-250	-289	-253	-192	-163	-282	-533	-921
순이익률	-46%	-54%	-40%	-51%	-55%	-69%	-77%	-64%	-51%	-36%	-30%	-37%	-48%	-66%

자료: 유니티, 미래에셋증권 리서치센터

표 33. 연간 매출액, P/S 추이

	2020	2021	2022	2023F	2024F
매출액(백만USD)	772	1,111	1,391	2,176	2,620
P/S(배)	41.0	31.1	12.7	5.5	4.6

주: 과거 멀티플 수치는 평균값으로 자체 추정
 자료: 블룸버그, 미래에셋증권 리서치센터 추정



(Analyst) 박연주 yeonju.park@miraeasset.com



 미래에셋증권

TSLA US · 전기차/자율주행 · 미국

테슬라

도조의 의미와 옵티머스의 잠재력

(유지)

매수

목표주가

USD 340

상승여력

37.9%

현재주가(23/10/03)

USD 246.53

나스닥 종합(p)	13,059.47	시가총액(십억USD)	782.48	시가총액(조원)	1,066.07	유통주식수(백만주)	2,760.14
-----------	-----------	-------------	--------	----------	----------	------------	----------

Report summary

도조의 의미와 옵티머스의 잠재력

GPU 공급 부족이 심각한 상황에서 테슬라는 자체 개발한 슈퍼 컴퓨터 도조의 가동을 시작했다. 독보적인 실 주행 데이터, End to End 신경망을 쓸 수 있는 기술력과 함께 연산 능력이 24년 10월 100억사플롭(23년 초 대비 30배 이상 증가)까지 확장된다면 24년 중 레벨 2+ 도심 자율주행의 기술 수준이 소비자들의 옵션 채택률이 높아질 정도로 충분히 개선될 가능성이 높아 보인다. 또한 최근 공개된 옵티머스 영상에서 확인하듯 대형언어모델은 로봇의 인지/판단 능력을 획기적으로 개선시킬 수 있어 해당 분야 잠재력도 크다.

AI 상용화에 따른 기업 가치 상승 잠재력이 가장 큰 기업, 단기 경기 우려에도 Top Pick 유지

테슬라는 단기적으로 3분기 판매량 둔화, 4분기 금리 급등에 따른 자동차 수요 둔화 가능성 등으로 주가 변동성이 커질 수 있다. 그러나 경기 둔화는 모든 업체에 영향을 미치는 반면 테슬라의 핵심 경쟁력은 24년으로 갈수록 강해질 전망이다. AI 기술의 상용화를 통해 중장기 기업 가치가 확대될 잠재력이 높다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	0.6	26.6	1.7
상대주가	7.4	23.4	-11.5

Earnings and valuation metrics

계산기 (12월)	2020	2021	2022	2023F	2024F
매출액 (백만USD)	31.5	53.8	81.5	98.1	129.1
영업이익 (백만USD)	2.0	6.5	13.7	11.0	19.9
영업이익률 (%)	6.3	12.1	16.8	11.2	15.4
순이익 (백만USD)	0.7	5.5	12.6	10.2	15.3
EPS (USD)	0.2	1.6	3.6	2.9	4.4
ROE (%)	3.2	18.3	28.1	18.5	21.7
P/E (배)	972.9	217.4	33.7	83.1	55.6
P/B (배)	30.5	36.2	8.7	15.4	12.0

주: GAAP 기준, 순이익은 지배주주 귀속 순이익
 자료: 테슬라, Refinitiv, 미래에셋증권 리서치센터

GPU 공급 부족 시대 도조의 의미: 경쟁 업체와의 격차 확대

생성 AI 어플리케이션으로 GPU 공급 부족이 심각한 상황에서 테슬라는 자체 개발한 슈퍼 컴퓨터 도조의 가동을 시작했다. 도조는 기존 GPU 대비 원가가 낮기도 하지만 자율주행에 특화된 전용 모델이고 확장이 용이해 연산 능력이 빠르게 증가할 수 있는 점에서 의미가 크다.

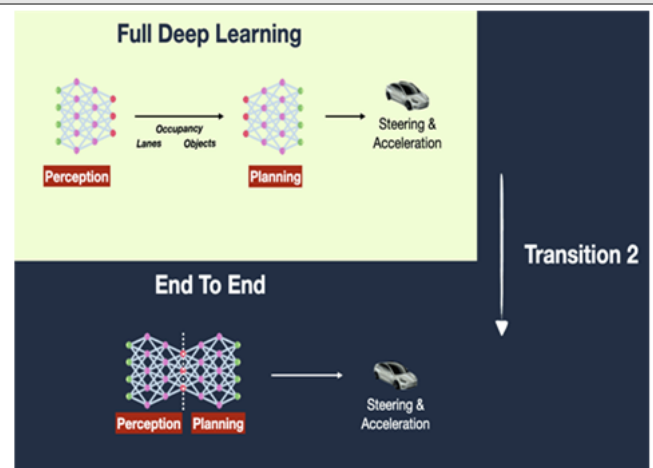
이미 독보적인 실 주행 데이터, End to End 신경망을 쓸 수 있는 기술력과 함께 연산 능력이 24년 10월 100엑사플롭(23년 초 대비 30배 이상 증가)까지 확장된다면 24년 중 레벨 2+ 도심 자율주행의 기술 수준이 소비자들의 옵션 채택률이 높아질 정도로 충분히 개선될 가능성이 높아 보인다. 더 중요한 것은 향후 1년 내에 경쟁 업체가 이 같은 기술과 연산 능력을 확보하는 것은 매우 어려울 것이라는 점이다. 이는 다른 자동차 업체들이 테슬라의 자율주행 기술을 라이선스할 가능성이 높다는 것을 의미한다.

AI 기술의 가장 큰 난제는 Long Tail, 즉 흔치 않은 상황에 대한 대응력을 키우는 것인데 기술 완성도를 높이고자 할수록 필요한 학습 리소스는 기하 급수적으로 늘어난다. 테슬라는 이미 200만대 이상의 자동차를 통해 실 도로 주행 데이터를 빠르게 확보하고 있다. AI 기술의 발전 속도도 빠르다. FSD V12부터 End to End 인공신경망을 도입하는데 현재 시스템은 주행 환경을 인식하는 인지 분야만 인공신경망을 주로 사용하고 판단 및 제어 분야는 사람이 프로그래밍하는 비중이 큰데 End to End 방식으로 가면 인지부터 판단 제어까지 하나의 인공신경망을 통해 이루어지기 때문에 사람의 개입이 줄고 기술의 발전 속도가 획기적으로 빨라질 수 있다. 이미 연초 이후 테슬라의 도심 자율주행 기술은 빠르게 발전 중인데 여기에 도조 도입으로 연산 능력이 24년 10월 100ExaFlops까지 늘어난다면 24년 중에는 레벨 2+ 도심 자율주행의 기술이 소비자들이 돈을 내고 살 수준까지 올라올 가능성이 높아 보인다.

그림 193. 테슬라의 도조 연산 능력 전망



그림 194. FSD V12의 End to End 모델 개념도



자료: 테슬라, 미래에셋증권 리서치센터

자료: autonomous.ai, 미래에셋증권 리서치센터

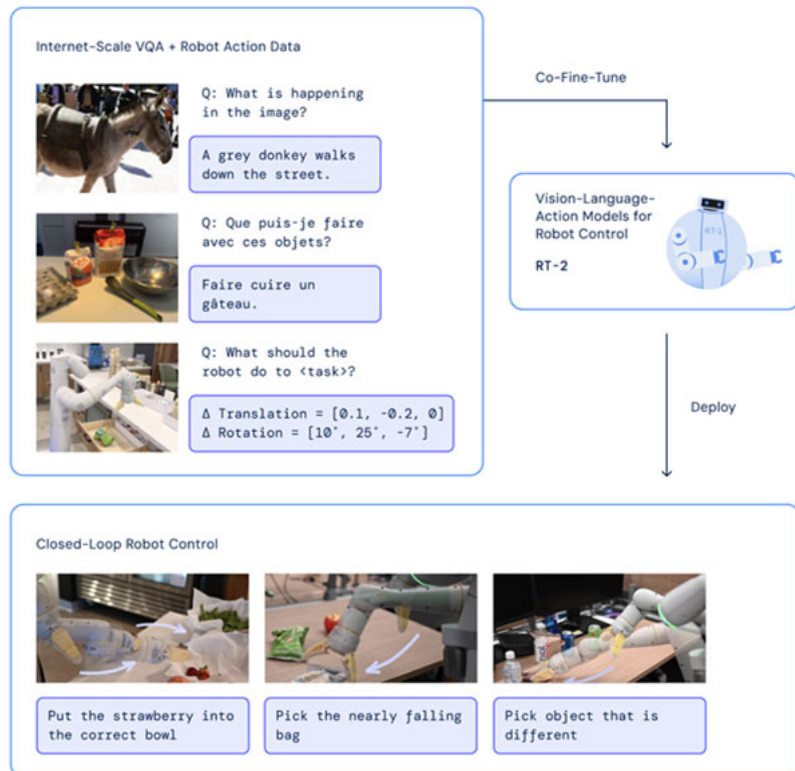
언어 모델과 오퍼머스의 잠재력

23년 9월 테슬라는 AI 로봇 오퍼머스 시연 영상을 공개하였다. 22년 AI Day에서 제대로 걷지도 못했던 오퍼머스는 이번 영상에서 사람의 방해를 피해 물건을 옮기거나 잘못 놓아진 물건을 바로 세우기도 하는 등 다양한 동작들을 선보였다. 테슬라는 이러한 로봇의 행동들이 end to end 방식, 즉 시각을 통해 상황을 인지하고 적절한 행동을 취하는 과정이 로봇에 내재된 하나의 인공지능경망을 통해 이루어지고 있다고 설명하였다.

기존의 로봇은 사전에 프로그래밍된 대로만 움직일 수 있기 때문에 수 많은 비정형적인 일들이 시시각각으로 일어나는 현실 세계에서 할 수 있는 일이 제한적이었다. 그러나 대형 언어 모델은 인터넷 등에 있는 수 많은 데이터를 학습해 로봇의 상식과 추론 능력을 개선시킬 수 있고 사용자들과 컴퓨터 언어가 아닌 자연어로 소통할 수 있다.

23년 7월 구글에서 발표한 'Robotic Transformer 2'라는 모델을 참고해볼 필요가 있다. 'RT 2'는 대형 시각 언어 모델(텍스트뿐 아니라 이미지 인식 기능을 갖춘 모델)을 이용해 로봇을 직접 제어하는 모델이다. 과거의 로봇은 "쓰레기를 치우라"는 명령을 수행하기 위해 세상에 존재하는 수 많은 쓰레기를 식별하기 위한 훈련을 해야 했다. 그러나 RT 2는 인터넷 상의 많은 데이터로 학습했기 때문에 따로 학습하지 않아도 쓰레기가 무엇인지 알고 상식과 추론에 근거해 적절한 해결 방안을 찾아 수행할 수 있다.

그림 195. 23년 7월 구글에서 발표한 RT-2: 대형 시각-언어 모델과 로봇의 결합



자료: 구글, 미래에셋증권 리서치센터

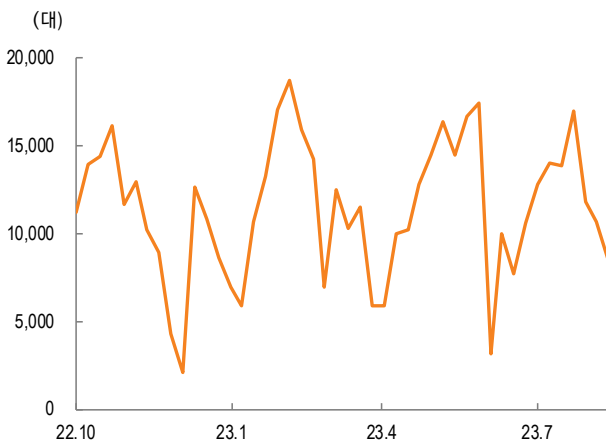
단기 경기 우려와 AI의 수익화 잠재력

테슬라 주가는 3분기 판매량이 예상보다 다소 부진했던 가운데 금리 급등에 따른 자동차 수요 둔화 우려, 중국 경기 부진 등으로 단기적으로 변동성이 확대될 수 있다. 특히 높아진 오토론 금리가 자동차 수요에 어느 정도 영향을 줄 지는 4분기 지켜볼 필요가 있다. 다만 경기 둔화에 따른 리스크는 모든 업체에 해당하는 것으로 테슬라의 핵심 경쟁력은 24년으로 갈수록 더 강해질 전망이다.

테슬라는 AI 기술을 상용화해서 기업 가치의 상승 폭이 가장 클 수 있는 글로벌 기업 중 하나이다. 자율주행 기술은 연산 능력 및 데이터 측면에서 경쟁 업체와의 격차가 계속 벌어지고 있다. 연초 이후 FSD의 도심 주행 능력은 빠르게 개선되고 있으며 24년 End to End 신경망의 도입과 연산 능력의 급격한 발전이 함께 한다면 레벨 2+ 도심 자율주행의 편의성이 충분히 높아져 본격적으로 수익에 기여할 가능성이 높아 보인다.

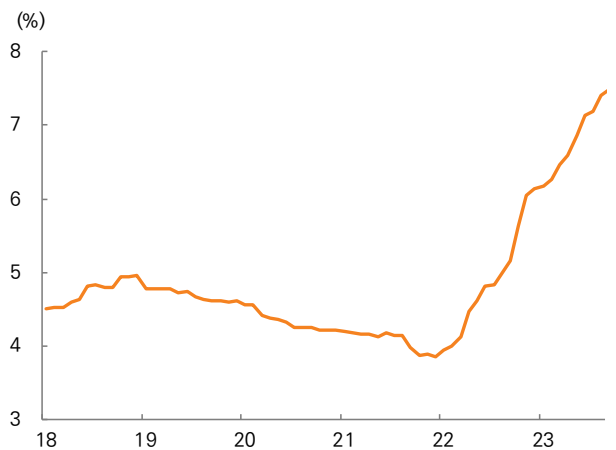
중기적으로 로보틱스 분야의 잠재력도 크다. 이번 시연에서 보여주듯 언어 모델 등 AI 기술 발전에 따라 로봇의 인지 및 판단 능력이 크게 개선될 수 있기 때문이다. 로보틱스 분야에서 테슬라의 자율주행 기술, 도조의 연산 능력, 제조 경쟁력 등은 테슬라의 경쟁 우위로 작용할 전망이다. 아직 테슬라 기업 가치에 로보틱스 가치가 포함되어 있지 않지만 향후 AI 데이 등을 통해 구체화되어 나갈 것으로 예상된다.

그림 196. 중국 테슬라 보험 등록 대수 추이



자료: Cnevpost, 미래에셋증권 리서치센터

그림 197. 미국 오토론 금리 추이



자료: Bloomberg, 미래에셋증권 리서치센터

테슬라 (TSLA US)

예상 포괄손익계산서 (요약)

(백만USD)	2021	2022	2023F	2024F
매출액	53,823	81,462	98,051	129,119
매출원가	40,217	60,609	78,871	99,763
매출총이익	13,606	20,853	19,180	29,356
영업비용	7,083	7,197	8,181	9,408
영업이익	6,523	13,656	10,999	19,948
비영업손익	-180	63	346	461
세전사업손익	6,343	13,719	11,345	20,409
법인세비용	699	1,132	567	4,286
비배주주순이익	125	31	539	806
지배주주순이익	5,519	12,556	10,239	15,317

Growth (YoY)	2021	2022	2023F	2024F
매출액증가율	70.7%	51.4%	20.4%	31.7%
매출총이익증가율	105.2%	53.3%	-8.0%	53.1%
영업이익증가율	227.1%	109.4%	-19.5%	81.4%
지배주주순이익증가율	665.5%	127.5%	-18.5%	49.6%

Margins	2021	2022	2023F	2024F
매출총이익률	25.3%	25.6%	19.6%	22.7%
영업이익률	12.1%	16.8%	11.2%	15.4%
지배주주순이익률	10.3%	15.4%	10.4%	11.9%

예상 현금흐름표 (요약)

(백만USD)	2021	2022	2023F	2024F
영업활동 현금흐름	11,497	14,724	17,330	22,948
당기순이익(지배주주지분)	5,519	12,556	10,239	15,317
비현금수익비용가감(감가상각비)	2,911	3,747	4,710	5,783
운전자본변동	3,447	-1,691	2,380	1,849
기타영업활동	-380	112	0	0
투자활동 현금흐름	-7,868	-11,973	-7,150	-8,580
자본적지출	-7,710	-6,231	-7,150	-8,580
장기금융자산의 증가(감소)	0	0	0	0
기타투자활동	-158	-5,742	0	0
재무활동 현금흐름	-5,203	-3,527	-2,000	-1,000
장단기금융부채의 증가(감소)	-5,732	-3,866	-2,000	-1,000
자본의 증가(감소)	707	541	0	0
기타재무활동	-178	-202	0	0
FX rate effect	-183	-444	0	0
현금의 증가	-1,757	-1,220	8,180	13,368
기초현금	19,384	17,576	16,253	24,356
기말현금	17,576	16,253	24,356	37,724

주: TTB는 '흑자전환', RR은 '적자지속'
 자료: 테슬라, 미래에셋증권 리서치센터

예상 재무상태표 (요약)

(백만USD)	2021	2022	2023F	2024F
유동자산	27,100	40,917	48,156	68,198
현금 및 현금성 자산	17,707	22,185	24,356	37,724
매출채권 및 기타채권	1,913	2,952	3,553	4,679
재고자산	5,757	12,839	16,708	21,133
기타유동자산	1,723	2,941	3,540	4,662
비유동자산	35,031	41,421	50,775	54,378
유형자산	23,395	26,111	28,551	31,348
기타비유동자산	11,636	15,310	22,224	23,030
자산총계	62,131	82,338	98,931	122,576
유동부채	19,705	26,709	34,158	42,680
매입채무 및 기타채무	15,376	21,912	28,514	36,067
단기금융부채	1,957	1,987	1,987	1,987
기타유동부채	2,372	2,810	3,657	4,625
비유동부채	10,843	9,731	7,731	6,731
장기금융부채	6,916	3,761	1,761	761
기타비유동부채	3,927	5,970	5,970	5,970
부채총계	30,548	36,440	41,889	49,411
자본금	1	3	3	3
자본잉여금	29,803	32,177	32,177	32,177
이익잉여금	331	12,885	23,124	38,441
기타포괄손익누계액	54	-361	5	5
비지배주주지분	1,394	1,194	1,733	2,539
자본총계	31,583	45,898	57,042	73,165
부채 및 자본총계	62,131	82,338	98,931	122,576

예상 주당가치 및 valuation (요약)

	2021	2022	2023F	2024F
P/E (배)	217.4	33.7	83.1	55.6
P/B (배)	36.2	8.7	15.4	12.0
EV/EBITDA (배)	114.9	21.5	39.6	23.7
EPS (USD)	1.6	3.6	2.9	4.4
BPS (USD)	8.9	12.9	15.9	20.3
매출채권회전율 (회)	28.3	27.6	27.6	27.6
재고자산회전율 (회)	8.2	6.3	5.9	6.1
매입채무회전율 (회)	5.2	3.7	3.4	3.6
ROA (%)	9.7	30.5	20.7	25.0
ROE (%)	18.3	28.1	18.5	21.7
ROIC (%)	15.0	74.0	57.0	114.9
부채비율 (%)	96.7	79.4	73.4	67.5
유동비율 (%)	137.5	153.2	141.0	159.8
순부채비율 (%)	-28.0	-35.8	-36.1	-47.8



(Analyst) 임희석 heeseok.lim@miraeeasset.com

MIRAE ASSET
 미래에셋증권

035420 · 인터넷

NAVER

국가대표 AI 기업

(유지)

매수

목표주가

290,000원

상승여력

43.9%

현재주가(23/9/27)

201,500원

KOSPI	2,465.07	시가총액(십억원)	33,056	발행주식수(백만주)	164	외국인 보유비중(%)	47.1
-------	----------	-----------	--------	------------	-----	-------------	------

Report summary

광고 단가 상승 및 수수료 BM 기대

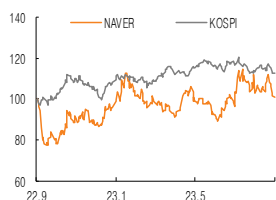
생성 AI 검색 도입에 따른 타겟팅 고도화로 광고 단가 상승으로 이어질 것으로 판단한다. 생성 AI 도입 이후 광고 단가가 YoY 20% 이상으로 성장세가 가속화될 것이다. 챗봇에 광고가 추가되기 시작하면 구독 외에도 중개 및 판매수수료 BM이 추가적으로 가능할 것이다. 개별 카테고리에서의 신규 광고 매출 외에도 구매까지 연결되었을 경우 Take rate를 수취하는 등의 중개 수수료가 신규로 발생할 것으로 예상된다.

1Q24부터 강력한 모멘텀 기대

동사를 업종 내 Top Pick으로 지속 유지한다. 24F 23배 수준으로 밸류에이션 부담이 여전히 낮으며 긍정적 지표 변화가 나타날 1Q24부터 다시금 AI 모멘텀이 발생할 것으로 판단한다. 11월 이벤트가 예고된 CUE와 CLOVA for AD가 광고 매출과 직결되는 부분이기 때문에 중요도가 가장 높다. 서비스 도입 후 발표할 것으로 기대되는 기존 검색 및 광고 대비 제로클릭률과 구매전환율이 얼마나 개선되었는지가 관전 포인트다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	-4.5	1.1	-1.5
상대주가	-2.4	-1.2	-11.1

Earnings and valuation metrics

계산기 (12월)	2021	2022	2023F	2024F	2025F
매출액 (십억원)	6,818	8,220	9,764	11,087	12,516
영업이익 (십억원)	1,325	1,305	1,513	1,822	2,392
영업이익률 (%)	19.4	15.9	15.5	16.4	19.1
순이익 (십억원)	16,490	760	955	1,441	1,881
EPS (원)	100,400	4,634	5,824	8,784	11,465
ROE (%)	106.7	3.3	4.1	6.0	7.3
P/E (배)	3.8	38.3	35.5	23.5	18.0
P/B (배)	2.5	1.2	1.4	1.3	1.2
배당수익률 (%)	0.1	0.5	0.4	0.4	0.4

주: K-IFRS 연결 기준, 순이익은 지배주주 귀속 순이익
 자료: NAVER, 미래에셋증권 리서치센터

광고 단가 상승 및 수수료 BM 기대

생성 AI 검색 도입에 따른 타겟팅 고도화로 광고 단가 상승으로 이어질 것으로 판단한다. CLOVA for AD의 경우 광고주는 생성AI를 원하는 방향으로 정보를 학습시켜 검색자와의 대화 중 적절한 맥락에서 광고를 제공할 수 있게 된다. 20년 GFA 도입 이후와 비슷한 효과가 발생할 것으로 예상된다. 생성 AI 광고 도입 이후 광고 단가가 YoY 20% 이상으로 성장세가 가속화될 것으로 전망한다.

챗봇에 광고가 추가되기 시작하면 구독 외에도 중개 및 판매수수료 BM이 추가적으로 가능할 것으로 판단한다. 챗봇은 서비스 및 상품 구매의 목적성이 강하게 나타나는 창구이기에 각 카테고리별 대형 브랜드와의 연결을 통한 시너지 효과가 클 것이다. 개별 카테고리에서의 신규 광고 매출 외에도 구매까지 연결되었을 경우 일정 수준의 Take rate를 수취하는 등의 중개 수수료가 신규로 발생할 전망이다.

커머스, 콘텐츠 영향도 긍정적

커머스에서는 AI 기반으로 이미지, 음성, 텍스트를 동시에 인식하는 멀티모달 서비스가 도입되면서 이용자가 원하는 상품을 보다 효율적으로 찾을 수 있게 변할 것이다. 생성 AI의 도입으로 쇼핑라이브 성장도 가속화될 것이다. 생성 AI를 통해 특정 품목에 관심이 있을 잠재 고객을 타겟팅 할 수 있게 되고, 대본 생성을 보조해 판매자 생산성이 향상되며, AI 휴먼 및 다국어 라이브 서비스도 가능해질 전망이다.

생성형 AI들의 도입은 양질의 웹툰/웹소설 콘텐츠 증가로 이어져 플랫폼 업체의 수혜로 이어질 것이다. 이미 웹툰/웹소설 시나리오를 생성하는 플랫폼들이 등장하고 있으며 클릭 한 번으로 AI가 채색을 완성시켜주는 툴도 고도화되고 있다. 다만 생성 AI 툴 도입에 대한 업계 내 기준이 확립될 필요가 있어 신규 콘텐츠 공급 확대에 따른 웹툰 산업 고성장세 회복은 25년 이후에 이어질 것으로 판단한다.

1Q24부터 강력한 모멘텀 기대

NAVER에 대해 투자 의견 '매수', 목표주가 290,000원, 업종 내 Top Pick으로 지속 유지한다. 24F 23배 수준으로 밸류에이션 부담이 여전히 낮으며 생성 AI 서비스 이후의 긍정적 지표 변화가 나타날 1Q24부터 다시금 AI 모멘텀이 발생할 것으로 판단한다. 동사는 CLOVA for Writing, 10월 CLOVA Studio, Neurocloud for HyperCLOVA X 오픈, 11월 통합 검색 내 CUE 적용, CLOVA for AD 선공개 등 다양한 서비스들을 공개 중이다. 11월에 이벤트가 예고된 CUE와 CLOVA for AD가 광고 매출과 직결되는 부분이기에 중요도가 가장 높다. 서비스 도입 후 발표할 것으로 기대되는 기존 검색 및 광고 대비 체로클릭률과 구매전환율이 얼마나 개선되었는지가 관전 포인트다.

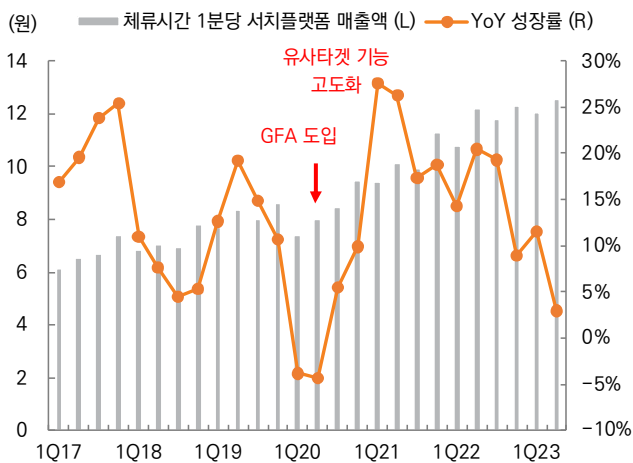
1. 광고

타겟팅 고도화에 따른 단가 상승

자연어 단위로 검색이 이루어지는 생성 AI 검색 도입은 광고 플랫폼 업체들의 타겟팅 능력을 고도화시킬 것이며 이는 곧 광고 단가 상승으로 이어질 전망이다. 질문 과정에서 노출되는 구체적인 조건들은 광고 플랫폼들이 검색자가 원하는 정보를 찾아주는 일의 난이도를 대폭 낮춰준다. 생성 AI 검색자는 서비스 또는 물품에 대한 구매의 목적성이 높기에 기존 대비 수배 이상 되는 구매 전환율을 보여주게 될 것이다. 빠른 목적 달성에 따른 검색 쿼리 및 체류시간(Q)이 일부 감소할 수 있으나 광고 단가(P) 상승폭이 이를 충분히 상쇄할 것으로 판단한다.

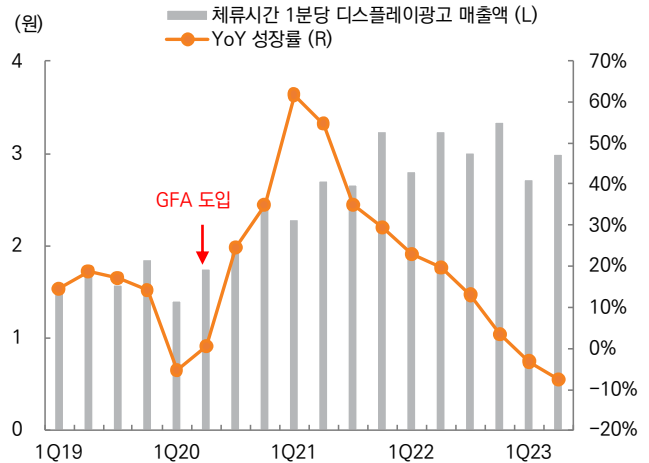
생성 AI 기반 광고를 통한 신규 매출은 1H24부터 본격화될 전망이다. 생성 AI 기반 광고 도입이 늦어지고 있는 글로벌과 달리 NAVER는 8월 CLOVA for AD를 일찌감치 공개하며 광고 상품의 대대적인 변화를 예고했다. CLOVA for AD에서 광고주는 생성 AI를 원하는 방향으로 정보를 학습시켜 검색자와의 대화 중 적절한 맥락에서 광고를 제공할 수 있게 된다. NAVER는 20년에도 성과형 디스플레이광고(GFA)를 도입하고 타겟 기능을 고도화 하면서 광고 단가를 상승시킨 경험이 있다. 이번 생성형 AI 광고가 도입된 이후에도 광고 단가가 YoY 20% 이상으로 성장세가 가속화될 것으로 전망한다.

그림 198. NAVER 체류시간 1분당 서치플랫폼 매출액 추정



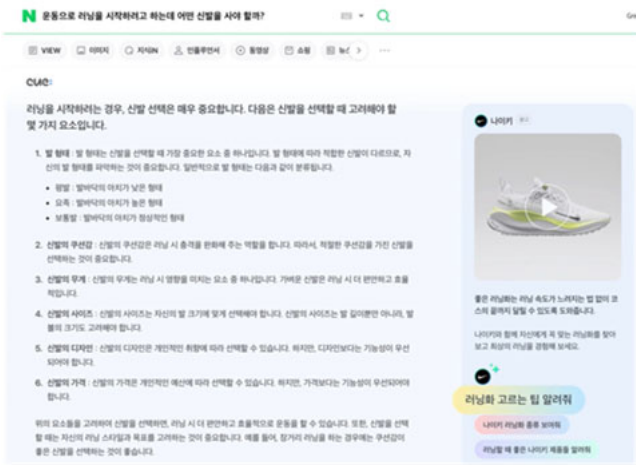
자료: 미래에셋증권 리서치센터 추정

그림 199. NAVER 체류시간 1분당 DA 매출액 추정



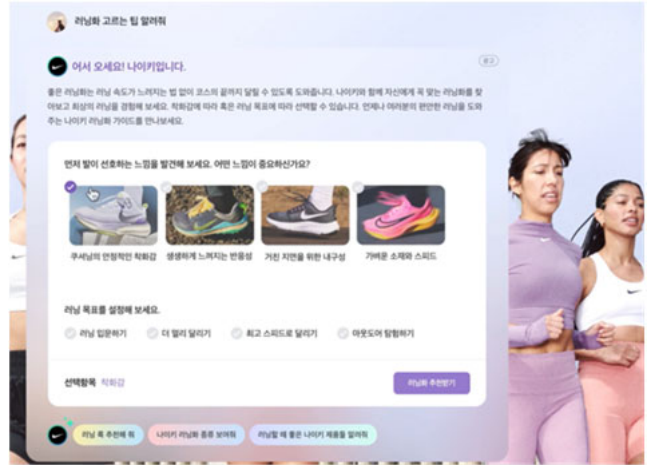
자료: 미래에셋증권 리서치센터 추정

그림 200. CLOVA for AD 예시1



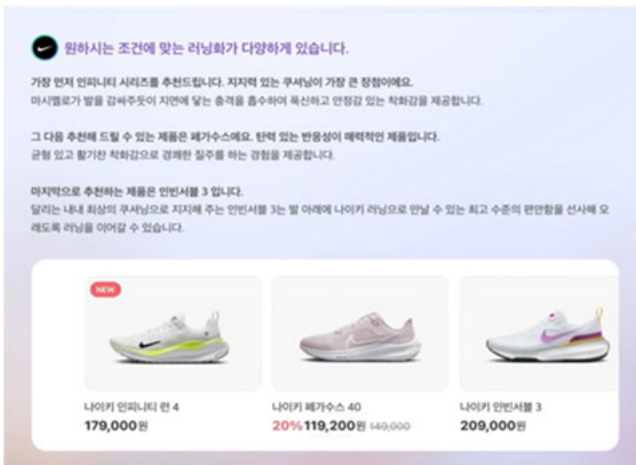
자료: NAVER, 미래에셋증권 리서치센터

그림 201. CLOVA for AD 예시2



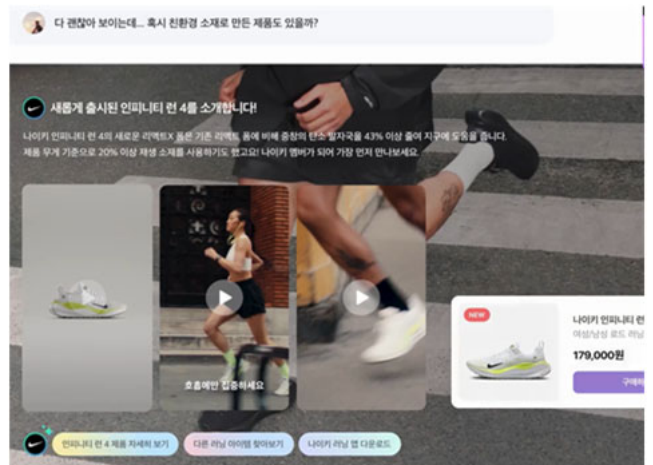
자료: NAVER, 미래에셋증권 리서치센터

그림 202. CLOVA for AD 예시3: 조건에 최적화된 상품 추천



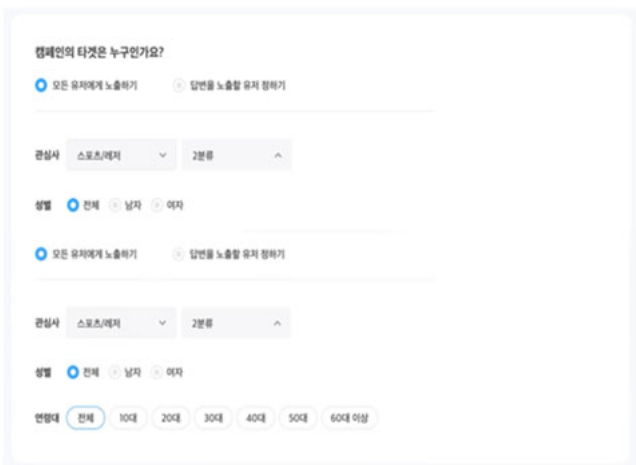
자료: NAVER, 미래에셋증권 리서치센터

그림 203. CLOVA for AD 예시4: 대화를 통한 상품 추천



자료: NAVER, 미래에셋증권 리서치센터

그림 204. CLOVA for AD 예시5: 타겟층 세부 설정



자료: NAVER, 미래에셋증권 리서치센터

그림 205. CLOVA for AD 예시6: 브랜드 말투 학습



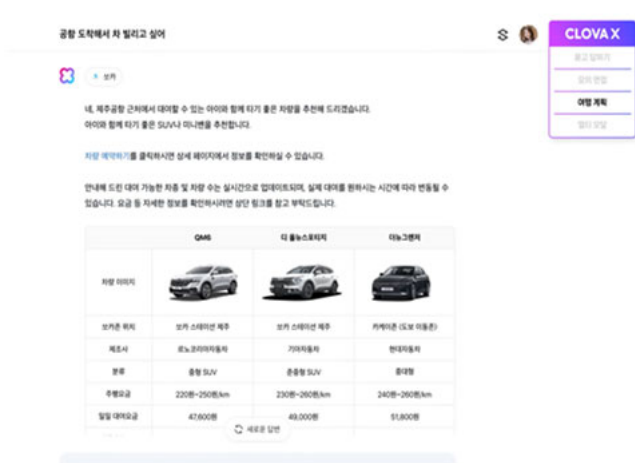
자료: NAVER, 미래에셋증권 리서치센터

중개 및 판매 수수료 BM 추가 가능성

챗봇에 광고가 추가되기 시작하면 구독 외에도 중개 및 판매수수료 BM이 추가적으로 가능할 것으로 판단한다. 그리고 다양한 버티컬 서비스를 보유하고 있는 플랫폼의 챗봇 서비스가 가장 큰 경쟁력을 보여줄 것이다. 챗봇을 업무에 보조하는 용도로 주로 사용하는 사람들도 있을 것이나 대다수의 사람들은 챗봇을 어떠한 계획을 세우거나 무엇을 구매하기 위해 사용할 가능성이 높다.

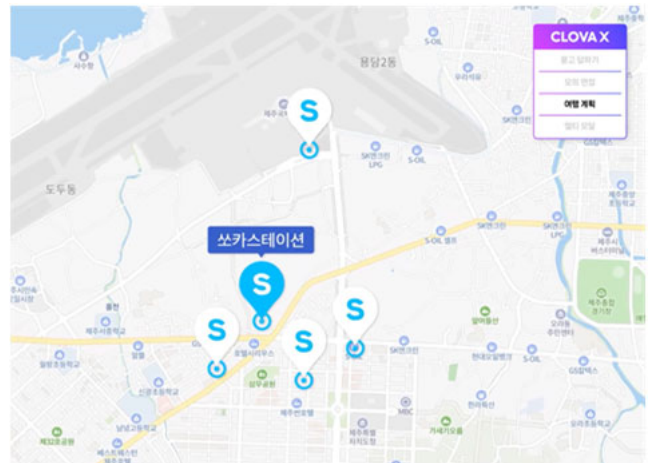
계획은 여행에 대한 계획(행선지, 숙소, 교통편 등), 장소 방문에 대한 계획(음식점, 병원, 헤어샵 등) 등이 대표적일 것이다. 계획도 결국 상품/서비스 등 무언가의 구매로 이어진다. 챗봇은 구매로 이어질 가능성이 높은 창구이기에 필연적으로 광고 수요가 강하게 발생할 것이다. 개별 카테고리에서의 신규 광고 매출 외에도 구매까지 연결되었을 경우 일정 수준의 Take rate를 수취하는 등의 중개 수수료가 신규로 발생할 전망이다.

그림 206. 클로바X 예시 1



자료: NAVER, 미래에셋증권 리서치센터

그림 207. 클로바X 예시 2



자료: NAVER, 미래에셋증권 리서치센터

그림 208. 클로바X 예시 3

9월 30일에서 10월 4일 일정으로 제주도 가는데 30만원대 서귀포 숙소 5곳 추천해줘

- 제주도 여행을 계획 중이시군요! 고객님의 여행 일정에 맞는 숙소를 추천해 드리겠습니다.
- 제주롯데호텔: 시설과 뷰가 좋으며 수영장과 조식이 매우 좋은 곳입니다.
 - 시절인연: 감성적인 분위기와 자쿠지가 있는 숙소로, 가격도 숙소 퀄리티에 비해 생각보다 괜찮은 편입니다.
 - 더그랜드 성오름 호텔: 수영장과 룸 컨디션이 좋으며, 부대시설과 가격이 착한 숙소입니다.
 - 풀라밍고 다이 라이 리조트: 숙박 + 양복 픽업서비스 + 석식 + 조식 구성으로 20만원대 후반으로 예약이 가능하며 독채 숙소와 수영장을 이용할 수 있습니다.
 - 롯데호텔: 시설과 뷰가 좋으며 소핑몰과 붙어있는 것이 장점입니다. 미리 예약하지 않으면 풀복일 수 있으니 미리 예약하는 것이 좋습니다. 1박 가격은 3월 평일 기준으로 20만원대부터 있습니다.
- 사용자님의 여행 일정과 선호하는 숙소 유형을 말씀해 주시면 더욱 정확한 추천을 해드릴 수 있습니다.

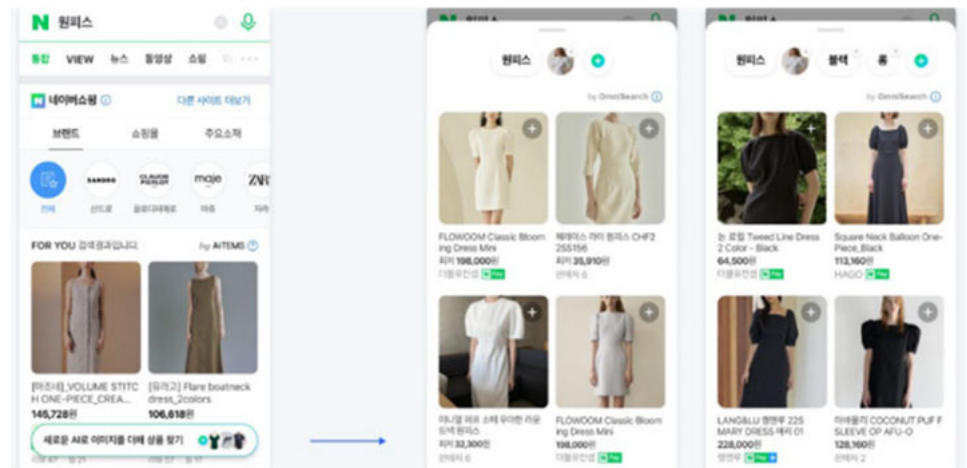
자료: 본인 화면, 클로바X, 미래에셋증권 리서치센터

2. 커머스

AI 기반 멀티모달 서비스 도입

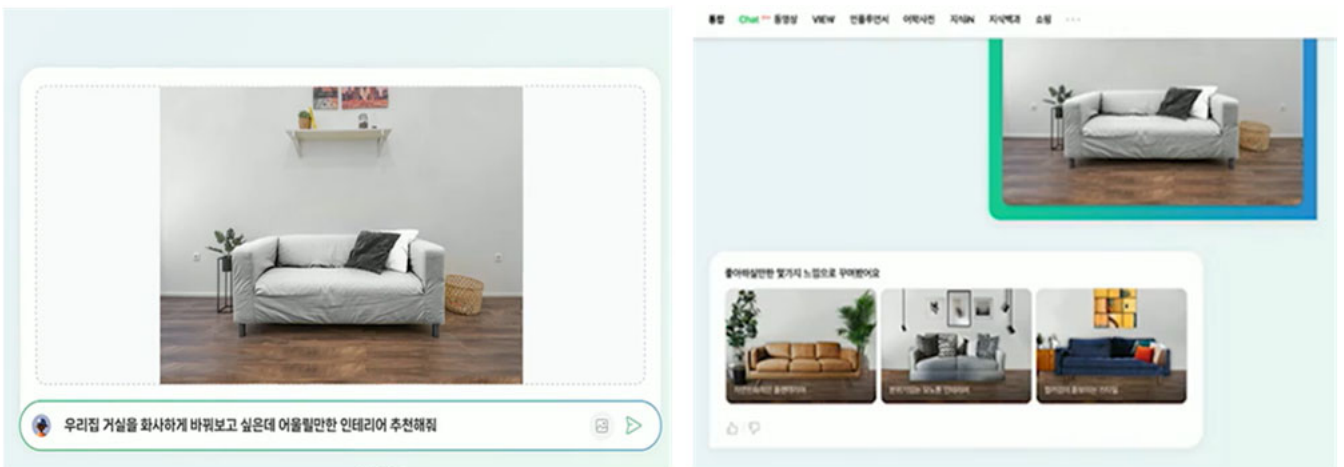
AI 기반으로 이미지, 음성, 텍스트를 동시에 인식하는 멀티모달 서비스가 도입되면서 이용자가 원하는 상품을 보다 효율적으로 찾을 수 있게 변할 것이다. 멀티모달 서비스 도입으로 쇼핑 이용자가 고민할 만한 상품을 보기까지 걸리는 시간이 대폭 단축될 것이다. 그 결과 쇼핑 피로도는 낮아지고 구매 전환율이 상승하게 된다. 또한 검색 이용자는 배경, 공간 등에 어울리는 상품 추천까지 받을 수 있을 것으로 보인다. 더 나아가면 공간에 대한 코디와 더불어 개인의 사진에 대한 AI 커스터마이징이 가능해질 것으로 기대된다.

그림 209. NAVER 멀티모달 AI 서비스 음니서치



자료: NAVER, 미래에셋증권 리서치센터

그림 210. NAVER 이미지+텍스트 결합형 검색 예시: "우리집 거실을 화사하게 바꿔보고 싶은데 어울릴만한 인테리어 추천해줘"



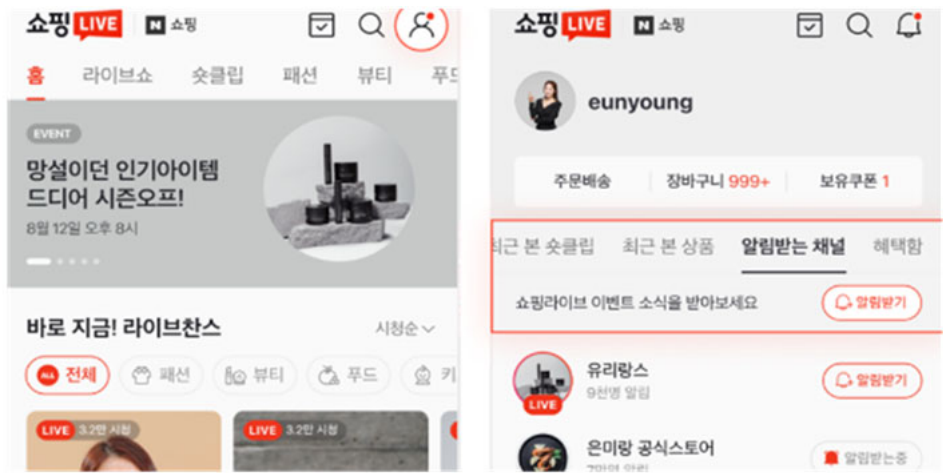
자료: NAVER, 미래에셋증권 리서치센터

라이브커머스 성장 가속화

생성 AI의 도입으로 쇼핑라이브 성장도 가속화될 것이다. 생성 AI의 도입으로 1) 특정 품목에 관심이 있을 잠재 고객을 타겟팅 할 수 있게 되고, 2) 대본 생성을 통한 판매자 생산성이 향상되며, 3) AI 휴먼 라이브 방송이 가능해지고, 4) 다국어 라이브 스트리밍이 가능해질 전망이다.

AI 기반 유저 타겟팅이 고도화되면 될수록 TV홈쇼핑 시장의 붕괴는 가속화되고 라이브 커머스 침투율이 올라갈 것이다. TV홈쇼핑이 쇠퇴 원인은 TV 시청자의 급감도 있지만 홈쇼핑 채널들이 고객들에 대한 능동적인 타겟팅이 불가능하다는 점에 있기 때문이다. AI 기반 메시지 마케팅은 특정 상품을 사길 원하는 구매 의지가 높은 고객을 해당 방송 시간에 맞춰 데려올 수 있다. 그 결과 판매자의 판매 효율과, 구매자의 시간 및 구매 효율성 모두가 향상된다.

그림 211. NAVER 쇼핑라이브 알림 기능



자료: 네이버쇼핑, 미래에셋증권 리서치센터

생성 AI의 도입은 라이브커머스를 위한 대본 준비의 부담도 완화시켜 판매자 측면에서의 확장도 가속화되고 있다. 기존 홈쇼핑에서는 대본 작성에 대한 부담, 실시간 의견 취합 등의 이슈로 브랜드 규모가 있는 곳이 주로 이용했지만 생성 AI의 도입을 통해 라이브커머스의 기회가 중소상공인에게도 열리고 있다. NAVER의 AI 큐시트 헬퍼의 경우 상품 특징, 고객 리뷰 등을 바탕으로 전문 쇼호스트나 인플루언서의 화법이 반영된 대본을 작성해준다. 지난 6월부터 도입되어 시간이 많이 지나지 않았음에도 AI 큐시트 헬퍼 침투율은 쇼핑라이브 판매자의 10%까지 올라왔다.

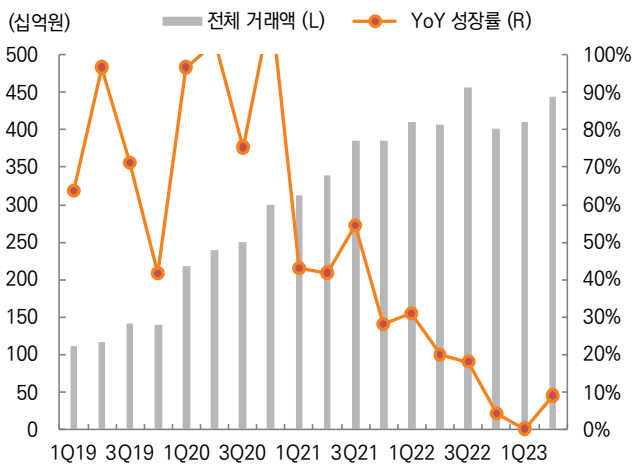
3. 콘텐츠: 생성 AI 도입을 통한 공급난 해소 기대

성장 둔화에 직면한 웹툰·웹소설 산업

생성형 AI의 도입은 현재 웹툰·웹소설 산업이 직면하고 있는 콘텐츠 공급난을 완화시켜 줄 것이다. 글로벌 웹툰·웹소설 산업은 22년 이후 성장세가 급격하게 둔화되고 있다. 글로벌 최대 웹툰 플랫폼인 NAVER와 카카오조차도 2H22부터 콘텐츠 거래액의 YoY 역성장이 관찰되고 있는 상황이다.

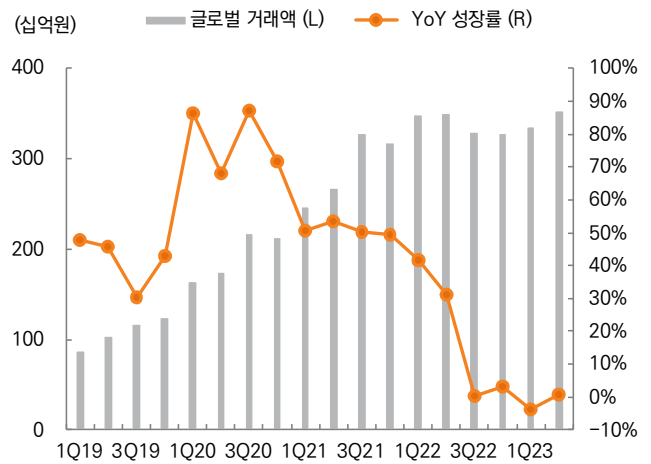
성장세의 급격한 둔화는 콘텐츠 공급 부족에 기인한다. 19~21년 웹툰 작가가 각광받으면서 신규 웹툰 작가들이 대거 유입되고 콘텐츠 공급량이 급격히 확대되었으나 현재는 신규 작가 유입 증가세가 일단락된 상태다. 현재 상황에서 콘텐츠 공급 확대를 위해 기대할 수 있는 부분은 생산성 증가 부분이다.

그림 212. 네이버웹툰 글로벌 분기 거래액



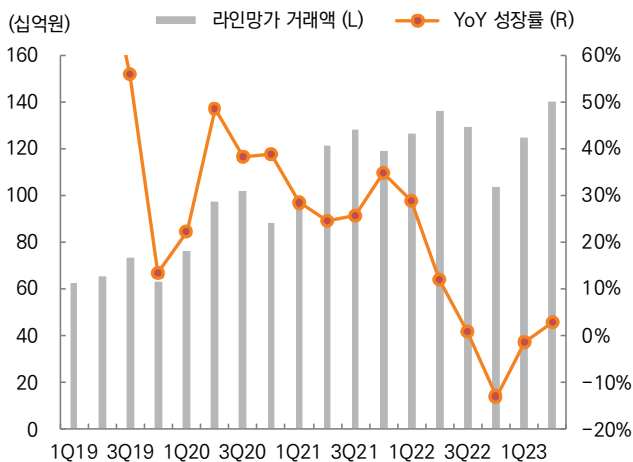
자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 213. 카카오 웹툰 글로벌 분기 거래액



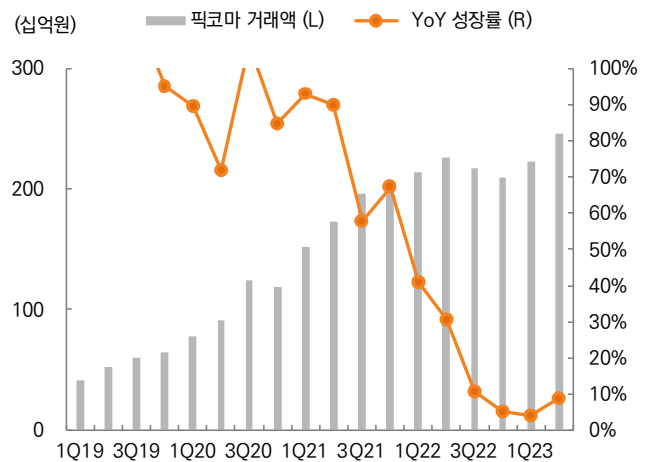
자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 214. 라인망가 분기 거래액



자료: 센서타워, 미래에셋증권 리서치센터 추정

그림 215. 픽코마 분기 거래액

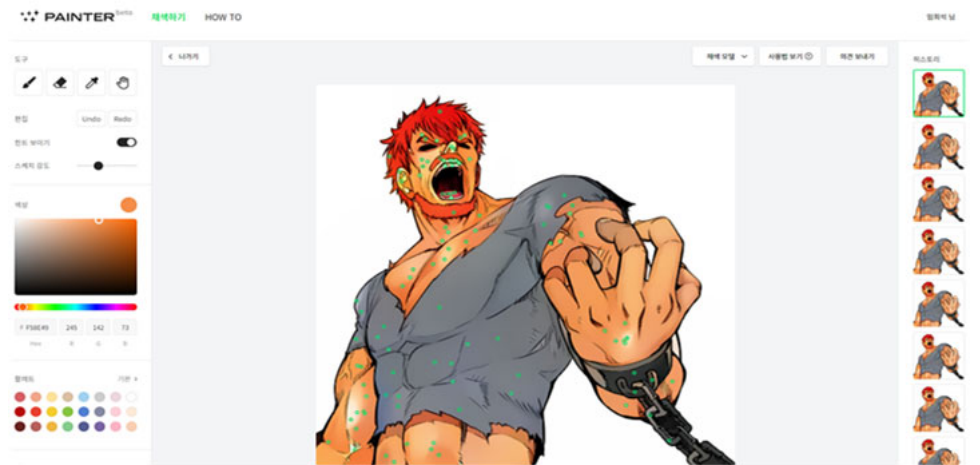


자료: 센서타워, 미래에셋증권 리서치센터 추정

생성 AI 도입이 불러올 생산성 향상

생성형 AI와 AI툴의 도입이 양질의 웹툰/웹소설 콘텐츠 증가로 이어질 전망이다. 이미 웹툰/웹소설 시나리오를 생성하는 플랫폼들이 등장하고 있으며 클릭 한 번으로 AI가 채색을 완성시켜주는 툴도 고도화되고 있다. 웹툰은 작가가 하루 평균 그림을 그리는 시간이 10시간 반(콘진원 2022년 웹툰 작가 실태 조사)에 달할 정도로 채색 노동 부담이 높다. AI 툴의 적용으로 채색 부담이 크게 완화될 것으로 기대된다.

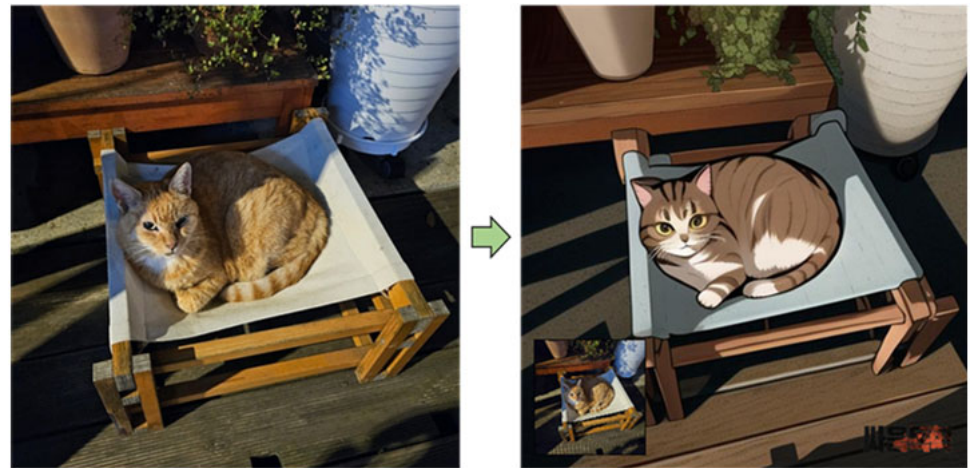
그림 216. AI 페인터: 클릭 몇 번만으로 손쉽게 채색이 가능하다.



자료: 본인 화면, AI 페인터, 미래에셋증권 리서치센터

글로벌 최대 웹툰 플랫폼 네이버웹툰은 AI 기술을 웹툰에 발빠르게 도입하고 있다. AI 페인터를 통해서 자동으로 채색할 수 있도록 도와주며 튠필터의 도입을 통해 사진이나 영상을 웹툰으로 변환해주기도 한다. NAVER의 AI 채색 툴 AI 페인터는 그림 쪽으로는 문외한인 필자조차도 10분 정도면 한 장을 완성할 정도까지 고도화되었다.

그림 217. 튠필터 이용 예시: 집주변 길냥이를 싸움독학 웹툰 그림체로 변환해 보았다.



자료: 본인 화면, 튠필터, 미래에셋증권 리서치센터

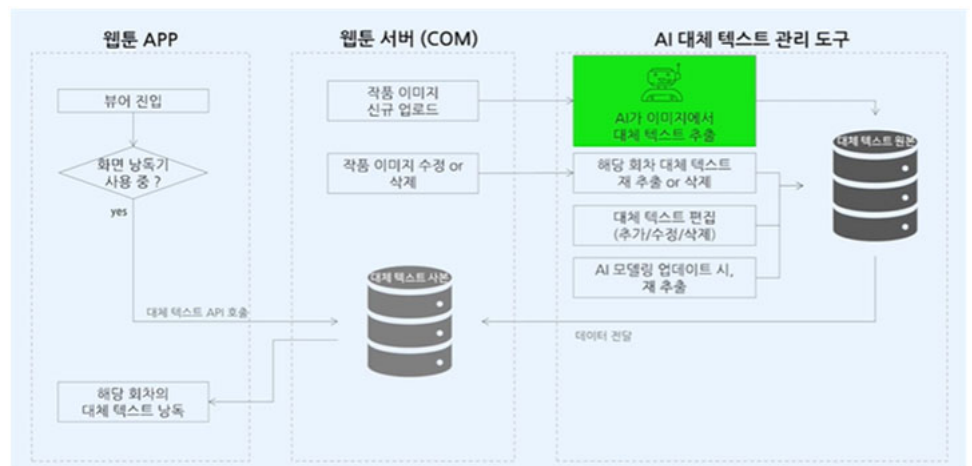
채색뿐만 아니라 이미지의 대체 텍스트화 측면에서도 AI는 혁신을 일으키고 있다. 모든 웹툰 이미지에 대해 사람이 일일이 텍스트 정보를 입력하는 데는 한계가 존재했으나 AI는 무리 없이 가능하기 때문이다. AI 기술을 통해 웹툰 이미지를 텍스트로 변환하고 이를 다시 청각 정보로 변환해 시각 약자도 즐길 수 있는 배리어프리 웹툰을 제공할 수 있을 것으로 기대된다. 시각 약자도 웹툰 콘텐츠를 즐길 수 있게 되면서 이용자 측면에서의 확장이 추가적으로 일어날 전망이다.

그림 218. AI 기술 적용을 통한 이미지의 청각 정보 변환



자료: 네이버웹툰, 미래에셋증권 리서치센터

그림 219. 네이버웹툰 AI 대체 텍스트 서비스 구조도



자료: 네이버웹툰, 미래에셋증권 리서치센터

실적 전망

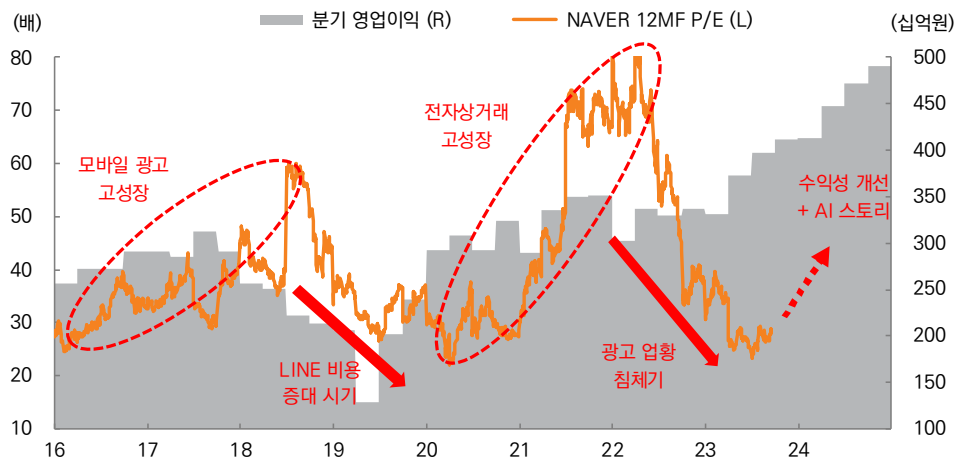
표 34. 분기 및 연간 실적 전망

(십억원)

	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23F	4Q23F	2021	2022	2023F	2024F
매출액	1,845	2,046	2,057	2,272	2,280	2,408	2,506	2,570	6,818	8,220	9,764	11,087
(% YoY)	23.1%	23.0%	19.1%	17.8%	23.6%	17.7%	21.8%	13.1%	28.5%	20.6%	18.8%	13.5%
서치플랫폼	850	906	896	916	852	910	925	967	3,308	3,568	3,655	4,000
검색	623	661	665	660	656	689	705	713	2,400	2,609	2,763	3,039
디스플레이	221	241	228	250	192	217	217	250	885	941	877	945
기타	6	4	3	6	4	4	4	4	23	19	15	17
커머스	416	439	458	487	606	633	659	703	1,488	1,801	2,600	3,088
커머스 광고	261	274	268	276	264	280	294	316	944	1,078	1,155	1,415
중개 및 판매	132	142	157	173	303	310	321	340	476	604	1,274	1,463
멤버십	24	24	34	37	39	42	44	46	69	119	172	210
핀테크	275	296	296	320	318	340	350	371	979	1,187	1,379	1,648
콘텐츠	212	300	312	437	411	420	463	417	660	1,262	1,712	1,847
웹툰	164	251	268	383	353	370	410	362	492	1,066	1,495	1,603
스노우	33	27	29	38	45	36	38	39	84	128	158	179
기타	15	22	14	16	13	15	15	16	84	67	59	66
클라우드	92	105	95	111	93	104	109	113	383	403	419	502
NCP/웍스(B2B)	86	92	92	100	91	99	102	105	329	370	398	454
클로바/랩스	7	13	3	11	2	5	6	8	53	33	21	49
영업비용	1,543	1,710	1,727	1,935	1,950	2,035	2,108	2,158	5,492	6,915	8,251	9,265
개발/운영비용	472	515	522	562	635	644	659	670	1,753	2,070	2,609	2,830
파트너	624	720	722	898	832	866	901	924	2,172	2,964	3,522	3,985
인프라	133	141	152	154	133	142	149	155	496	581	579	684
마케팅	315	333	331	322	350	383	399	409	1,072	1,301	1,542	1,765
영업이익	302	336	330	336	330	373	398	412	1,325	1,305	1,513	1,822
(% YoY)	4.5%	0.2%	-5.6%	-4.2%	9.5%	10.9%	20.4%	22.5%	9.1%	-1.6%	16.0%	20.4%
영업이익률	16.4%	16.4%	16.1%	14.8%	14.5%	15.5%	15.9%	16.0%	19.4%	15.9%	15.5%	16.4%
지배주주순이익	171	183	257	149	55	268	310	323	16,490	760	955	1,441
NPM	9.3%	9.0%	12.5%	6.5%	2.4%	11.1%	12.4%	12.6%	241.9%	9.2%	9.8%	13.0%

자료: 미래에셋증권 리서치센터

그림 22. NAVER 12MF P/E, 분기 영업이익 추이



자료: 미래에셋증권 리서치센터 추정

NAVER (035420)

예상 포괄손익계산서 (요약)

(십억원)	2022	2023F	2024F	2025F
매출액	8,220	9,764	11,087	12,516
매출원가	0	0	0	0
매출총이익	8,220	9,764	11,087	12,516
판매비와관리비	6,915	8,251	9,265	10,123
조정영업이익	1,305	1,513	1,822	2,392
영업이익	1,305	1,513	1,822	2,392
비영업손익	-221	-138	84	61
금융손익	-13	-16	-10	13
관계기업등 투자손익	97	243	321	300
세전계속사업손익	1,084	1,375	1,906	2,453
계속사업법인세비용	411	458	572	711
계속사업이익	673	916	1,334	1,741
중단사업이익	0	0	0	0
당기순이익	673	916	1,334	1,741
지배주주	760	955	1,441	1,881
비지배주주	-87	-39	-107	-139
총포괄이익	-715	916	1,334	1,741
지배주주	-681	873	1,272	1,660
비지배주주	-34	43	63	82
EBITDA	1,866	2,041	2,367	2,920
FCF	753	1,550	1,674	2,017
EBITDA 마진율 (%)	22.7	20.9	21.3	23.3
영업이익률 (%)	15.9	15.5	16.4	19.1
지배주주귀속 순이익률 (%)	9.2	9.8	13.0	15.0

예상 재무상태표 (요약)

(십억원)	2022	2023F	2024F	2025F
유동자산	6,440	6,208	7,667	9,787
현금 및 현금성자산	2,724	2,420	3,804	5,847
매출채권 및 기타채권	0	0	0	0
재고자산	9	10	12	13
기타유동자산	3,707	3,778	3,851	3,927
비유동자산	27,459	29,131	28,942	28,765
관계기업투자등	18,541	18,541	18,541	18,541
유형자산	2,458	2,362	2,290	2,273
무형자산	1,481	3,249	3,131	2,971
자산총계	33,899	35,340	36,610	38,553
유동부채	5,481	5,778	5,804	6,093
매입채무 및 기타채무	899	917	936	954
단기금융부채	1,575	1,603	1,544	1,748
기타유동부채	3,007	3,258	3,324	3,391
비유동부채	4,968	5,333	5,380	5,430
장기금융부채	2,622	3,621	3,621	3,621
기타비유동부채	2,346	1,712	1,759	1,809
부채총계	10,449	11,111	11,184	11,523
지배주주지분	22,744	23,562	24,866	26,609
자본금	16	16	16	16
자본잉여금	1,556	1,556	1,556	1,556
이익잉여금	23,646	24,464	25,767	27,510
비지배주주지분	706	667	560	421
자본총계	23,450	24,229	25,426	27,030

예상 현금흐름표 (요약)

(십억원)	2022	2023F	2024F	2025F
영업활동으로 인한 현금흐름	1,453	1,800	1,974	2,367
당기순이익	673	916	1,334	1,741
비현금수익비용가감	1,318	968	1,091	1,191
유형자산감가상각비	519	346	372	367
무형자산상각비	42	182	173	160
기타	757	440	546	664
영업활동으로인한자산및부채의변동	307	355	96	98
매출채권 및 기타채권의 감소(증가)	-56	0	0	0
재고자산 감소(증가)	0	-2	-1	-2
매입채무 및 기타채무의 증가(감소)	133	0	0	0
법인세납부	-861	-458	-572	-711
투자활동으로 인한 현금흐름	-1,216	-2,238	-394	-390
유형자산처분(취득)	-689	-250	-300	-350
무형자산감소(증가)	-55	-1,950	-55	0
장단기금융자산의 감소(증가)	-444	-38	-39	-40
기타투자활동	-28	0	0	0
재무활동으로 인한 현금흐름	-339	134	-196	66
장단기금융부채의 증가(감소)	235	271	-59	204
자본의 증가(감소)	66	0	0	0
배당금의 지급	-213	-138	-138	-138
기타재무활동	-427	1	1	0
현금의 증가	-57	-305	1,384	2,043
기초현금	2,781	2,724	2,420	3,804
기말현금	2,724	2,420	3,804	5,847

자료: NAVER, 미래에셋증권 리서치센터

예상 주당가치 및 valuation (요약)

	2022	2023F	2024F	2025F
P/E (x)	38.3	35.5	23.5	18.0
P/CF (x)	14.6	18.0	14.0	11.6
P/B (x)	1.2	1.4	1.3	1.2
EV/EBITDA (x)	16.1	17.3	14.3	10.9
EPS (원)	4,634	5,824	8,784	11,465
CFPS (원)	12,138	11,485	14,785	17,876
BPS (원)	144,841	149,826	157,772	168,399
DPS (원)	914	914	914	914
배당성향 (%)	20.4	15.0	10.3	7.9
배당수익률 (%)	0.5	0.4	0.4	0.4
매출액증가율 (%)	20.6	18.8	13.5	12.9
EBITDA증가율 (%)	6.1	9.4	15.9	23.4
조정영업이익증가율 (%)	-1.6	16.0	20.4	31.3
EPS증가율 (%)	-95.4	25.7	50.8	30.5
매출채권 회전을 (회)	0.0	0.0	0.0	0.0
재고자산 회전을 (회)	1,111.4	1,031.7	1,010.4	1,007.6
매입채무 회전을 (회)	0.0	0.0	0.0	0.0
ROA (%)	2.0	2.6	3.7	4.6
ROE (%)	3.3	4.1	6.0	7.3
ROIC (%)	37.6	29.9	32.4	45.6
부채비율 (%)	44.6	45.9	44.0	42.6
유동비율 (%)	117.5	107.4	132.1	160.6
순차입금/자기자본 (%)	1.0	3.2	-2.8	-9.5
조정영업이익/금융비용 (x)	18.2	19.5	23.1	29.9



(Analyst) 김수진 soojin.kim@miraeasset.com



017670 · 유무선통신

SK텔레콤

AI에 대한 진심이 통하기 시작한다

(유지)

매수

목표주가

76,000원

상승여력

46.2%

현재주가(23/9/27)

52,000원

KOSPI	2,465.07	시가총액(십억원)	11,379	발행주식수(백만주)	219	외국인 보유비중(%)	41.4
-------	----------	-----------	--------	------------	-----	-------------	------

Report summary

통신사와 생성형AI의 긴밀한 시너지 기대 - SKT의 내공을 확인할 때

18~19일 이틀간 SKT의 AI 비즈니스 현황을 주제로 NDR을 진행했다. 이번 NDR을 통해 SKT의 AI 전략이 글로벌 스탠다드에 맞춰 앞서 나가고 있다는 점을 확인했다. 동사는 1) 에이닷 등 생성AI 직접 개발 뿐만 아니라 엔트로픽처럼 글로벌 대표 AI 기업에 투자를 통해 기술력을 확보해 나가고 있으며, 2) 도이치텔레콤, 싱텔 등 글로벌 통신사와 얼라이언스를 통해 통신 산업에서 AI 시장 확대를 주도, 사업화 하려고 시도 중이고, 3) SK텔레콤 아메리카(SKTA) 추가 출자로 AI 비즈니스의 글로벌 성장 루트까지 확보했다.

현재 SKT에서 생성AI를 사용하는 대표 서비스로는 1) AICC 활용이 있고, 2) 에이닷 앱을 통해 안드로이드 OS 발신 전화의 텍스트 전환 및 음성 요약 기능을 제공하고 있다. 향후 아이폰에서도 서비스 할 예정이다.

현재 여러 산업군에서 생성AI 앱을 개발중이고, 이 가운데 클라우드 인프라가 반드시 필요하므로 동사의 엔터프라이즈 사업부 매출부터 점진적으로 증가할 수 있을 것으로 예상된다. 통신서비스 톱픽으로 추천한다.

Key data

Price performance



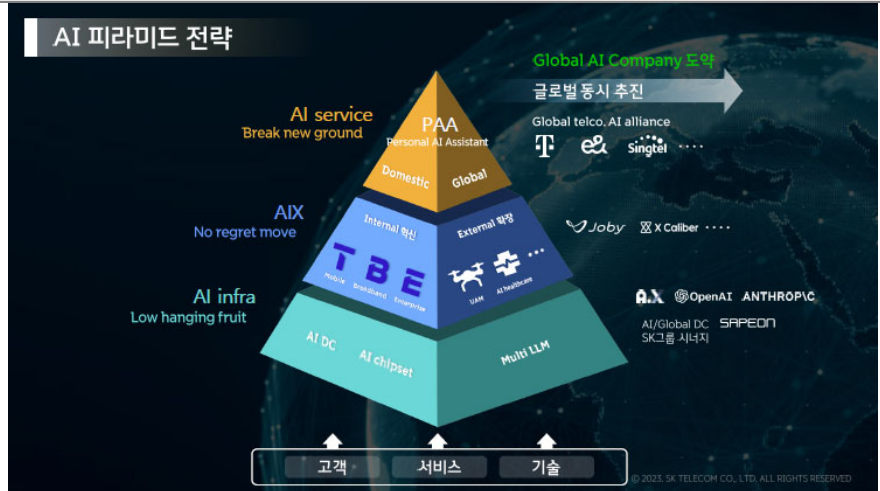
(%)	1M	6M	12M
절대주가	10.3	7.4	-1.0
상대주가	12.7	5.0	-10.6

Earnings and valuation metrics

계산기 (12월)	2020	2021	2022	2023F	2024F
매출액 (십억원)	16,088	16,749	17,305	17,588	17,958
영업이익 (십억원)	1,249	1,387	1,612	1,749	1,818
영업이익률 (%)	7.8	8.3	9.3	9.9	10.1
순이익 (십억원)	1,504	2,408	912	1,090	1,123
EPS (원)	3,726	6,841	4,169	4,982	5,130
ROE (%)	6.4	13.6	8.0	9.3	9.2
P/E (배)	12.8	8.5	11.4	10.4	10.1
P/B (배)	0.7	1.1	0.9	0.9	0.9
배당수익률 (%)	4.2	4.6	7.0	6.4	6.4

주: K-IFRS 연결 기준, 순이익은 지배주주 귀속 순이익
자료: SK텔레콤, 미래에셋증권 리서치센터

그림 223. SKT가 9월 26일 발표한 AI 피라미드 전략



SKT가 9월 26일 기자간담회를 통해 대대적으로 AI 피라미드 전략을 발표했다. 데이터 센터 규모를 2030년까지 두 배 늘리고, AI 관련 투자를 향후 5년간 전체 투자 대비 33%로 지난 5년 대비 3배 늘리겠다는 목표를 내세웠다. AI 관련 매출 비중은 지난해 9%에서 2030년 36%로 상향조정했다.

이날 에이닷을 공식 런칭했다. 지난해 5월 베타 버전 이후 1년 4개월 만의 공식 발표다. 추가로 SKT가 개발하는 LLM 브랜드명을 '에이닷엑스(A.X)'로 확정했다고 발표했다. 동사는 자사의 LLM 에이닷엑스와, 투자한 기업 '엔트로픽', '코난테크놀로지' 등의 기술력으로 생성형AI 투트랙 전략을 펼치고 있다. 또 AI 반도체 사피온과 클라우드 인프라로 AI 풀스택을 제공하는 AI 비즈니스가 가능하다.

그림 224. SKT의 AI 제품화 전략

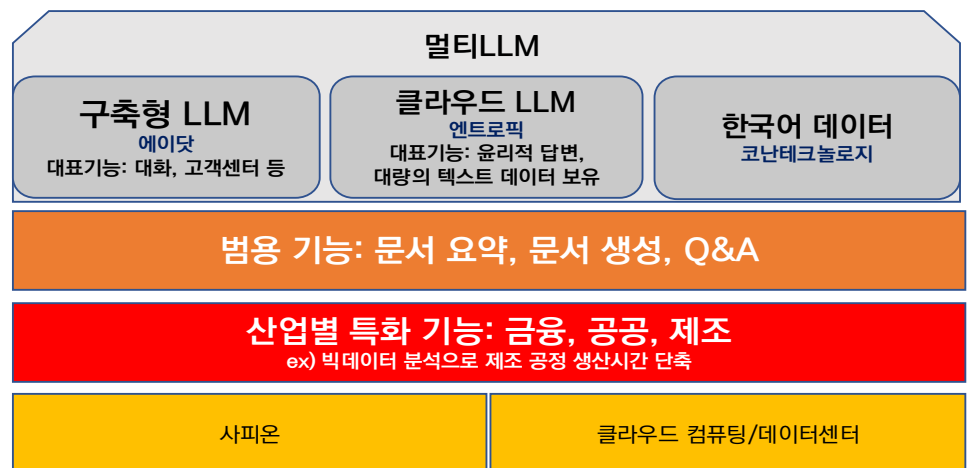
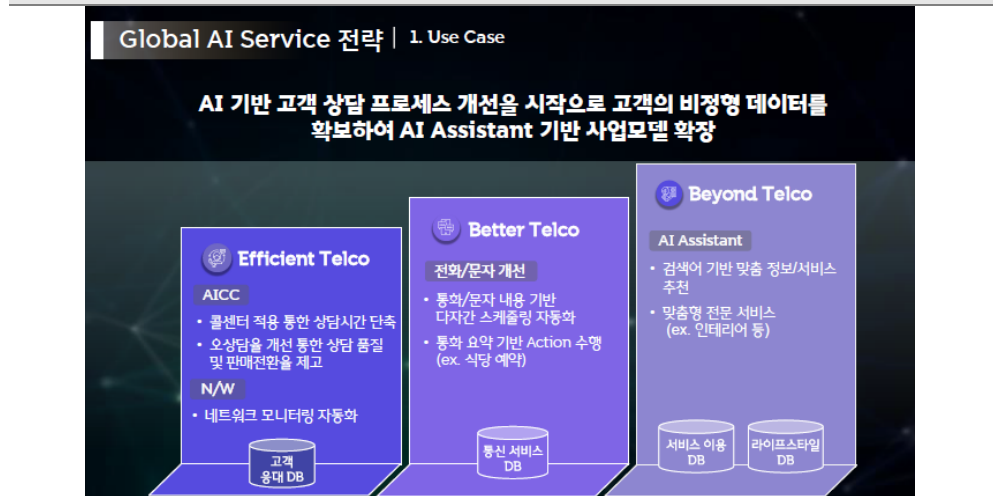


그림 225. SKT의 Global AI 서비스 전략 - 3단계 Use Case



자료: SK텔레콤, 미래에셋증권 리서치센터

동사는 AI를 본격적으로 활용함으로써 통신업에서의 시너지를 낼 수 있을 것으로 기대한다. 1) 우선 통신업에서 큰 비용으로 작용하는 콜센터에 생성형AI를 적용해 비용 절감 효과를 가져오고, 소비자의 편리성을 극대화할 수 있다. 또 네트워크 데이터를 처리할 때 생성AI를 활용하는 것도 비용 절감의 방안이 될 수 있다.

2) 뿐만 아니라 유무선 가입자의 락인 효과를 강화할 수 있다. 전화와 문자 서비스에 AI 기능을 접목시킴으로서, 최근 점유율이 확대되고 있는 알뜰폰과의 차별화 된 서비스를 제공할 수 있기 때문이다. 3) 나아가 지금까지 없었던 새로운 서비스는 ARPU 인상의 동기부여로 연결될 수 있을 것으로 판단한다. 새로운 서비스의 영역은 향후 무궁무진하게 발전할 수 있다. 에이닷의 번역 서비스가 될 수도 있고, 음성 요약 서비스가 될 수도 있다.

통신사는 그 어떤 산업보다 개개인의 다양한 데이터를 활용할 수 있다는 강점을 보유하고 있다. 따라서 기존 통신사의 강점에 선진적인 기술력이 더해진다면, 새로운 생성 AI 시대에 주도권을 잡을 수 있을 것이라고 판단한다.

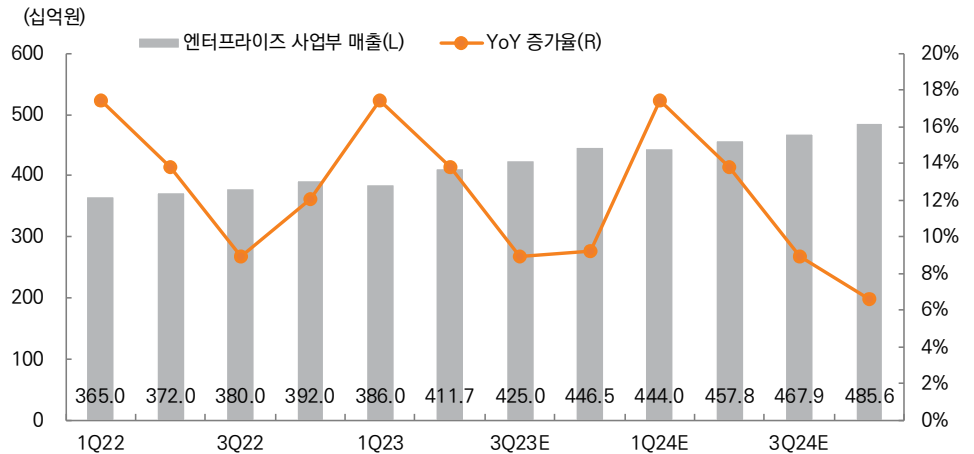
단기적으로는 엔터프라이즈 사업부의 매출 증가를 기대한다. 23년 데이터센터와 클라우드 사업부 매출은 각각 2,034억원(+30% YoY)과 1,813억원(+43% YoY)을 기록할 것으로 예상된다. 엔터프라이즈 LLM 비즈니스의 확대는 당장 데이터센터 및 클라우드 사업부의 성장으로 연결될 수 있으므로, 인프라 관련 매출과 B2B사업부의 매출부터 점진적으로 증가할 것으로 기대한다.

표 35. SKT의 AI 관련 기업 투자 현황

기업명	분야	SKT와의 관계 및 특징	투자규모	투자시기	그 외 투자자들
사피온 (SAPEON)	AI 반도체 설계	- 본래 SKT의 AI 반도체 계열사로, 2022년 SKT로부터 분사 - 사피온의 AI 반도체는 엔비디아의 GPU 대비 가격은 절반 수준 but 답러닝 연산은 1.5배 빠르고 전력 소모량은 80% 수준에 불과	500억원	2022년 2월	어센트에쿼티파트너스, 대보정보통신, 미래에셋벤처투자, 위벤처스, GS네오텍, GS건설 등
씨메스 (CMES)	AI 로봇틱스	- 2016년에 9억원 투자 후 2022에 100억을 투자해 씨메스의 2대주주가 됨 - 씨메스의 로봇 제어 기술과 SKT의 비전 AI 기술을 결합한 'AI 물류 로봇 공동 사업' 추진	100억원	2022년 3월	GS리테일, SBI Investment Korea, K2 Investment Partners, Kolon Investment
코난테크놀로지	LLM모델, 자연어처리	- 에이닷(A.ida) 서비스 품질 고도화 및 사피온과의 B2B 시장 창출 계획 - 코난테크놀로지의 상당지분을 인수함으로써 2대 주주 지위를 확보함	224억원	2022년 11월	한국항공우주
스캐터랩	감성형 AI챗봇	- 감성형 AI 모델 개발 경험을 기반으로 SKT와 사용자와 친밀한 대화를 지원하는 케어형 AI 대화 서비스를 공동개발하고자 함	150억원	2023년 4월	
Anthropic	생성형 AI	- AI 모델 '클로드'를 통해 SKT와 함께 한국어 언어모델을 개발하고 글로벌 텔코 AI 얼라이언스를 통해 글로벌 이동통신사에 제공할 예정	1300억원 이상	2023년 8월	Spark Capital, SAP, Google, Zoom Ventures, Menlo Ventures, SVA, Salesforce Ventures, Raison, Sound Ventures 등
페르소나 AI	AI컨택센터 (AICC)	- 전략적 투자 통해 3대 주주 지위 획득 - SKT '누구' 서비스에 사용된 AI 음성인식과 페르소나의 자연어 처리 기술을 결합한 AICC 상품 '서비스형 AI콜더' 제공	미공개	2023년 8월	효성벤처스

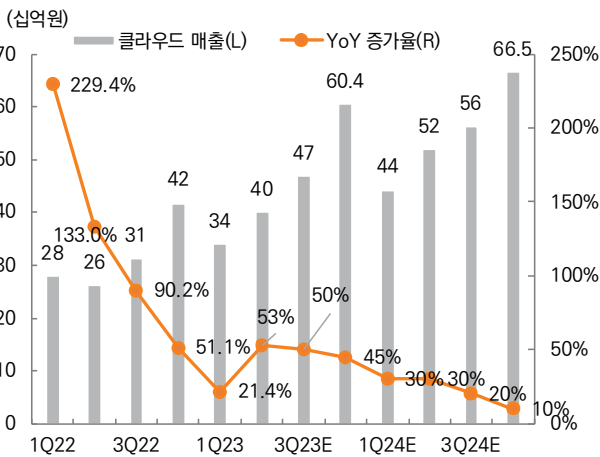
자료: 언론 종합, SK텔레콤, 미래에셋증권 리서치센터

그림 226. SKT 엔터프라이즈 사업부 매출 및 YoY 증가율 추이



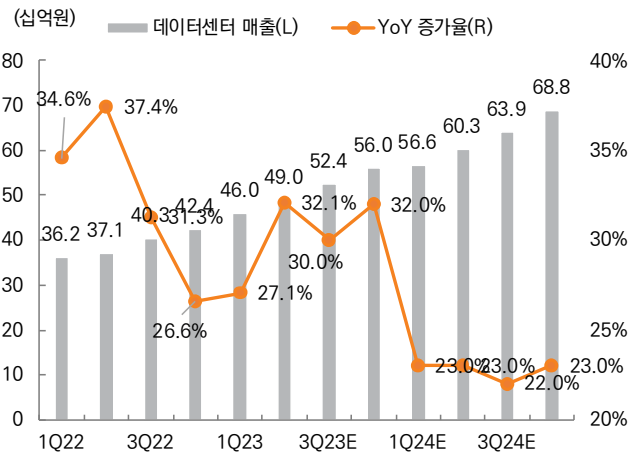
자료: SK텔레콤, 미래에셋증권 리서치센터

그림 227. SKT 클라우드 사업부 매출 및 YoY 증가율 추이



자료: SK텔레콤, 미래에셋증권 리서치센터

그림 228. SKT 데이터센터 사업부 매출 및 YoY 증가율 추이



자료: SK텔레콤, 미래에셋증권 리서치센터

표 36. SKT 실적추정표

	1Q22	2Q22	3Q22	4Q22	1Q23	2Q23	3Q23F	4Q23F	2022	2023F	2024F
영업수익	4,277.2	4,290.0	4,343.4	4,394.0	4,372.0	4,306.0	4,492.1	4,417.2	17,305	17,587	17,958
SK텔레콤	3,066.4	3,117.9	3,123.0	3,096.3	3,117.0	3,162.7	3,322.0	3,232.6	12,090	12,404	13,006
이동전화수익	2,587.1	2,616.0	2,734.0	2,634.0	2,627.0	2,618.0	2,771.2	2,664.5	10,571	10,681	10,845
망접속정산수익	132.4	125.0	124.3	133.1	136.5	126.9	125.9	135.8	515	525	517
기타	346.9	373.0	372.7	379.3	353.5	417.8	424.9	432.4	1,472	1,629	1,645
SK브로드밴드	1,026.0	1,033.0	1,044.0	1,053.0	1,062.0	1,068.0	1,090.2	1,089.9	4,156	4,310	4,531
유선통신	238.9	257.0	259.0	272.0	264.0	265.0	270.1	273.4	1,027	1,072	1,112
유료방송(IPTV+CATV)	471.0	468.0	476.4	468.0	472.0	474.0	478.2	470.4	1,883	1,895	1,902
엔터프라이즈 사업부	386.0	411.7	425.0	446.5	386.0	411.7	425.0	446.5	1,669	1,669	1,855
기타자회사	89.5	79.5	74.1	86.0	77.0	73.0	80.0	94.8	329	325	421
영업비용	3,844.7	3,830.3	3,877.9	4,140.0	3,877.4	3,843.0	3,992.7	4,125.9	15,693	15,839	16,140
인건비	637.1	593.2	581.5	638.0	617.6	590.8	634.5	672.3	2,450	2,515	2,564
지급수수료	1,310.2	1,370.9	1,384.6	1,453.0	1,356.2	1,352.3	1,439.1	1,501.4	5,519	5,649	5,721
감가상각비	946.3	940.7	933.5	934.8	933.3	948.8	941.2	974.6	3,755	3,798	3,820
영업이익	432.5	459.7	465.5	254.0	494.6	463.0	499.4	291.4	1,612	1,748	1,818
영업이익률	10.1%	10.7%	10.7%	5.8%	11.3%	10.8%	11.1%	6.6%	9.3%	9.9%	10.1%
EBITDA	1,378.8	1,400.4	1,399.0	1,188.8	1,427.9	1,411.8	1,440.6	1,266.0	5,367.0	5,546.2	5,638

자료: SK텔레콤, 미래에셋증권 리서치센터

SK텔레콤 (017670)

예상 포괄손익계산서 (요약)

(십억원)	2021	2022	2023F	2024F
매출액	16,749	17,305	17,588	17,958
매출원가	0	0	0	0
매출총이익	16,749	17,305	17,588	17,958
판매비와관리비	15,361	15,693	15,839	16,140
조정영업이익	1,387	1,612	1,749	1,818
영업이익	1,387	1,612	1,749	1,818
비영업손익	331	-376	-198	-239
금융손익	-243	-270	-318	-333
관계기업등 투자손익	446	-82	44	0
세전계속사업손익	1,718	1,236	1,551	1,579
계속사업법인세비용	447	288	415	410
계속사업이익	1,271	948	1,135	1,169
중단사업이익	1,148	0	0	0
당기순이익	2,419	948	1,135	1,169
지배주주	2,408	912	1,090	1,123
비지배주주	11	35	45	46
총포괄이익	3,781	642	1,212	1,169
지배주주	3,473	601	1,176	1,134
비지배주주	308	40	36	35
EBITDA	5,502	5,367	5,548	5,638
FCF	2,115	2,251	3,145	2,909
EBITDA 마진율 (%)	32.8	31.0	31.5	31.4
영업이익률 (%)	8.3	9.3	9.9	10.1
지배주주귀속 순이익률 (%)	14.4	5.3	6.2	6.3

예상 재무상태표 (요약)

(십억원)	2021	2022	2023F	2024F
유동자산	6,353	7,219	7,232	6,462
현금 및 현금성자산	873	1,882	1,867	941
매출채권 및 기타채권	2,462	2,450	2,463	2,535
재고자산	205	166	167	172
기타유동자산	2,813	2,721	2,735	2,814
비유동자산	24,559	24,089	24,863	26,303
관계기업투자등	2,197	1,889	1,899	1,955
유형자산	12,871	13,322	14,257	15,637
무형자산	5,942	5,400	5,237	5,237
자산총계	30,911	31,308	32,094	32,765
유동부채	6,960	8,047	7,956	8,119
매입채무 및 기타채무	2,262	2,517	2,530	2,604
단기금융부채	2,192	2,896	2,779	2,790
기타유동부채	2,506	2,634	2,647	2,725
비유동부채	11,616	11,107	11,304	11,369
장기금융부채	8,896	8,890	9,077	9,077
기타비유동부채	2,720	2,217	2,227	2,292
부채총계	18,576	19,153	19,260	19,488
지배주주지분	11,579	11,318	12,007	12,405
자본금	30	30	30	30
자본잉여금	1,771	1,771	1,771	1,771
이익잉여금	22,437	22,464	22,806	23,203
비지배주주지분	756	837	827	873
자본총계	12,335	12,155	12,834	13,278

예상 현금흐름표 (요약)

(십억원)	2021	2022	2023F	2024F
영업활동으로 인한 현금흐름	5,031	5,159	5,810	5,228
당기순이익	2,419	948	1,135	1,169
비현금수익비용가감	3,474	4,719	4,590	4,563
유형자산감가상각비	4,114	3,755	3,799	3,820
무형자산상각비	0	0	0	0
기타	-640	964	791	743
영업활동으로인한자산및부채의변동	-569	118	592	74
매출채권 및 기타채권의 감소(증가)	-164	-80	17	-58
재고자산 감소(증가)	-71	40	-5	-5
매입채무 및 기타채무의 증가(감소)	-97	268	-432	3
법인세납부	-351	-435	-297	-410
투자활동으로 인한 현금흐름	-3,486	-2,808	-2,629	-2,338
유형자산처분(취득)	-2,854	-2,892	-2,662	-2,319
무형자산감소(증가)	-378	-127	-7	0
장단기금융자산의 감소(증가)	911	155	-3	-20
기타투자활동	-1,165	56	43	1
재무활동으로 인한 현금흐름	-2,054	-1,350	-1,094	-879
장단기금융부채의 증가(감소)	-1,867	698	69	11
자본의 증가(감소)	-1,159	0	0	0
배당금의 지급	-1,029	-904	-544	-725
기타재무활동	2,001	-1,144	-619	-165
현금의 증가	-497	1,010	-15	-926
기초현금	1,370	873	1,882	1,867
기말현금	873	1,882	1,867	941

예상 주당가치 및 valuation (요약)

	2021	2022	2023F	2024F
P/E (x)	8.5	11.4	10.4	10.1
P/CF (x)	3.5	1.8	2.0	2.0
P/B (x)	1.1	0.9	0.9	0.9
EV/EBITDA (x)	4.2	3.8	3.9	4.0
EPS (원)	6,841	4,169	4,982	5,130
CFPS (원)	16,745	25,898	26,162	26,194
BPS (원)	53,176	51,889	54,946	56,762
DPS (원)	2,660	3,320	3,320	3,320
배당성향 (%)	29.6	76.4	63.9	62.1
배당수익률 (%)	4.6	7.0	6.4	6.4
매출액증가율 (%)	4.1	3.3	1.6	2.1
EBITDA증가율 (%)	1.5	-2.4	3.4	1.6
조정영업이익증가율 (%)	11.1	16.2	8.5	4.0
EPS증가율 (%)	83.6	-39.1	19.5	3.0
매출채권 회전을 (회)	8.2	8.9	8.9	8.9
재고자산 회전을 (회)	89.1	93.3	105.5	105.8
매입채무 회전을 (회)	0.0	0.0	0.0	0.0
ROA (%)	6.1	3.0	3.6	3.6
ROE (%)	13.6	8.0	9.3	9.2
ROIC (%)	4.8	6.1	6.2	6.3
부채비율 (%)	150.6	157.6	150.1	146.8
유동비율 (%)	91.3	89.7	90.9	79.6
순차입금/자기자본 (%)	77.9	77.5	74.0	78.5
조정영업이익/금융비용 (x)	5.0	4.9	4.7	4.9

자료: SK텔레콤, 미래에셋증권 리서치센터

(Analyst) 김수진 soojin.kim@miraeasset.com

018260 · IT서비스

삼성SDS

한국의 기업용 생성AI 솔루션 공개, 새 시장 열린다

(유지)

매수

목표주가

185,000원

상승여력

36.8%

현재주가(23/9/27)

135,200원

KOSPI	2,465.07	시가총액(십억원)	10,461	발행주식수(백만주)	77	외국인 보유비중(%)	15.6
-------	----------	-----------	--------	------------	----	-------------	------

Report summary

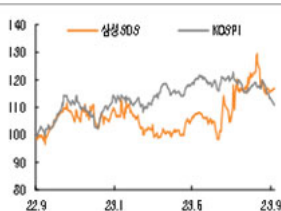
생성AI 솔루션, 발빠르게 공개 - 3Q23 실적은 매출 3.37조원 전망

지난 12일 생성AI 솔루션 1) Brity Copilot과 2) 패브릭스(FabriX)를 발표했다. 제품 출시 본격화는 추후 진행될 것으로 보이나, 관련 프로젝트가 집행되며 클라우드 매출이 점진적으로 증가해 나갈 것으로 기대한다. Brity Copilot은 메일, 메신저, 화상미팅 등 AI를 통해 기업내 생산성 솔루션의 효율성을 극대화해주는 제품이다. FabriX는 IT부서에서 사용하는 미들웨어로, 원하는 LLM(네이버 하이퍼클로바, OpenAI의 GPT-4 등)을 선택해 기업들이 어플리케이션을 개발할 수 있도록 해주는 기업용 서비스 플랫폼이다.

3Q23 매출은 3조3,710억원(+2.4% QoQ)로 예상한다. IT 서비스 매출은 1조6,610억원(+11.7% YoY), 물류 매출은 1.7조원(-4% QoQ)로 예상한다. 물류 운임 회복이 예상보다 지속 부진할 것으로 보인다. IT 서비스는 클라우드 사업부 고성장세에 힘입어 견조할 것이나 매크로 영향으로 IT 지출이 완전히 회복되지는 않았다. 영업이익률은 5.6%로, 생성AI와 클라우드 투자로 전분기 대비 소폭 감소할 것으로 전망한다.

Key data

Price performance



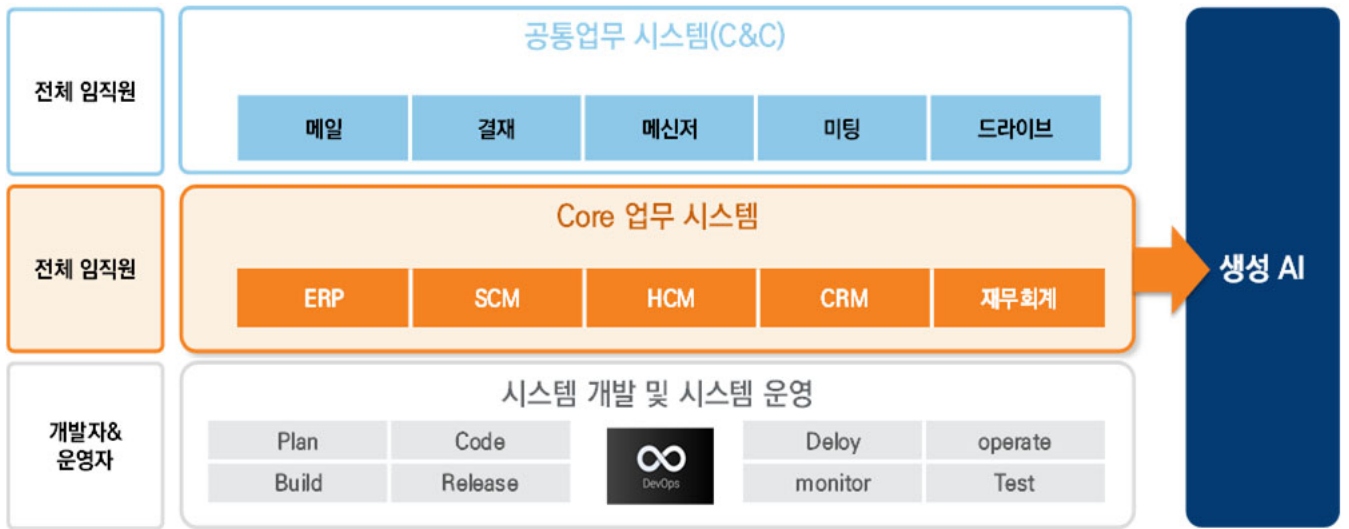
(%)	1M	6M	12M
절대주가	-1.7	13.8	16.1
상대주가	-0.7	10.1	6.5

Earnings and valuation metrics

계산기 (12월)	2020	2021	2022	2023F	2024F
매출액 (십억원)	11,017	13,630	17,235	13,527	14,760
영업이익 (십억원)	872	808	916	783	864
영업이익률 (%)	7.9	5.9	5.3	5.8	5.9
순이익 (십억원)	443	611	1,100	730	680
EPS (원)	5,731	7,899	14,213	9,437	8,788
ROE (%)	6.7	8.8	14.2	8.6	7.5
P/E (배)	31.1	19.8	8.7	14.5	15.6
P/B (배)	2.1	1.7	1.2	1.2	1.1
배당수익률 (%)	1.3	1.5	2.6	2.3	2.3

주: K-IFRS 연결 기준, 순이익은 지배주주 귀속 순이익
자료: 삼성SDS, 미래에셋증권 리서치센터

그림 229. 삼성SDS는 엔터프라이즈 솔루션 영역에 생성AI를 본격적으로 도입



자료: 삼성SDS, 미래에셋증권 리서치센터

표 37. 삼성SDS 연결 실적 추정치 및 시장 컨센서스 비교(연결 기준)

(십억원)

	3Q23				2023			
	당사 추정치	YoY	컨센서스	차이 %	당사 추정치	YoY	컨센서스	차이
매출액	3,371.2	-19.70%	3,431.9	-1.8%	13,527.2	-21.5%	13,771.3	-1.8%
영업이익	187.6	1.39%	209.2	-10.3%	782.7	-15.5%	834.2	-6.2%
지배순이익	173.0	-57%	183.8	-5.9%	730.0	-35.4%	741	-2.2%

자료: fnguide, 미래에셋증권 리서치센터

표 38. 삼성SDS 사업부별 수익 예상 변경(연결 기준)

(십억원)

(십억원/억원)	변경전			변경후			변경률		
	3Q23F	2023F	2024F	3Q23F	2023F	2024F	3Q23F	2023F	2024F
매출액	3,417.7	13,689.6	15,012.6	3,371.2	13,527.2	14,759.6	-1.4%	-1.2%	-1.7%
IT서비스	1,601.3	6,271.7	7,167.2	1,661.4	6,398.0	7,330.3	3.8%	2.0%	2.3%
클라우드	463.8	1,810.4	2,359.0	479.4	1,871.9	2,522.1	3.4%	3.4%	6.9%
MSP	226.6	913.7	1,269.7	241.4	963.1	1,391.6	6.5%	5.4%	9.6%
CSP	156.4	586.2	721.8	157.1	598.2	763.0	0.5%	2.1%	5.7%
SaaS	80.9	310.5	367.5	80.9	310.5	367.5	0.0%	0.0%	0.0%
SI	318.7	1,218.4	1,383.3	353.4	1,275.5	1,383.3	10.9%	4.7%	0.0%
ITO	818.7	3,243.0	3,424.9	828.6	3,250.6	3,424.9	1.2%	0.2%	0.0%
물류	1,816.4	7,418.3	7,845.4	1,709.7	7,129.6	7,429.3	-5.9%	-3.9%	-5.3%
영업이익	188.4	793.4	830.0	187.6	782.7	864.3	-0.5%	-1.3%	4.1%
EPS	1,886.8	8,668.0	8,581.6	2,085.9	8,913.2	8,524.4	10.6%	2.8%	-0.7%

자료: 미래에셋증권 리서치센터

삼성SDS (018260)

예상 포괄손익계산서 (요약)

(십억원)	2021	2022	2023F	2024F
매출액	13,630	17,235	13,527	14,760
매출원가	11,947	15,334	11,825	12,997
매출총이익	1,683	1,901	1,702	1,763
판매비와관리비	875	984	920	898
조정영업이익	808	916	783	864
영업이익	808	916	783	864
비영업손익	50	216	218	103
금융손익	43	82	136	137
관계기업등 투자손익	6	7	-27	-74
세전계속사업손익	858	1,132	1,001	967
계속사업법인세비용	224	2	271	294
계속사업이익	633	1,130	730	673
중단사업이익	0	0	0	0
당기순이익	633	1,130	730	673
지배주주	611	1,100	730	680
비지배주주	22	30	0	-7
총포괄이익	773	1,212	827	673
지배주주	745	1,180	808	657
비지배주주	28	31	19	15
EBITDA	1,250	1,428	1,071	950
FCF	694	698	-2,800	-4,650
EBITDA 마진율 (%)	9.2	8.3	7.9	6.4
영업이익률 (%)	5.9	5.3	5.8	5.9
지배주주귀속 순이익률 (%)	4.5	6.4	5.4	4.6

예상 재무상태표 (요약)

(십억원)	2021	2022	2023F	2024F
유동자산	7,576	8,006	6,591	7,354
현금 및 현금성자산	1,081	1,392	1,200	1,295
매출채권 및 기타채권	2,609	2,551	2,079	2,337
재고자산	44	41	33	37
기타유동자산	3,842	4,022	3,279	3,685
비유동자산	2,941	3,947	5,479	5,405
관계기업투자등	110	119	97	109
유형자산	1,179	1,520	2,536	2,474
무형자산	715	713	1,260	1,237
자산총계	10,517	11,952	12,070	12,759
유동부채	2,370	2,493	2,082	2,314
매입채무 및 기타채무	892	797	650	730
단기금융부채	149	193	207	207
기타유동부채	1,329	1,503	1,225	1,377
비유동부채	703	992	944	976
장기금융부채	313	676	687	687
기타비유동부채	390	316	257	289
부채총계	3,074	3,485	3,027	3,290
지배주주지분	7,229	8,223	8,802	9,235
자본금	39	39	39	39
자본잉여금	1,297	1,297	1,297	1,297
이익잉여금	6,087	7,001	7,484	7,916
비지배주주지분	215	244	242	234
자본총계	7,444	8,467	9,044	9,469

예상 현금흐름표 (요약)

(십억원)	2021	2022	2023F	2024F
영업활동으로 인한 현금흐름	980	1,290	889	722
당기순이익	633	1,130	730	673
비현금수익비용가감	757	464	438	241
유형자산감가상각비	387	462	261	62
무형자산감가상각비	55	51	27	23
기타	315	-49	150	156
영업활동으로인한자산및부채의변동	-266	-100	-337	-36
매출채권 및 기타채권의 감소(증가)	-377	107	258	-167
재고자산 감소(증가)	0	0	4	-4
매입채무 및 기타채무의 증가(감소)	241	-49	-160	68
법인세납부	-189	-272	-195	-294
투자활동으로 인한 현금흐름	-827	-645	-2,470	-4,739
유형자산처분(취득)	-286	-587	-3,186	-4,372
무형자산감소(증가)	-33	-60	-5	0
장단기금융자산의 감소(증가)	-515	-127	672	-367
기타투자활동	7	129	49	0
재무활동으로 인한 현금흐름	-323	-355	-49	-248
장단기금융부채의 증가(감소)	96	407	25	0
자본의 증가(감소)	0	0	0	0
배당금의 지급	-188	-188	0	-248
기타재무활동	-231	-574	-74	0
현금의 증가	-118	311	-192	95
기초현금	1,199	1,081	1,392	1,200
기말현금	1,081	1,392	1,200	1,295

자료: 삼성에스디에스, 미래에셋증권 리서치센터

예상 주당가치 및 valuation (요약)

	2021	2022	2023F	2024F
P/E (x)	19.8	8.7	14.5	15.6
P/CF (x)	8.7	6.0	9.1	11.6
P/B (x)	1.7	1.2	1.2	1.1
EV/EBITDA (x)	6.6	3.9	7.1	7.5
EPS (원)	7,899	14,213	9,437	8,788
CFPS (원)	17,975	20,600	15,096	11,809
BPS (원)	93,442	106,294	113,750	119,339
DPS (원)	2,400	3,200	3,200	3,200
배당성향 (%)	29.3	21.9	33.9	36.8
배당수익률 (%)	1.5	2.6	2.3	2.3
매출액증가율 (%)	23.7	26.4	-21.5	9.1
EBITDA증가율 (%)	-5.5	14.2	-25.0	-11.3
조정영업이익증가율 (%)	-7.3	13.4	-14.6	10.4
EPS증가율 (%)	37.8	79.9	-33.6	-6.9
매출채권 회전을 (회)	8.0	9.7	9.0	10.3
재고자산 회전을 (회)	389.2	404.6	364.4	416.9
매입채무 회전을 (회)	17.1	20.5	19.3	22.2
ROA (%)	6.4	10.1	6.1	5.4
ROE (%)	8.8	14.2	8.6	7.5
ROIC (%)	17.7	23.1	10.4	10.6
부채비율 (%)	41.3	41.2	33.5	34.8
유동비율 (%)	319.6	321.1	316.6	317.8
순차입금/자기자본 (%)	-55.5	-49.1	-36.2	-39.4
조정영업이익/금융비용 (x)	45.8	31.8	37.2	41.0



(Analyst) 박연주 yeonju.park@miraeasset.com



AIQ US · 생성 AI · 미국

Global X Artificial Intelligence & Technology

생성 AI, 대 전환의 시작

Report summary

주요 기업들의 생성 AI 수익화 시작: 응용 서비스 시장의 성장 속도는 시장 기대를 상회할 전망

글로벌 주요 기업들이 본격적으로 생성AI 기술을 제품화하기 시작했다. 빅테크부터 소프트웨어, 하드웨어 업체에 이르기까지 주요 기업들은 앞다퉀 생성 AI 전략을 발표하고 있고 그 설득력에 따라 주가도 움직이기 시작했다. 생성 AI를 통해 점유율을 확대시키거나 판매 가격을 인상시키고, 원가를 절감시키는 등 기업 가치를 변화시킬 가능성이 크기 때문이다. 23년 4분기~24년에는 관련 제품들이 쏟아질 것으로 예상된다.

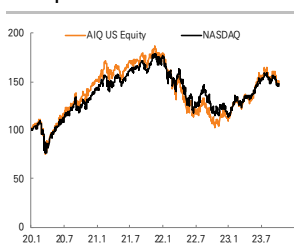
아직 생성 AI 기술이 완전하지 않지만 주요 기업들의 제품화를 통해 방대한 고객군과 적극적인 마케팅 능력, 저작권 등 법적 이슈에 대한 보완 등이 더해진다면 응용 서비스 시장의 성장 속도는 시장 기대를 상회할 가능성이 높다. 생성 AI 기술은 특히 기반 기술 자체의 발전 속도가 빨라 중장기적으로 잠재 시장 규모는 지금보다 훨씬 더 커질 수 있다.

개화하는 생성 AI 산업에 분산 투자할 수 있는 ETF

과거 인터넷, 모바일 혁명과 같은 기반 기술 혁명에서도 기술이 제품화되면서 본격적으로 부가가치를 창출했고 치열한 경쟁에서 살아남은 일부 과점 업체들이 그 수혜를 누렸다. 생성 AI 혁명에서 그 주인공이 누가 될지는 2~3년 내에 결정되겠지만, 분명한 것은 산업의 성장 속도는 시장 기대를 상회할 것이고 막대한 부가가 창출될 것이라는 점이다. 글로벌 AI 업체에 분산 투자할 수 있는 AIQ ETF에 관심을 가질 이유다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	-7.8	8.5	36.0
상대주가	-1.6	5.2	21.9

Earnings and valuation metrics

운용사	Global X	순자산 (백만USD)	570.37
설정일	2018-05-11	Total Expense Ratio (%)	0.68
수수료율 (%)	0.68	구성 종목 수	88

구성종목 상위 10개

순위	기업명	비중 (%)	순위	기업명	비중 (%)
1	알파벳	3.54	6	엔비디아	3.19
2	인텔	3.31	7	어도비	3.18
3	메타	3.30	8	아마존	3.16
4	시스코	3.28	9	액센츄어	3.10
5	인터내셔널 비즈니스 머신스	3.22	10	테슬라	3.07

주: K-IFRS 연결 기준, 순이익은 지배주주 귀속 순이익
자료: GlobalX, 미래에셋증권 리서치센터



(Analyst) 김수진 soojin.kim@miraeasset.com



BUG US · 글로벌 사이버보안 · 미국

Global X Cyber Security ETF

생성시 시대 투자 1순위는 우선 사이버보안이다

Report summary

생성시와 사이버보안은 동반 성장할 수밖에 없다

생성시 시대에 사이버보안 산업의 구조적 성장으로 관련 기업의 매출 확대를 예상한다. 생성시로 인해 해킹 공격도 늘어나고, 보안 및 대응 방안도 증가할 수 있다. 기업들의 디지털 의존도 역시 강화되고 있다. 그러므로 기업의 IT 투자가 사이버보안 기술로 몰리면서 산업의 성장이 필연적으로 연결되는 상황이다.

마이크로소프트 매출의 10%가 사이버보안에서 발생 - 그만큼 유망한 산업이라는 증거

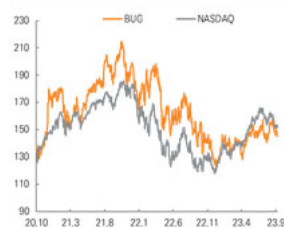
대표적으로 마이크로소프트가 보안 산업에 대대적으로 진출하며 산업의 구조적 성장이 동반될 것으로 기대한다. MS는 지난해 5월 XDR, 올해 SIEM 등 신규 보안 솔루션을 적극 발표하며 생성시를 기반으로 한 보안 솔루션 확대에 주력 중이다. 이는 경쟁사 등장이란 관점보다는 산업 확대를 입증하는 근거로 판단한다.

매출 50% 이상 사이버보안에서 발생하는 기업에 투자하는 ETF: 어려운 종목 투자 대신 ETF가 나은 선택

최근 고성장하는 보안 카테고리로는 클라우드 보안, ID 접근 보안, 데이터 프라이버시 보안, XDR, 관제 보안 등이 있다. 생성시로 클라우드 인프라 사용 비중이 더욱 커지는 상황이므로, 기업들의 투자 최우선순위는 계속해서 사이버보안과 생성시가 될 것으로 전망한다. 산업 전체가 확장하는 만큼, 어려운 보안 기업 투자보다 ETF 투자를 추천한다. BUG US는 매출의 50% 이상 사이버보안에서 발생하는 대표 보안 기업 중심으로 포트폴리오가 구성되어 있는 만큼 적절한 투자 대안으로 판단한다.

Key data

Price performance



(%)	1M	6M	12M
절대주가	4.2	8.8	3.6
상대주가	6.2	0.0	-11.4

Earnings and valuation metrics

운용사	Global X	순자산 (백만USD)	604.73
설정일	2019.10.25	Total Expense Ratio (%)	0.51%
수수료율 (%)	0.5%	구성 종목 수	23

구성종목 상위 10개

순위	기업명	비중 (%)	순위	기업명	비중 (%)
1	지스케일러	8.9%	6	퀵리스	5.4%
2	클라우드스트라이크	7.0%	7	다크트레이스	5.4%
3	팔로 알토 네트워크	6.5%	8	바로니스 시스템즈	5.2%
4	체크포인트 소프트웨어	6.0%	9	사이버아크 소프트웨어	5.1%
5	옥타	5.7%	10	테너블	5.0%

주: K-IFRS 연결 기준, 순이익은 지배주주 귀속 순이익
자료: GlobalX, 미래에셋증권 리서치센터

표 39. ETF 보유 종목 (23년 9월 22일 기준)

비중(%)	Ticker	Name	Market Price (\$)	Shares Held	Market Value (\$)
8.86	ZS	ZSCALER INC	151.71	353,336	53,604,604.56
7.06	CRWD	CROWDSTRIKE HO-A	162.57	262,714	42,709,414.98
6.53	PANW	PALO ALTO NETWORKS INC	228.51	172,850	39,497,953.50
5.95	CHKP	CHECK POINT SOFTWARE TECH	137.4	261,747	35,964,037.80
5.69	OKTA	OKTA INC	80.68	426,240	34,389,043.20
5.44	QLYS	QUALYS INC	153.12	214,863	32,899,822.56
5.39	DARK LN	DARKTRACE PLC	5.15	6,334,754	32,594,896.33
5.23	VRNS	VARONIS SYSTEMS INC	30.54	1,034,920	31,606,456.80
5.12	CYBR	CYBERARK SOFTWARE	167.67	184,748	30,976,697.16
5.04	TENB	TENABLE HOLDINGS INC	45.66	667,775	30,490,606.50
4.88	FTNT	FORTINET INC	58.24	506,350	29,489,824.00
4.54	BB	BLACKBERRY LTD	5.16	5,315,744	27,429,239.04
4.4	GEN	GEN DIGITAL INC	18.86	1,411,613	26,623,021.18
4.1	ATEN	A10 NETWORKS INC	14.93	1,661,262	24,802,641.66
3.95	RPD	RAPID7 INC	46.85	509,372	23,864,078.20
3.83	S	SENTINELONE INC -CLASS A	15.8	1,466,121	23,164,711.80
3.46	RDWR	RADWARE LTD	16.32	1,282,012	20,922,435.84
2.98	4704 JP	TREND MICRO INC	38.57	466,919	18,007,052.77
2.5	053800 KS	AHNLAB INC	47.2	319,799	15,095,240.05
2.35	OSPN	ONESPAN INC	11.09	1,278,622	14,179,917.98
2.26	2326 JP	DIGITAL ARTS INC	31.06	440,603	13,686,557.04
0.3	TLS	TELOS CORPORATION	1.93	953,782	1,840,799.26
0.12	ARQQ	ARQIT QUANTUM INC	0.53	1,380,794	734,444.33

자료: Global X, 미래에셋증권 리서치센터

표 40. 마이크로소프트 사이버보안 제품

제품명	기능	대표 peer	대형 보안 기업 출시 제품
Microsoft Security Copilot	생성AI로 위협 발견, 조사 및 대응	클라우드스트라이크(CRWD US) 외 등장 중	
Microsoft Sentinel	SIEM	Splunk(SPLK US), Elastic, Rapid7	아마존, 포티넷, IBM, 포티넷, 팔로알토네트웍스, 오라클
Microsoft Defender	XDR	클라우드스트라이크(CRWD US), 센티널원(S US), 트렌드마이크로(4704 JP)	포티넷, 시만텍
Microsoft Intune	모바일 디바이스 관리 보안	Jamf(JAMF US)	VM웨어, 시스코, Sophos, 시트릭스, AWS, IBM
Microsoft Entra	ID 보안	OKTA(OKTA US)	포티넷, 아마존, 구글, IBM
Microsoft Purview	데이터/거버넌스 보안	informatica(INFA US)	IBM, SAP

자료: 마이크로소프트, 미래에셋증권 리서치센터

표 41. 사이버보안 산업 세부 카테고리별 성장률

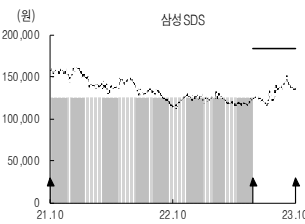
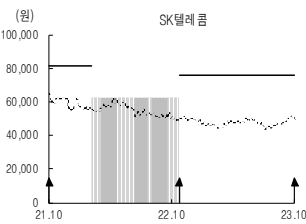
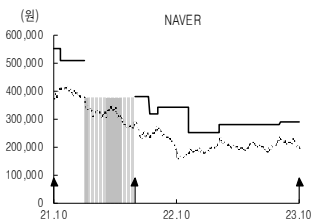
세부 시장	2021	2022	2023	2024	2025	2026	2027	CAGR 2022 ~2027
Application Security	4,550.4	5,047.6	5,765.2	6,670.3	7,632.8	8,642.1	9,619.6	13.8%
YoY growth(%)		10.9%	14.2%	15.7%	14.4%	13.2%	11.3%	
전체 대비 비중(%)	3.1%	3.1%	3.1%	3.1%	3.2%	3.3%	3.3%	
Cloud Security	3,619.2	4,487.4	5,616.7	7,002.6	8,656.3	10,625.9	12,811.1	23.3%
YoY growth(%)		24.0%	25.2%	24.7%	23.6%	22.8%	20.6%	
전체 대비 비중(%)	2.4%	2.7%	3.0%	3.3%	3.6%	4.0%	4.4%	
Data Privacy	1,027.9	1,129.2	1,338.7	1,667.3	1,970.8	2,300.8	2,667.5	18.8%
YoY growth(%)		9.9%	18.5%	24.6%	18.2%	16.7%	15.9%	
전체 대비 비중(%)	0.7%	0.7%	0.7%	0.8%	0.8%	0.9%	0.9%	
Data Security	2,532.2	3,072.9	3,692.1	4,333.3	4,989.8	5,596.5	6,171.2	15.0%
YoY growth(%)		21.4%	20.1%	17.4%	15.2%	12.2%	10.3%	
전체 대비 비중(%)	1.7%	1.9%	2.0%	2.0%	2.1%	2.1%	2.1%	
Identity Access Management	12,271.5	13,944.1	16,169.1	18,556.5	20,753.3	22,781.5	24,863.4	12.3%
YoY growth(%)		13.6%	16.0%	14.8%	11.8%	9.8%	9.1%	
전체 대비 비중(%)	8.2%	8.5%	8.6%	8.6%	8.6%	8.6%	8.6%	
Infrastructure Protection	20,089.7	24,089.0	28,359.6	33,319.6	38,172.1	42,832.4	47,396.2	14.5%
YoY growth(%)		19.9%	17.7%	17.5%	14.6%	12.2%	10.7%	
전체 대비 비중(%)	13.5%	14.6%	15.1%	15.5%	15.9%	16.2%	16.4%	
Integrated Risk Management	4,706.4	5,157.3	5,687.1	6,277.7	6,696.8	7,052.4	7,392.6	7.5%
YoY growth(%)		9.6%	10.3%	10.4%	6.7%	5.3%	4.8%	
전체 대비 비중(%)	3.2%	3.1%	3.0%	2.9%	2.8%	2.7%	2.6%	
Network Security Equipment	16,924.4	18,932.5	21,383.6	24,360.1	27,285.6	30,235.1	33,103.7	11.8%
YoY growth(%)		11.9%	12.9%	13.9%	12.0%	10.8%	9.5%	
전체 대비 비중(%)	11.4%	11.5%	11.4%	11.3%	11.3%	11.4%	11.5%	
Other Information Security Software	5,349.1	8,029.8	11,365.4	14,362.8	16,520.4	17,502.5	18,456.7	18.1%
YoY growth(%)		50.1%	41.5%	26.4%	15.0%	5.9%	5.5%	
전체 대비 비중(%)	3.6%	4.9%	6.0%	6.7%	6.9%	6.6%	6.4%	
Security Services	70,641.9	73,394.7	80,835.7	89,996.7	98,884.9	107,717.7	116,255.8	9.6%
YoY growth(%)		3.9%	10.1%	11.3%	9.9%	8.9%	7.9%	
전체 대비 비중(%)	47.4%	44.6%	43.0%	41.9%	41.1%	40.7%	40.3%	
Consumer Security Software	7,234.7	7,443.4	7,901.7	8,406.7	8,847.9	9,300.9	9,759.4	5.6%
YoY growth(%)		2.9%	6.2%	6.4%	5.2%	5.1%	4.9%	
전체 대비 비중(%)	4.9%	4.5%	4.2%	3.9%	3.7%	3.5%	3.4%	
Total	148,947.4	164,728.0	188,114.8	214,953.7	240,410.7	264,587.7	288,497.2	11.9%
YoY growth(%)		10.6%	14.2%	14.3%	11.8%	10.1%	9.0%	

자료: 가트너(2023.7), 미래에셋증권 리서치센터

투자 의견 및 목표주가 변동 추이

제시일자	투자 의견	목표주가(원)	과리율(%)		제시일자	투자 의견	목표주가(원)	과리율(%)	
			평균주가대비	최고(최저)주가대비				평균주가대비	최고(최저)주가대비
NAVER (035420)					2021.07.07				
2023.08.07	매수	290,000	-	-	매수	550,000	-23.41	-17.45	
2023.02.06	매수	280,000	-27.39	-16.79	SK텔레콤 (017670)				
2022.11.08	매수	250,000	-24.17	-10.60	2022.10.25	매수	76,000	-	-
2022.08.08	매수	340,000	-39.33	-21.32	2022.02.07	분석 대상 제외	-	-	-
2022.07.14	매수	320,000	-21.09	-14.53	2021.07.15	매수	82,000	-26.90	-20.24
2022.05.31	매수	380,000	-34.29	-23.68	삼성SDS (018260)				
2022.01.03	분석 대상 제외	-	-	-	2023.05.31	매수	185,000	-	-
2021.10.22	매수	510,000	-22.15	-18.14	2021.03.22	분석 대상 제외	-	-	-

* 과리율 산정: 수정주가 적용, 목표주가 대상시점은 1년이며 목표주가를 변경하는 경우 해당 조사분석자료의 공표일 전일까지 기간을 대상으로 함



투자 의견 분류 및 적용 기준

기업	산업
매수 : 향후 12개월 기준 절대수익률 20% 이상의 초과수익 예상	비중확대 : 향후 12개월 기준 업종지수상승률이 시장수익률 대비 높거나 상승
Trading Buy : 향후 12개월 기준 절대수익률 10% 이상의 초과수익 예상	중립 : 향후 12개월 기준 업종지수상승률이 시장수익률 수준
중립 : 향후 12개월 기준 절대수익률 -10~10% 이내의 등락이 예상	비중축소 : 향후 12개월 기준 업종지수상승률이 시장수익률 대비 낮거나 약화
매도 : 향후 12개월 기준 절대수익률 -10% 이상의 주가하락이 예상	

매수(▲), Trading Buy(■), 중립(●), 매도(◆), 주가(—), 목표주가(→), Not covered(■)

투자 의견 비율

매수(매수)	Trading Buy(매수)	중립(중립)	매도
89.09%	7.88%	2.42%	0.61%

* 2023년 06월 30일 기준으로 최근 1년간 금융투자상품에 대하여 공표한 최근일 투자등급의 비율

Compliance Notice

- 당사는 자료 작성일 현재 NAVER 발행주식총수의 1% 이상을 보유하고 있습니다.
- 당사는 자료 작성일 현재 SK텔레콤, 삼성SDS, NAVER 을(를) 기초자산으로 하는 주식워런트증권에 대해 유동성공급자(LP)업무를 수행하고 있습니다.
- 당사의 애널리스트는 자료 작성일 기준으로 1개월 이내에 SK텔레콤 의 기업설명회 등에 해당 법인의 비용으로 참석한 사실이 있습니다.
- 당사는 본 자료를 제3자에게 사전 제공한 사실이 없습니다.
- 본 자료를 작성한 애널리스트 임희석은(는) 자료작성일 현재 유니티 2000주 보유하고 있습니다.
- 본 자료는 외부의 부당한 압력이나 간섭없이 애널리스트의 의견이 정확하게 반영되었음을 확인합니다.

본 조사분석자료는 당사의 리서치센터가 신뢰할 수 있는 자료 및 정보로부터 얻은 것이나, 당사가 그 정확성이나 완전성을 보장할 수 없으므로 투자자 자신의 판단과 책임하에 종목 선택이나 투자시기에 대한 최종 결정을 하시기 바랍니다. 따라서 본 조사분석자료는 어떠한 경우에도 고객의 증권투자 결과에 대한 법적 책임소재의 증빙자료로 사용될 수 없습니다. 본 조사분석자료의 지적재산권은 당사에 있으므로 당사의 허락 없이 무단 복제 및 배포할 수 없습니다.