Strategy Idea

글로벌 투자전략 - 선진국
Analyst **황수욱**
02. 6454-4896
soowook.hwang@meritz.co.kr

AI 변곡점에서 투자 이정표

- ✓ 딥시크 모먼트: 고성능 AI LLM의 상품화. 여기서 비롯된 AI 가격 하락
- ✓ 상품화된 고성능 모델 기반의 AI S/W 응용 확장 → AI 금융, AI 헬스케어 시장 주목
- ✓ AI 모델 속도 경쟁 → AI H/W: 저전력 추론 연산 프로세서(ASIC), 네트워크 시스템 수혜 예상

딥시크 모먼트: AI LLM의 '상품화'

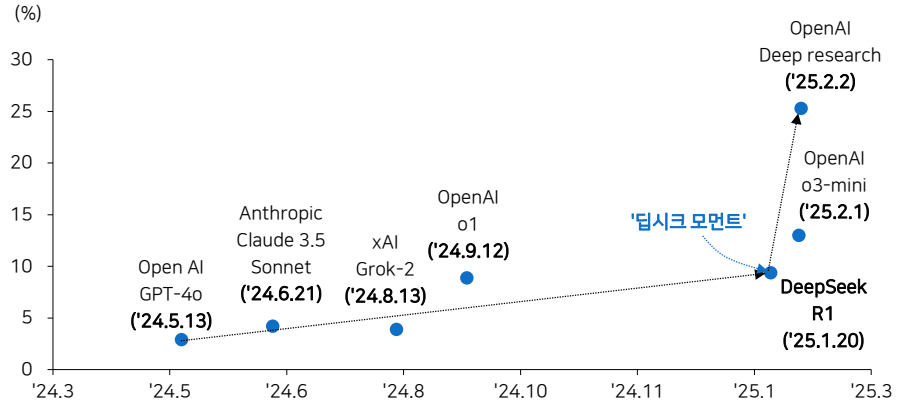
딥시크 R1에 출시에 기민하게 대응하는 OpenAI

딥시크 충격 이후 가장 바쁜 기업은 OpenAI로 보인다. 시장이 딥시크 R1을 인지한 1월 27일 이후 1주일도 안되어 OpenAI는 새 AI 모델 o3-mini를 공개했다. 여기서 멈추지 않고 더 정교하고 뛰어난 성능을 가진 Deep research를 바로 출시하면서 AI 모델 경쟁에 박차를 가했다.

덕분에 AI 모델 성능 향상 속도는 변곡점을 맞이함

딥시크 출현 이전과 이후, 새로 공개되는 모델의 기능 향상 속도가 다르다. AI를 평가하는 모델 중 가장 어려운 것으로 알려져 있는 'Humanity's Last Exam(인류 최후의 시험)' 성적 기준으로 보면, 딥시크 이전까지 AI 모델이 등장한 이후 가장 진보된 모델이라는 o1의 정답률이 10%가 되지 않았다. 딥시크가 공개되면서 새로 공개된 모델로 인해 불과 4개월만에 이 시험에 대한 최신 AI 모델의 정답률은 25%까지 급상승했다.

그림1 AI 모델 'Humanity's Last Exam' 성적: 딥시크 출시 이후 고성능 모델 빠르게 공개



주: 1) Text-only results, Deep research는 web browsing + python tools 사용
2) 각 모델의 첫 출시일 기준으로 작성; 모델 버전의 업데이트 날짜와 다를 수 있음
자료: Center for AI Safety, 'Humanity's Last Exam(2025)', 메리츠증권 리서치센터

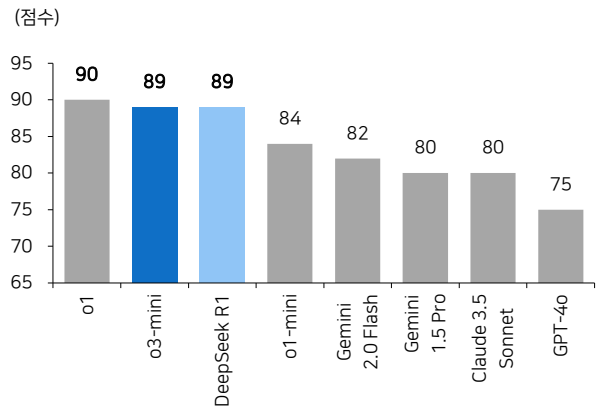
고성능+가격 하락 경쟁. 높은 유용성 레벨에서 상품화된 AI

팔란티어 CTO 샴 산카르는 이번 실적 발표회에서 딥시크 R1의 명백한 교훈 중 하나는 AI 모델이 상품화(Commodity, 대체 가능하고 공급자 관계 없이 동일한 가치와 특성을 가지는 제품을 의미, 일반화되고 대중이 쉽게 이용 가능해진 상태) 되고 있는 것이라고 말했다. Generative AI 단계에서의 모델들의 상품화는 계속 진행 중이었지만, R1 모델이 이런 흐름을 대중에게 각인시킨 효과가 큰 것이다.

'이 가격에 이정도 성능을?' 딥시크 충격이 컸던 이유는 '꽤나 쓸모 있는 수준'까지 발전한 AI 성능도 한 몫 했을 것

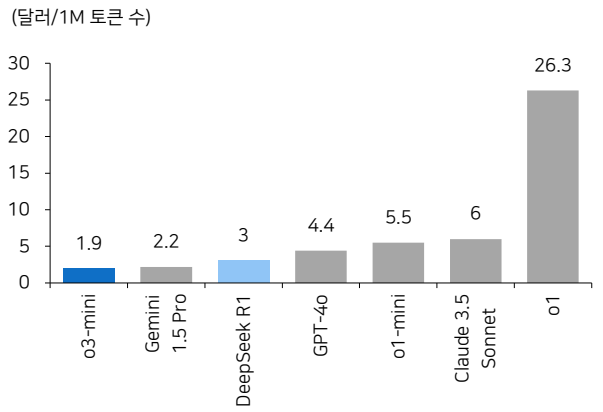
R1의 각인 효과가 컸던 이유는 낮은 비용도 중요했지만 모델의 '실력'도 중요했던 것으로 생각한다. AI 모델이 수준이 '꽤나 쓸모 있는' 수준에 도달하기 시작했기 때문이다. 새로 공개된 o3-mini 모델도 딥시크보다 저렴한 API 비용이 필요한, 박사 수준의 전문가 지식을 갖춘 모델이다(GPQA 벤치마크 점수 기준). o1 모델이 유사한 성능이었지만, 비용(그림 3)을 고려하면 딥시크의 R1과 o3-mini는 낮은 비용으로 쓸모 있는 높은 수준의 성능을 갖춘 최초의 모델인 것이다. 코딩 작업에서의 AI의 효용성도 2년 사이 급격한 발전을 이루었다.

그림2 Artificial Analysis Quality Score: 높을수록 좋음



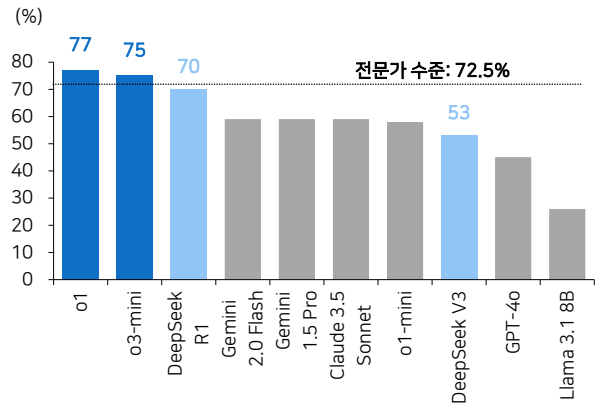
자료: Artificial Analysis, 메리츠증권 리서치센터

그림3 AI 모델 토큰 가격: 낮을수록 좋음



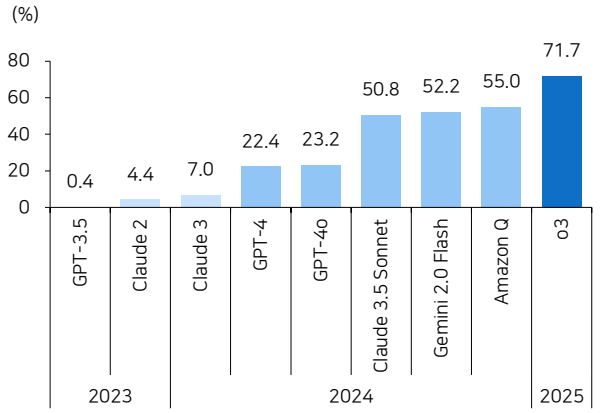
자료: Artificial Analysis, 메리츠증권 리서치센터

그림4 GPQA(전문지식평가): 전문가 수준에 도달한 AI 모델



주: GPQA: Graduate-Level Google-Proof. 박사학위 레벨 전문가는 해당 테스트에서 평균 65% 정확도를, 비전문가는 34% 정확도 수준을 보임
자료: Rein et al.(2023), Artificial Analysis, 메리츠증권 리서치센터

그림5 SWE-bench(verified): 복잡한 코딩 작성 기여도



주: SWE-bench 복잡한 코딩 작성 능력 수준을 평가하는 벤치마크
자료: SWE-bench Leaderboard, ARK Investment, 메리츠증권 리서치센터

AI 인프라 → 소프트웨어 중심으로 AI usage 포트폴리오 전환 가속

고성능 AI의 범용화 → AI usage로 포트 전환 가속

우리는 작년 11월부터 2025년에는 AI 인프라 포트폴리오(표 1)에서 AI usage 포트폴리오(표 2)로의 전환 타이밍이 핵심이라고 주장하고 있다. 1월 중순 트럼프 정책이 나올 때까지만 하더라도 AI 인프라 포트폴리오 흐름이 강화되는 듯 했으나, 딥시크 모먼트로 포트폴리오의 전환 타이밍이 가속화된 것으로 판단한다.

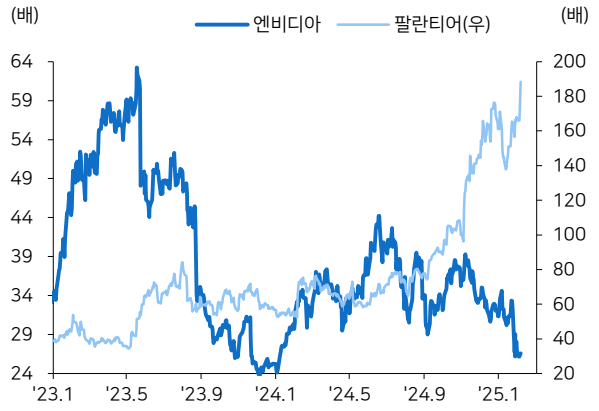
AI 성능 체험이 늘수록 AI B2C 인기는 커질 것

딥시크 모먼트로 AI 성능을 체험한 사람들을 중심으로 AI B2C 확산은 빨라질 것이고 AI usage(AI의 활용; 소프트웨어/자율주행/로봇) 관련 종목들의 '인기'도 계속 커질 것으로 생각한다.

우상향 고밸류 주식 = 주도주

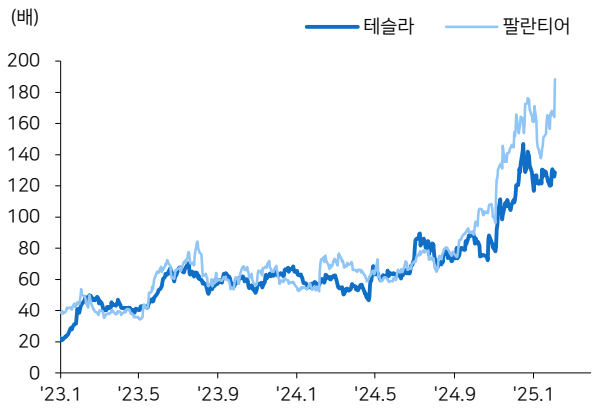
엔비디아는 2년간 시장을 이끈 주도주였는데, 실적의 고속 성장에도 밸류에이션은 IT 섹터 평균까지 하락했다. 반면, 팔란티어는 테슬라처럼 고밸류에이션을 계속 영위하며 주가가 급등 중이다. 너무 높은 밸류에 우려의 시선도 많지만 고밸류를 계속 정당화하며 주가 상승 흐름이 이어진, 최근 5년 사이 주도주였던 테슬라, 엔비디아의 흐름을 상기해봐야 한다.

그림6 엔비디아 vs 팔란티어 12MF PER: 인기 vs 비인기?



자료: Bloomberg, 메리츠증권 리서치센터

그림7 테슬라 vs 팔란티어 12MF PER: 인기 주식의 공통점?



자료: Bloomberg, 메리츠증권 리서치센터

표1 2024년 관심가졌던 AI 인프라 포트폴리오

기업명 (%)	YTD 수익률 (2/4)	매출 성장률		GP 마진		ROE
		2025E	2026E	2025E	2026E	최근 분기
엔비디아	-11.7	53.8	-	73.5	-	127.2
컨스텔레이션에너지	36.9	-4.7	-	48.2	-	23.6
GE 버노바	8.4	5.1	8.6	20.6	22.9	18.3
버티브	-2.1	16.6	-	38.0	-	33.0
마이크로소프트	-2.2	13.8	-	68.2	-	34.3
아마존	10.3	10.7	-	50.2	-	22.6
구글	0.7	-0.4	11.6	63.6	63.9	32.9

자료: Bloomberg, 메리츠증권 리서치센터

표2 2025년 전망에서 제시한 AI usage 포트폴리오('24.11)

기업명 (%)	YTD 수익률 (2/4)	매출 성장률		GP 마진		ROE
		2025E	2026E	2025E	2026E	최근 분기
애플	-7.0	4.9	7.8	46.9	47.1	136.5
테슬라	-2.9	16.6	21.1	18.7	19.8	10.5
브로드컴	-4.1	18.6	16.4	77.5	76.9	12.9
팔란티어	37.3	29.2	27.3	83.1	83.3	10.9
서비스나우	-4.7	18.8	19.6	81.6	81.7	16.5
오라클	0.8	12.7	-	71.8	-	132.0
어도비	-1.0	9.3	10.2	89.8	89.7	36.3

자료: Bloomberg, 메리츠증권 리서치센터

AI 인프라는 독점 구도 → AI 모델의 상품화에서는 수혜 종목의 확산 구도

팔란티어 Peer인 서비스나우는 주가 차별화 되었으나

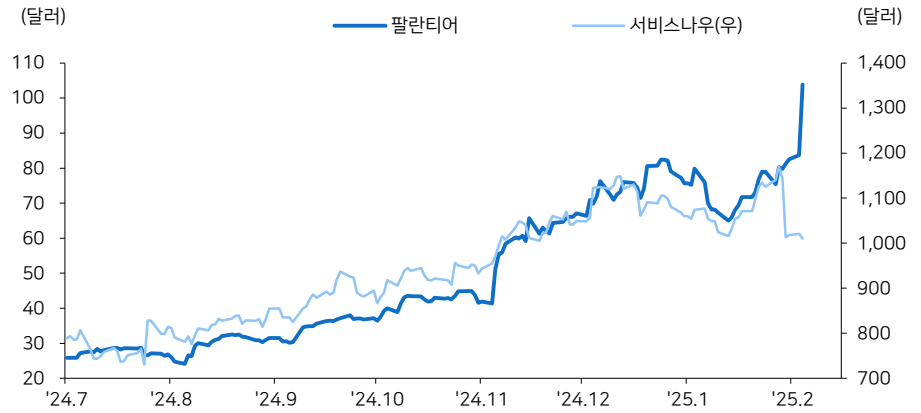
우리가 포트폴리오에 주목했던 종목 중 팔란티어와 서비스나우는 4Q24 실적발표 이후 주가 흐름이 차별화되었다. 팔란티어는 시장 기대를 상회하는 실적과 긍정적이고 자신감 있는 가이던스로 주가 상승이 가팔랐으나, 서비스나우는 실적 미스에 따른 주가 하락이 나타났다.

AI의 최적 적용을 기반으로 하는 팔란티어는 지속적인 고객사 증가 속도에 기반해 주가 상승 가도를 이어가고 있다. 반면 서비스나우는 4Q24 구독 매출 컨센서스 하회, 연간 구독 성장률 전망치도 시장 눈높이를 하회하면서 주가가 하락했다.

AI 상품화에 따른 가격 하락 수혜 확산시 기회 고려

다만 컨콜에 따르면 AI 비용 하락은 서비스나우의 마진 개선 요인이다. 여기에 해당 회사의 고객 기업들이 자체적으로 하드웨어에 투자하기보다 저렴해진 소프트웨어를 구매(서비스나우 소프트웨어 구매)하게 되는 변화까지 기대해볼 수 있다.

그림8 4Q24 실적발표 이후 차별화된 AI S/W 대표기업 팔란티어 vs 서비스나우 주가



자료: Bloomberg, 메리츠증권 리서치센터

표3 팔란티어와 서비스나우 4Q24 실적 컨퍼런스콜 주요 내용 정리

	팔란티어	서비스나우
주요 코멘트	▪ 현재까지 AI 시장이 공급 측면에 집중해왔기 때문에 AI를 생산에 적용하는데 한계를 겪는 기관들이 발생. 이러한 상황에서 팔란티어는 AI 수요를 충족시킬 수 있는 입지를 다져왔고, AIP는 이를 염두에 두고 구축된 것 → AI 수요 측면에서의 "missing piece"를 제공 중 ▪ AI 수요 관련 서비스의 핵심은 Ontology, AI를 통해 생산성을 개선하기 위해서는 AI와 기업체가 상호작용할 수 있게끔 하는 중간 매개체가 있어야 함. 그리고 팔란티어의 Ontology가 엔터프라이즈 데이터를 구조화함으로써 중재자 역할을 하고 있는 것	▪ AI 성과: Agentic AI +150% YoY 달성, 기업 오너들은 레거시 아키텍처를 경량화하고 싶어함. 그 점을 활용해서 Now Assistant는 1,000개 이상의 고객을 확보 ▪ 신규 가격 산정 방식: AI 에이전트가 가치 창출의 동인이 되고 있는 만큼 consumption-based pricing 도입을 추진하는 중. Agentic AI는 기존 모델보다 몇배의 쿼리를 소모하게 될 것
Q & A	**DeepSeek의 영향? Palantir는 어떻게 대응할 것인지?** ▪ DeepSeek의 시사점은 AI 모델이 상품화되고 있다는 것, 추론 가격이 떨어졌다는 것. 진짜 가치는 AI 모델 자체에 있는 것이 아니라, 그걸 어떻게 기업 안에서 작동하게 할 것인지 ▪ 중국과의 AI 군비 경쟁이 이루어지고 있다는 것, winner-take-all 구조이기 때문에 더욱 경쟁적으로 바라봐야 함 **정부의 변화 관련해서 팔란티어가 받을 영향은?** ▪ 이득일 것. 정부의 문제는 책임소재가 불분명한 여러 IT 프로젝트인데, 팔란티어가 나서서 정부 효율화에 도움이 될 수 있음	**DeepSeek 영향은?** ▪ 서비스나우에게는 좋은 소식. 모델들이 빠른 속도로 상품화 될 경우 마진 향상의 결과로 이어지게 될 것. 고객들이 하드웨어, 인프라, LLM 등에 투자할 비용을 줄이는 대신 소프트웨어에 투자할 여력이 생길 것 **글로벌 IT 트렌드 차원에서 기업들의 지출이 여전히 늘어나고 있는지?** ▪ 이전과 크게 다르지 않음. 다만 AI에 대한 단순 관심에서 비즈니스 모델 혁신으로 기업들의 포커스가 이동하고 있으며 기업들은 AI 후발주자가 될 것을 두려워하고 있음. 따라서 Agentic AI에 대한 수요가 기업의 지갑을 열게 할 것으로 예상

자료: 메리츠증권 리서치센터 정리

AI 상품화, 비용 절감 수혜 산업의 확장: AI 금융, 헬스케어

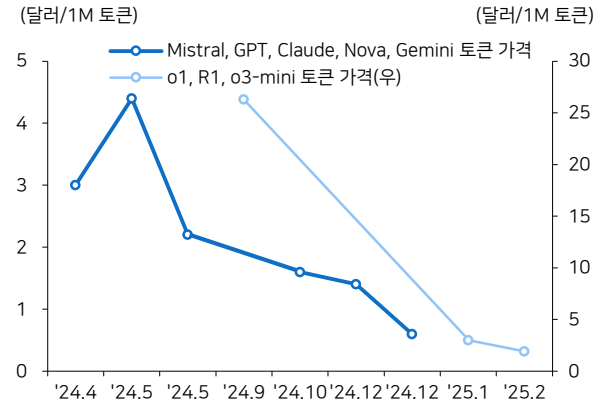
누적된 AI 하드웨어 투자가 기여한 AI 상품화

새로운 AI 모델의 기술 진보가 기본적인 가격 경쟁의 밑바탕이며, 여기에 AI 하드웨어 투자의 누적도 AI 추론 가격 하락에 영향을 주었을 것이다. 핵심 AI 연산 프로세서인 H100의 시간당 임대료는 2024년 피크 시점 8달러에서 최근 1.3달러까지 1/6 수준으로 하락했다.

AI 상품화 → 비용절감 수혜, AI 금융/헬스케어 주목

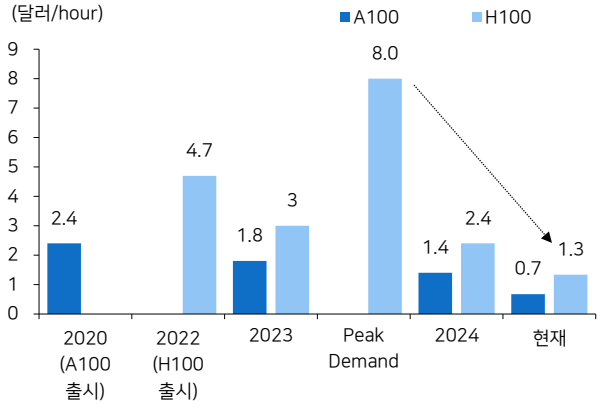
AI 추론 가격 하락, AI의 상품화 국면에서 아이디어의 확장은 AI 파운데이션 모델을 활용해 특정 분야에 응용하는 AI 어플리케이션 분야까지 수혜 확장을 기대할 수 있다. 대표적인 분야로 AI 금융 결제 시스템, 바이오/헬스케어 분야를 주목해볼 수 있다. AI 기술을 응용해 정밀 의학 솔루션을 제공하는 Tempus AI는 펠로시 전 하원의장의 거래 내역이 공개되며 1월에 주가가 급등하기도 했다.

그림9 AI 모델 출시를 거듭할수록 낮아지는 API 토큰 가격



자료: Artificial Analysis, 메리츠증권 리서치센터

그림10 A100, H100 시간당 임대료 변화



자료: Gurufocus, The Medium, Vast AI, 메리츠증권 리서치센터

표4 AI 소프트웨어 어플리케이션 가능성 (1): 금융

종목명	티커	시가총액 (십억달러)	수익률 (%) 1W	1M	1Y
비자	V	695.0	3.4	9.8	25.7
마스터카드	MA	517.7	3.0	8.3	23.1
인튜이트	INTU	168.2	-1.0	-4.5	-5.5
페어 아이작	FICO	44.7	-0.8	-7.0	45.6
소파이	SOFI	16.7	-4.2	3.8	95.4
업스타트	UPST	5.8	-2.8	-1.4	93.2
렌딩클럽	LC	1.6	-22.1	-26.0	43.4

자료: Bloomberg, 메리츠증권 리서치센터

표5 AI 소프트웨어 응용 가능성 (2): 바이오/헬스케어

종목명	티커	시가총액 (십억달러)	수익률 (%) 1W	1M	1Y
일루미나	ILMN	20.8	-5.0	-3.6	-6.5
템퍼스 AI	TEM	9.7	18.7	60.9	-
가튼트 헬스	GH	5.9	1.3	40.9	114.4
라드넷	RDNT	4.9	5.2	-6.9	76.8
리커전 파마수티컬스	RXRX	2.8	3.1	-7.1	-22.0
슈뢰딩거	SDGR	1.8	2.7	18.4	-6.7
텔레닥 헬스	TDOC	1.7	-1.6	5.6	-49.2

주: Tempus AI의 상장일은 2024.6.14
자료: Bloomberg, 메리츠증권 리서치센터

딥시크 모먼트에서 AI H/W의 기회

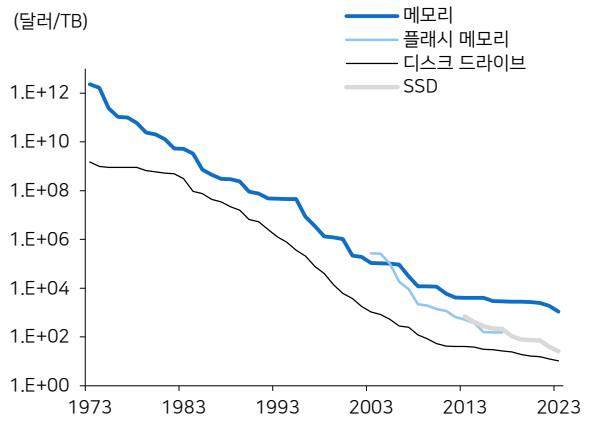
큰 틀에서는 AI S/W 확산은 AI H/W의 확산도 유발할 것

주식시장의 관심이 AI 소프트웨어로 집중되는 상황에서 AI 하드웨어의 기회도 재점검해야 한다. 제본스의 역설은 기술 발전이 자원 활용 효율성을 높이면 오히려 자원 소비가 증가함을 의미한다. 지금 AI도 마찬가지로 AI 추론 가격 하락에 따른 AI 하드웨어 활용의 효율성이 높아지면, AI 하드웨어 활용량의 증가로 이어질 것이다. 직전 거대 기술 사이클이었던 PC 사이클에서는 PC 가격의 하락이 보급을 가속화시키며 산업의 전체 규모를 키웠다.

AI ASIC, 네트워크 시스템 스토리 가속화 가능성 고려

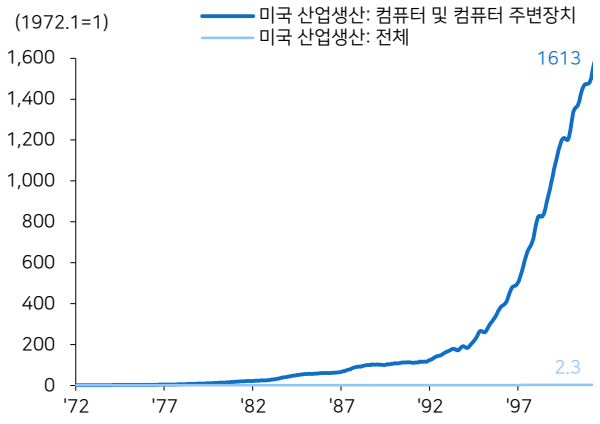
시장의 관심이 현재는 AI 소프트웨어에 집중되고 있지만, 큰 틀에서는 AI 하드웨어의 동반성장도 염두에 두어야 한다. 딥시크 모먼트로 인해 주식의 인기 관점에서 2023년 초부터 시장을 주도했던 AI GPU(엔비디아), AI CSP(구글, 마이크로소프트, 아마존)의 독점 프리미엄은 희석될 수 있지만, 그 다음 이어진 AI 하드웨어 사이클(ASIC, 네트워크 시스템)은 AI 경쟁을 위한 Capex 투자 가운데 모멘텀이 가속화될 가능성이 있다.

그림11 PC 사이클: exponential 비용 절감은 기술진보 과정



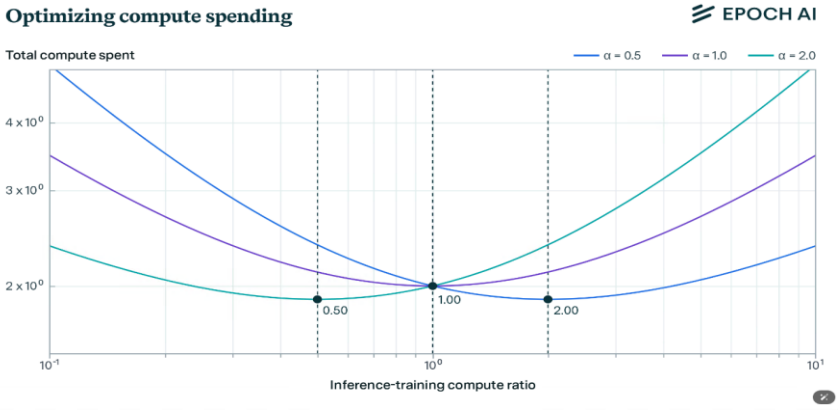
자료: Our World in Data, 메리츠증권 리서치센터

그림12 비용 절감은 기술 확산을 가속시키는 요인일 뿐



자료: Federal Reserve, 메리츠증권 리서치센터

그림13 Epoch AI: AI 최적 컴퓨팅 비율은 훈련과 추론에 5:5로 할당하는 것



주: AI 훈련과 추론에 할당하는 컴퓨팅 비율이 5대5가 최적임을 나타내는 그림
자료: Epoch AI, "Optimally Allocating Compute Between Inference and Training" ('24.3.29)

Scaling Law:
학습 단계의 더많은 컴퓨팅 자원
→ 더 나은 모델
Time Scaling Law:
추론 단계의 더많은 시간
→ 더 나은 결과

지금까지의 AI 모델 성능 향상 방정식인 Scaling Law에서는 1) 학습 단계에서 2) 더 많은 AI 학습 자원(모델 파라미터 수, 학습 데이터, 연산 자원(컴퓨팅 프로세서))이 확보될수록 더 강력한 AI 모델을 만들 수 있다. 이와 대비되는 Time Scaling Law는 1) 추론 단계에서, 2) 더 많은 계산 시간을 투입하면 더 나은 결과를 도출할 수 있다는 법칙이다. 이 기술이 적용된 모델은 생각하는 과정을 거쳐 적게는 수 초, 길게는 수십 초간 '생각'을 하다가 답을 내놓는다. Time Scaling Law 하에서는 AI 훈련 비용은 줄일 수 있지만, 추론에 필요한 인프라와 비용이 더 필요하게 된다.

o3-mini도 추론 과정이 있음, 최신 모델 경쟁 트렌드는 Test Time Scaling이 전제된 것으로 추측

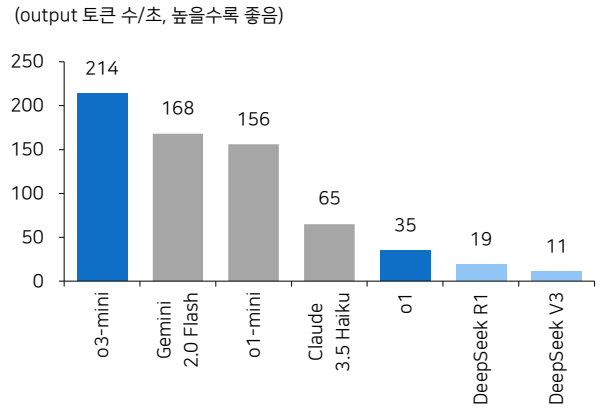
더 적은 비용으로 R1과 유사한 성능을 발휘하는 OpenAI의 새 모델 o3-mini도 추론 과정을 갖는다. 즉, 최신 모델의 경쟁 트렌드는 Test Time Scaling이 전제되어 있는 것으로 보인다. o3-mini의 경우 R1보다 더 빠른 결과 산출 속도, 더 낮은 레이턴시를 보여주는데, o3-mini 기반 모델(아마 o3일 것)의 높은 성능이 배경이 되어 추론 연산량을 낮춰도 기존 모델들의 성능과 유사하기 때문에 더 빠른 속도로 경쟁할 수 있는 것으로 해석된다. 이는 기반 모델 성능의 중요성과 distillation의 유용성을 의미한다.

표6 Scaling Law와 Time Scaling Law 비교

구분	전통적인 Scaling Law	Time Scaling Law(Test Time Scaling)
초점	- **훈련(Training) 과정** - 모델 크기, 학습 데이터, 연산 자원의 양	- **추론(Inference) 과정** - 예측/생성시 연산 자원 - 시간 투자량
요점	- **모델을 더 크게+더 많은 데이터+더 많은 GPU** → 더 강력한 모델	- 추론시 **더 많은 계산(시간)** 투입 → 더 나은 결과(정확/정교)
산업 영향	- AI 거대 모델 경쟁 가속 - AI 훈련 비용 대폭 증가 - 엔비디아 GPU 등 수요 폭증	- 훈련 비용 뿐만 아니라 **추론 비용 또한 확대** - 클라우드/GPU/ASIC 추론 인프라에 대한 투자 지속
기술 예시	- GPT-4 등 모델 훈련 비용 수억 달러 - LLaMA 등 주요 빅테크의 초거대 모델	- Beam Search 확장, 재추론, Prompt Chaining으로 품질 향상 - 일부 모델(o1, R1)에서 Test time Scaling 개념 시도
진화 방향	- 파라미터 수를 지속적으로 늘림 - 거대 데이터셋 확보 - 컴퓨팅 파워의 제한이 시장을 결정	- 추론 효율성과 유연성 강화 - 주어진 상황에서 필요한 만큼 '추론 스케일'을 확장해 성능 강화

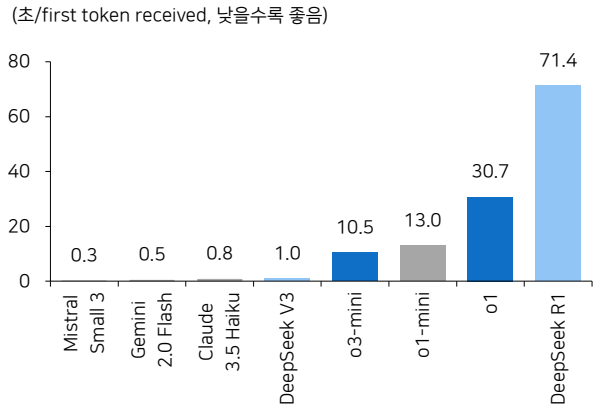
자료: 메리츠증권 리서치센터 정리

그림14 주요 AI 경쟁 모델 결과 산출 속도



자료: Artificial Analysis, 메리츠증권 리서치센터

그림15 주요 AI 모델 레이턴시(지연 속도)



자료: Artificial Analysis, 메리츠증권 리서치센터

Test Time Scaling이 트렌드라면, 향후 AI 모델 경쟁의 핵심은 '속도'일수도

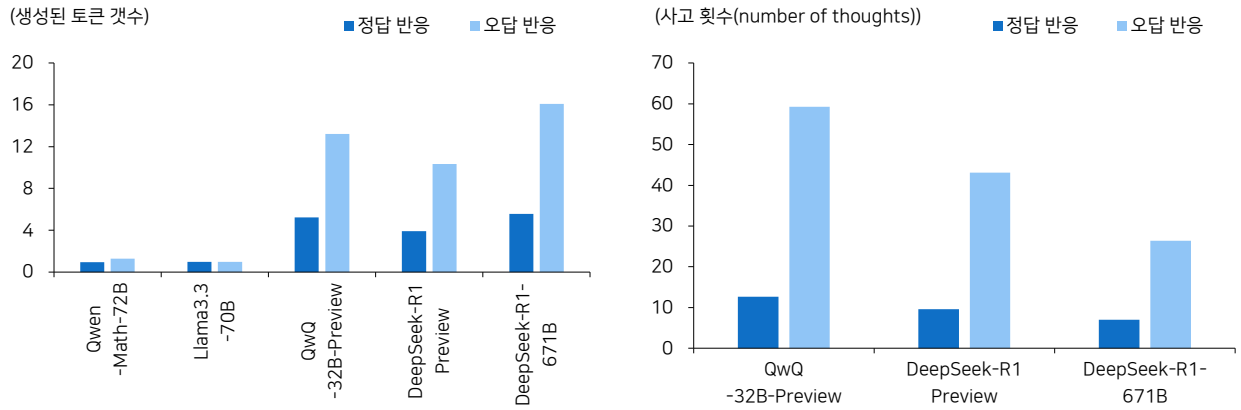
Test Time Scaling이 트렌드라면, 추론 시간 감축 여부가 향후 경쟁 요인일 수 있다고 생각

AI 모델에 투입되는 학습 비용을 낮추고, 추론 과정에서 성능을 향상시키는 Test Time Scaling이 AI 상품화 국면에서 모델 개선의 트렌드로 자리잡는 기술이라면, 향후 모델간 차별화의 핵심은 속도일 것으로 생각한다. 일정 수준 '쓸모 있는 수준'에 올라 서 있는 모델들의 수준을 고려하면, 비용 경쟁과 함께 부각될 수 있는 포인트는 모델이 생각하는 시간을 얼마나 줄일 수 있는 지일 것이다.

하드웨어 측면에서는 더 많은 프로세서를 투입하거나, 네트워크를 업그레이드 하거나

AI 모델의 속도를 개선시키는 일은 소프트웨어 개선과 하드웨어 개선 둘 다 가능하다. 소프트웨어 측면에서는 예를 들어 최근 Test Time Scaling이 적용된 모델에서 불필요한 추론 과정으로 비용을 유발하는 문제가 있어서, 강화학습 과정에 의도한 방향으로 보상을 주는 소프트웨어적 기술로 보완했다는 이야기가 있다. 하드웨어 측면에서는 더 많은 프로세서를 활용하거나, 네트워크 대역폭을 키우면 속도 개선이 가능하다.

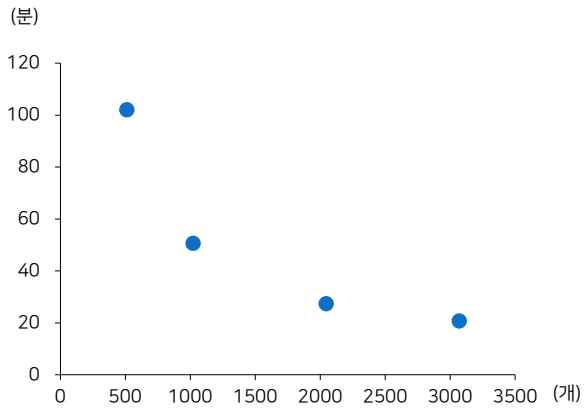
그림16 OpenAI, DeepSeek R1 같은 LLM 모델에서 나타나는 언더싱킹* 문제로 성능이 저하되는 문제 발견(Tencent AI)



주: * 모델이 문제 해결 과정에서 유망한 생각 경로를 충분히 탐색하지 않고 자주 다른 과정으로 전환해 깊이 있는 추론을 하지 못하는 현상을 의미. 특히 오답에서 더많은 불필요한 비용(토큰 생성량 및 사고 횟수 증가) 발생. 언더싱킹에 대해 점수화 및 디코딩 전략(S/W적 접근)으로 문제를 해결했다는 보고서

자료: Tencent AI Lab, 'Thoughts Are All over the Place: On the Underthinking of o1-Like LLMs('25.1.30)', 메리츠증권 리서치센터

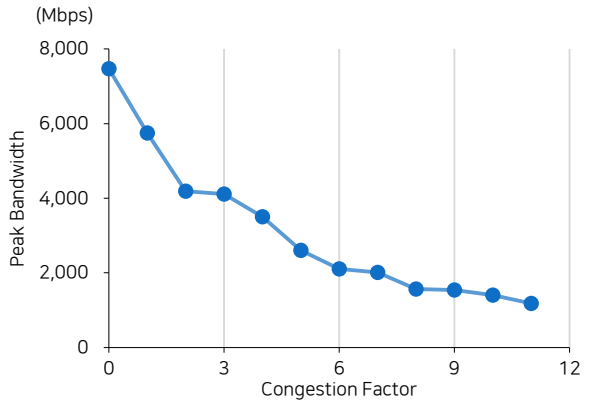
그림17 TPU 사용량 증가 → 레이턴시 하락



주: 사용 TPU는 Trillium으로 모두 동일

자료: MLCommons.org, 메리츠증권 리서치센터

그림18 대역폭이 높아질수록 Congestion Factor 하락



자료: Timo Schneider, Torsten Hoefler, 메리츠증권 리서치센터

AI 상품화 국면에서도 AI Capex가 필요한 이유

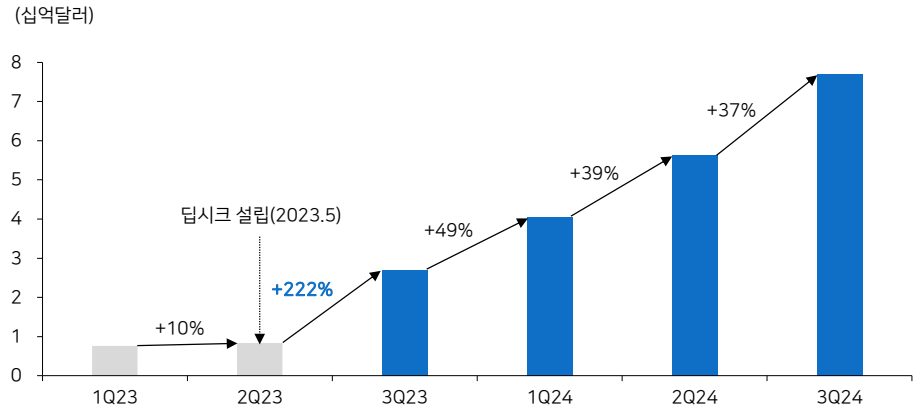
딥시크 R1과 이후에 출시된 o3-mini 모델의 경쟁 구도를 정리해 보면, AI 상품화 과정 속 모델들의 경쟁 과정에서 AI 하드웨어 투자가 필요함이 드러난다.

1) 기반 모델의 진보는 간소화 모델의 경쟁력에도 영향을 미치고, distillation을 통한 모델 개선까지도 고려하면 기반 모델의 성능 향상이 필요하고, 기본적인 GPU와 AI 기반의 Scaling Law는 여전히 유효하다.

2) 다만 1)보다 더 가속화를 기대할 수 있는 부분은 '추론 수요'이며, Test Time Scaling의 특성까지 고려해도 추론 수요 중심의 AI Capex 확장이 기대된다.

3) Time Scaling Law 기반의 모델 경쟁은 결국 '속도' 경쟁일 것이고 추론 연산에 더 많은 하드웨어 투입, 통신 네트워크 개선으로 경쟁력을 확보할 수 있다.

그림19 엔비디아의 싱가폴향 매출액 추이



자료: Bloomberg, 메리츠증권 리서치센터

표7 딥시크 공개 이후 주요 AI 인프라 기업 실적 코멘트 및 한국, 일본 대응 방향

구분	내용
마이크로소프트	- 매출 성장과 캐파 성장의 균형을 맞춰가야 할 것 - 딥시크로 어플리케이션 개발 비용 절감, 이후 활성화되면 어플리케이션 소비도 증가할 것 - AI PC 등으로 AI 앱이 보급된다면 서버와 PC 사업 모두 영위하는 MSFT에 좋은 소식
Meta	- 대규모 인프라를 구축하는 역량이 높은 품질의 서비스를 제공하기 때문에 여전히 Capex 대규모 투자가 전략적 우위 제공
코닝	- 매분기 신규 Generative AI 제품에 대한 수요 증가. 옵티컬 사업부 내 enterprise 매출이 4분기에 +93% 성장하는데 기여
구글	- 2025년 Capex 가이던스를 750억 달러로 제시(컨센서스 580억 달러) - 클라우드 고객들은 18개월 전보다 8배 이상의 컴퓨팅 용량을 AI 훈련 및 추론에 소비하고 있음 - 이를 충족하기 위해 인프라 투자를 지속할 것
이튼	- 하이퍼스케일러 고객만 2025년 3,000억 달러 자본지출 예상, 전년대비 30% 수준 - 2023년과 2024년 사이 데이터센터 건설속도 두배 증가 - 이 시장이 둔화될 것이라는 생각은 우리가 보는 어떤 데이터와도 일치하지 않음
일본	- 소프트뱅크 OpenAI와 인공지능 동맹, SB OpenAI 재팬 설립, 일본을 중심으로 AI 서비스 확대 추진 계획 - 일본 전역에 AI 개발용 데이터센터를 짓고 전력 수요를 충당할 발전 시설도 함께 건설할 방안 - 고급 엔터프라이즈 AI 크리스털 인텔리전스 소개, SB 전반에 통합하기 위해 연간 30억 달러 투자 계획
한국	- 정부, AI 컴퓨팅 인프라 전략 2월 발표. 연내 GPU 1만 5천개를 구매하겠다고 밝힘 - 기존에는 2030년까지 GPU 3만장 확보하는 계획이었으나, 늦어도 2027년초까지 3만장이 되어야 한다는 언급

자료: 언론 보도 종합, 각 기업 4Q24 실적 컨퍼런스 콜, 메리츠증권 리서치센터

ASIC, 통신 네트워크 관련주 중심 AI H/W 투자 유효

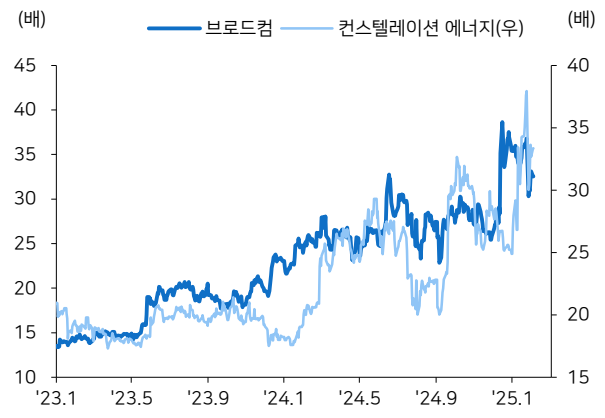
추론 중심의 AI Capex 확대 기대

GPU는 연산 속도, 효율 측면에서 ASIC 대비 불리하지만, 다양한 연산을 유연하게 할 수 있다는 장점이 있어 ASIC 대비 AI 학습에 유리하다. 반면 ASIC은 상대적으로 저전력, 고효율, 속도 측면에서 GPU 대비 AI 추론에 유리하다. 2월 4일 구글 실적발표 이후 Capex 계획 확대에 엔비디아의 주가반응보다 브로드컴의 주가 반응이 강했던 이유는 AI 추론 수요 중심의 Capex 확대에 대해서는 여전히 기대가 있고, '인기'가 살아 있는 것으로 생각한다.

ASIC, 네트워크 관련주의 기회

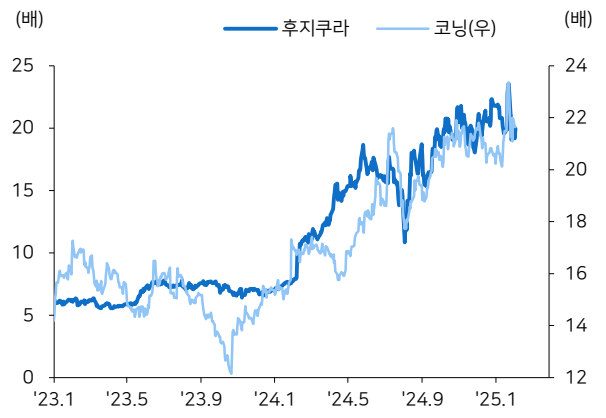
팔란티어처럼 PER이 우상향 추세에 여전히 남아있는 주식에 관심을 갖는다면, ASIC과 통신 네트워크 관련주들은 AI 하드웨어에 대한 관심이 소홀해진 시점이 오히려 기회일 수 있다는 의견이다.

그림20 브로드컴, 컨스텔레이션 에너지 12MF PER



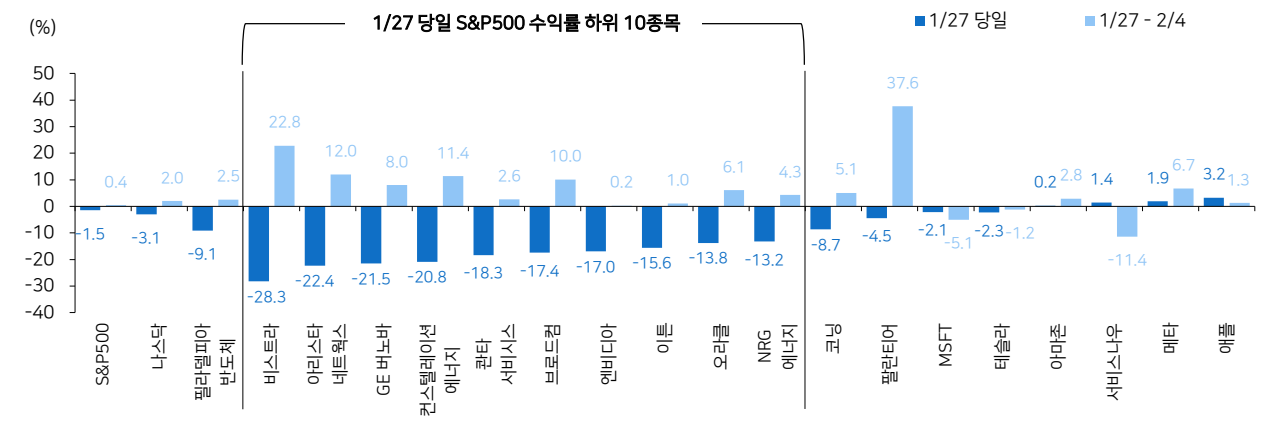
자료: Bloomberg, 메리츠증권 리서치센터

그림21 후지쿠라, 코닝 12MF PER



자료: Bloomberg, 메리츠증권 리서치센터

그림22 1월 27일 딥시크 충격 이후 주가 반응



자료: Bloomberg, 메리츠증권 리서치센터

Compliance Notice

본 조사분석자료는 제3자에게 사전 제공된 사실이 없습니다. 당사는 자료작성일 현재 본 조사분석자료에 언급된 종목의 지분을 1% 이상 보유하고 있지 않습니다.
본 자료를 작성한 애널리스트는 자료작성일 현재 추천 종목과 재산적 이해관계가 없습니다.
본 자료에 게재된 내용은 본인의 의견을 정확하게 반영하고 있으며, 외부의 부당한 압력이나 간섭 없이 신의 성실하게 작성되었음을 확인합니다.

본 자료는 투자자들의 투자판단에 참고가 되는 정보제공을 목적으로 배포되는 자료입니다. 본 자료에 수록된 내용은 당사 리서치센터의 추정치로서 오차가 발생할 수 있으며 정확성이나 완벽성은 보장하지 않습니다. 본 자료를 이용하시는 분은 본 자료와 관련한 투자의 최종 결정은 자신의 판단으로 하시기 바랍니다. 따라서 어떠한 경우에도 본 자료는 투자 결과와 관련한 법적 책임소재의 증빙자료로 사용될 수 없습니다. 본 조사분석자료는 당사 고객에 한하여 배포되는 자료로 당사의 허락 없이 복사, 대여, 배포 될 수 없습니다.