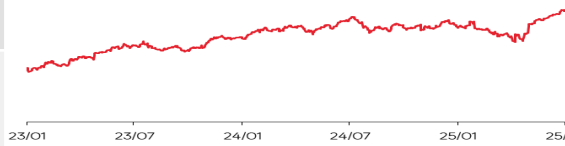
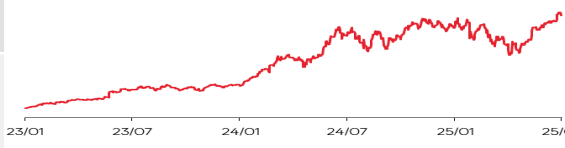

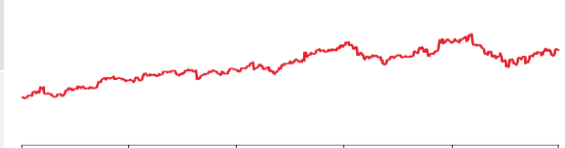
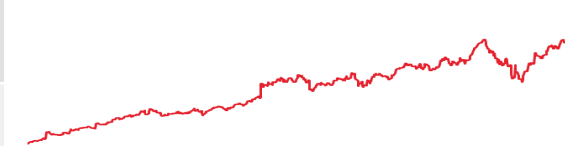
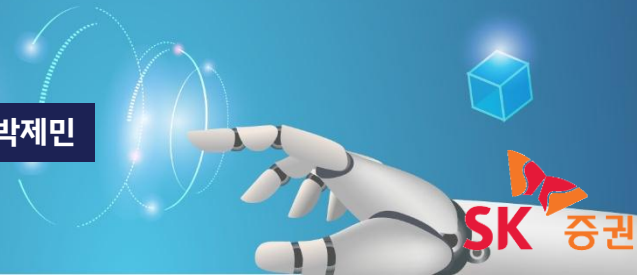


AI WEEKLY

SK증권 해외주식 박제민



종목 (티커)	시가총액(B, \$)	수익률 (% , 1W/1Y/커버리지 이후)			주가 추이 (23.01.01~)
마이크로소프트 (MSFT)	3,657	0.0	7.7	26.9	
- Azure로 AI 시대 하이퍼스케일러 포지션 구축, B2B 비즈니스 툴인 Office의 AI 침투율 제한은 염려 사항 - 강한 상승 이후 고점 부근 주가 횡보, Maia 2026년 연기 보도					
엔비디아 (NVDA)	3,740	-0.7	23.3	38.5	
- 추론 시대 도입에 따라 GPU 사용량 증가 지속, Test time scaling으로 훈련 수요도 여전 - 특별한 뉴스 없이 IB 추천 몰리며 6/25 신고가 갱신, 7/1 하반기 시작과 함께 차익실현으로 조정					
아마존 (AMZN)	2,340	4.0	11.8	21.1	
- 물류센터 추가 건설을 통한 배송 최적화 지속 진행 중, AWS 역시 비용 최적화 관점이 우선시 - 7/8부터 시작되는 Prime day 관련 기대감이 모이는 구간. Gen.AI 부문 핵심 인력 이탈 이슈					
알파벳 (GOOGL)	2,140	3.0	-3.9	10.5	
- 플랫폼 AI 제품 증가에 따른 Subscription 사업부 확장 주목. Search 부문 경쟁 우려는 지속 - EU DMA 관련 소송 리스크로 전 주에 빠졌던 주가를 일부 회복					
메타 (META)	1,808	1.5	42.5	25.4	
- GPU 전환, LLM 도입으로 피드, 광고 엔진 강화 진행 중. 광고 생성 엔진 가동 시작 - Super Intelligence Team 인재 영입, ASIC 관련 긍정적 뉴스 다수. 하반기 시작과 함께 일부 차익실현					



Enterprise AI로 영역을 확장하는 OpenAI

OpenAI B2B 서비스 구분

서비스 종류	비용	제공 내용 요약
Base models	최신 모델 기준 백만 토큰당 2달러	GPT-4o, o3 등의 최신 모델 사용 사용자 맞춤형은 불가능, 사전 정보를 쿼리에 input 해야
Fine-tuning	구형 대형 모델 기준 백만 토큰당 25달러	기업이 자체 데이터를 활용해 OpenAI 모델을 커스텀 터마이징할 수 있는 소프트웨어 아직 최신 모델(GPT-4o 등)은 파인튜닝 불가
Consulting + Fine-tuning	최소 1천만 달러 이상 계약 필요	OpenAI 엔지니어 팀이 고객의 기업 데이터를 바탕 으로 최신 모델을 파인튜닝 + 최종 결과물 테스트 및 개선 제안까지 포함한 통합 서비스 제공

자료 : SK증권

고객사 고유 데이터로 파인튜닝 + 컨설팅 서비스 제공 본격화

정부·대기업 대상 맞춤형 모델과 애플리케이션 개발을 진행. 최소 단가 \$10M로 설정, \$10B 단위 계약도 존재

고객 사례: 모건스탠리(자체 리서치 DB 기반 챗봇), Grab(동남아 스트리트맵 구축용 Vision 모델), 미국 국방부(전쟁 전략 수립용 AI 솔루션 개발)

Palantir FDE 전략 차용하여 컨설턴트 현장 파견하여 모델 커스터마이징. 모델 데이터 락인으로 중장기적으로 GPT 모델에 락인되는 효과 강할 것

OpenAI는 문서 협업 도구와 생산성 앱 시장 진입도 검토 중이라는 보도, MS Office, Google Workspace와 정면 경쟁 가능성 (TheInformation)

B2B SW 기업들은 걱정, OpenAI 관련 인프라 기업 주목 필요

OpenAI가 ChatGPT MAU 및 구독자 기반 Consumer Tech에서 Enterprise Tech 영역으로 확장하는 모습, 대부분 B2B SW 기업들에게 위협
AI 생산성 툴이 Fine-tuning 모델, Enterprise Agent 어느 형태로 진화하든, 특히 MS Office와 같은 현재의 비즈니스 생산성 툴은 위협받는 입장
MS Office의 상업용 유료 좌석(seat) 수는 현재 4.3억개, Copilot 적용 비중은 3% 이하 추정. 신규 채용 감소, 향후 인력 대체로 축소 전망
(6/29) Oracle은 Stargate 프로젝트로 추정되는 연간 \$30B 규모의 단일 클라우드 계약 체결 공개, 현재 인프라 매출의 3배. 발표 후 주가 6% 상승

연초 발표된 Stargate Project, OpenAI의 컴퓨팅 수요 담당할 DC 건설



자료 : kaohooninternational, SK증권

AI WEEKLY

SK증권 해외주식 박제민

손정의 회장의 ASI 계획

Cristal Intelligence Agent 소개



자료 : Softbank, SK증권

ARM 활용하여 Cloud 점유율 확장 계획



자료 : Softbank, SK증권

6/27 주주총회(Analyst Call) 진행, ASI 플랫폼 기업 선언

SoftBank의 존재 이유는 ASI 실현, ASI 시대의 산업 기반(플랫폼) 제공자로서의 포지셔닝하는 것이 목표

ASI는 인간 지능을 초월한 AI로, reasoning과 acting 가능한 AI Agent, 로봇이 현실화되는 시기

핵심 자산 1) ARM: 클라우드용 CPU 시장에서 50% 점유율 달성 목표, Graphcore, Ampere 인수로 역량 높이는 중

핵심 자산 2) OpenAI: 그간 4.8조엔 투자. Crystal Intelligence JV 출범. OpenAI는 향후 세계 최대 시가총액 기업이 될 것으로 기대

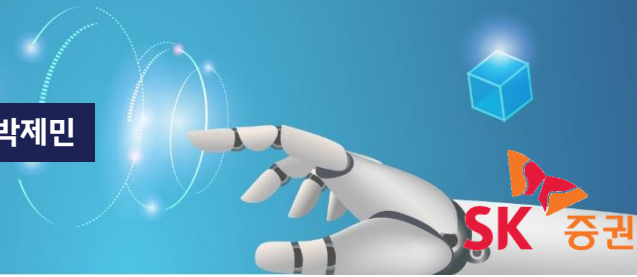
Crystal Intelligence는 Generative AI 기반 플랫폼. Agent들이 동작하는 플랫폼 + 생성 툴킷 + 상호작용 구조 전반을 포괄하는 개념

플랫폼 내 백오피스 Agent 우선 생성 중. 향후 Yahoo, LINE, PayPay 등 그룹사에 우선 적용 후 외부 기업 대상으로 확장 예정

Stargate에는 향후 4년간 최대 78조엔 투자 계획, 단계적으로 진행될 것이며 OpenAI와 공동으로 PF 방식 추진 중

AI WEEKLY

SK증권 해외주식 박제민

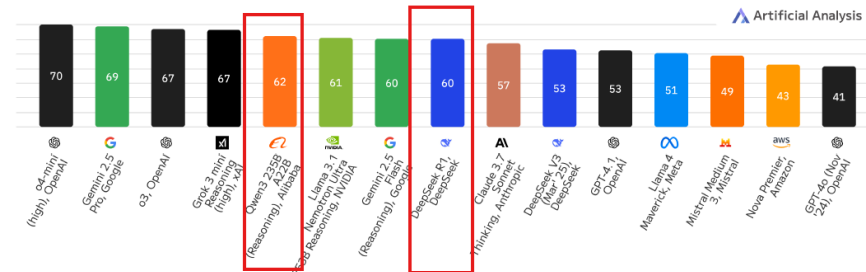


서방 세계의 공격에 휘청이는 DeepSeek

1Q25 LLM 모델 순위표

Leading Large Language Model (LLMs), by AI lab

Artificial Analysis Intelligence Index (incorporates MMLU-Pro, GPQA, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, MATH-500)



자료 : Artificial Intelligence, SK증권

AI 계에 큰 혁신을 불러올 것으로 기대를 모으는 R2



자료 : AIGrid, SK증권

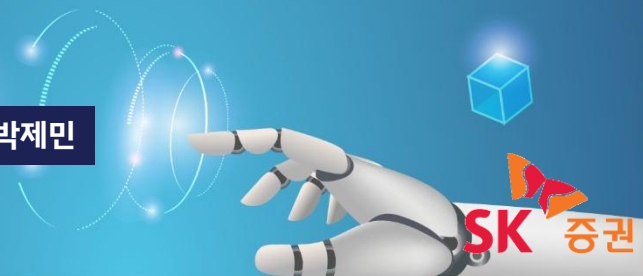
미국, 유럽 DeepSeek 퇴출 움직임

TikTok 퇴출 논리였던 '중국 국가 정보법 리스크'가 LLM 에도 번지는 중. 해당 법안으로 중국 기업은 정부 요청 시 데이터 및 기술을 당국에 제공해야 미국은 DeepSeek가 중국 군 및 정보기관에 기술을 제공하고 있다고 주장
독일 개인정보보호위원회 애플과 구글 스토어에서 Deepseek 어플 삭제 요청, 그 외 유럽 국가들도 공공 및 민간 부문 사용 통제 시작
각종 벤치마크에서 최신 Alibaba(Qwen3), DeepSeek(R1) 모델들을 모두 Frontier 모델 수준으로 평가, 추론 가성비 면에서는 더 뛰어난 상황
Llama 4.0의 부진으로 오픈소스 진영에서 더 활발하게 활용 중, DeepSeek의 Chatbot Arena Elo 순위 4위, 11위. LLM 경쟁구도에 큰 영향 예상

미국 H2O 수출 금지로 인해 DeepSeek 차세대 R2 출시 지연

현재 DeepSeek의 주력모델인 R1은 RL, MoE 등 신규 LLM 최적화 기술로 업계에 큰 방향, 조만간 R2 공개가 예정돼있었음

1) CEO 성능 불만족, 수정 단계. 칩 공급 부족 영향 추정 2) 출시 이후 배포에도 난관 예상. CUDA 최적화 모델로 H2O 보유 CSP만 배포 가능 전망도
Huawei 등 중국산 칩은 성능, 호환성 면에서 아직 불리하며 경쟁력이 부족한 상황



Meta Update: ASIC, Capex 그리고 AGI

랭킹 추천 모델 시행 시 MTIA vs GPU의 Perf/TCO(가성비) 추이

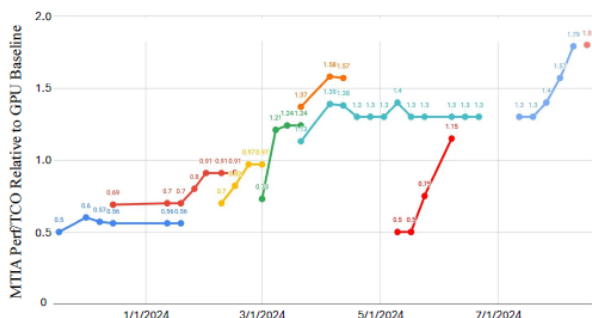


Figure 4: Continuous optimization improves MTIA 2i's Perf/TCO for a key ranking model used at Meta. The lines represent various variants of the model.

자료 : META, SK증권

주: 특정 case에서 TCO 80% 이상 우위 나타나나 전반적인 TCO 우위는 44% 수준

빨라지는 ASIC 적용, 대규모 Capex 소식

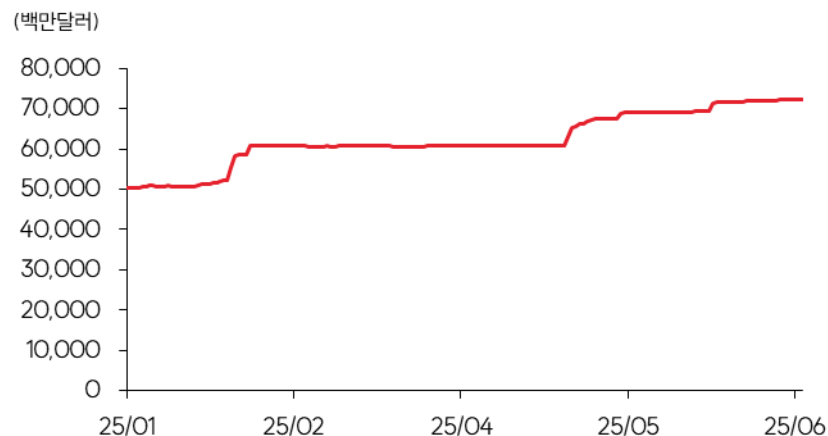
(6/25) ISCA 학회에서 ASIC인 MITA v2 스펙 공개. Hopper 대비 TCO 44% 절감, HBM 생략, SRAM 적용. 대규모 테스트 진행 완료 이후 MTIA ASIC을 2025년 Q4부터 양산, 배치 규모 100~150만개 추정 보도. (vs Google TPU: 150~200만)

META는 DAU 20억명의 광고, 피드 추천 엔진을 Transformer 기반으로 전환 중, 단일 AI 워크로드가 가장 큰 수준의 컴퓨팅 요구하는 것으로 판단 (6/30) \$29B 규모 데이터센터 프로젝트를 위한 자금 조달 뉴스. META의 올해 Capex 전망은 \$64B~\$72B, 추정치 계속 상향 중

META, AGI향 투자를 적극 늘려가는 중

Alexander Wang(Scale AI CEO)에 이어 Nat Friedman(GitHub CEO) 영입. OpenAI 출신 연구자 총 7명을 영입 보도(TheInformation) 지난달 Scale AI 인수 및 전담 조직 개설 소식은 Llama 4.0의 실패를 사실상 인정하는 행보, 그러나 주가는 오히려 상승. 이번 인재 추가 영입에도 상승 Meta는 기존 사업부가 Consumer 향 AI 서비스로 비즈니스 모델을 완전하게 구축한 것으로 판단, 이제 AGI 관련 시가총액이 붙는 초기 단계 투자자 입장에서는 기존 광고 서비스향 시가총액과 구분해서 바라보아야, 최악의 경우 현재 \$10B 이상의 적자를 내는 RL 사업부가 추가되는 꼴

META CY2025 Capex 전망치 추이



자료 : Bloomberg, SK증권