

G2의 AI 전쟁 바라보기

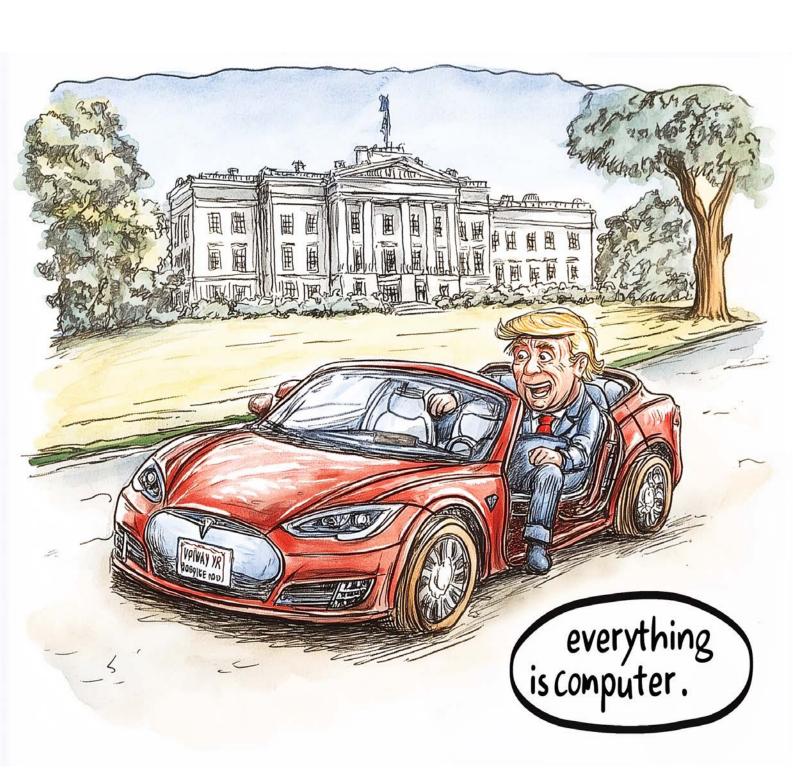
폐쇄소스 미국 vs 오픈소스 중국

한종목

김은지

chongmok.han@miraeasset.com

eunji.kim.a@miraeasset.com



CONTENTS

I. G2 tensions	5
1. 디지털 중상주의의 서막 (1) "소프트웨어가 세상을 삼키다" (2) 중국의 소프트파워 전환과 미국의 기술 우위 사수 노력 (3) 기술주의 "운율": 닷컴 버블의 학습효과와 빅테크의 대규모 투자 행보 2. 미국의 전방위적 견제 (1) OpenAI의 읍소 "DeepSeek 쓰지말자" (2) CSIS의 화웨이 보고서 (3) 에릭 슈미트의 새로운 논문 (4) 왜 이런 견제와 우려가 계속될까?	5 7 9 11 11 12 16 18
II. Agent war	21
1. 중국의 Manus Al (1) 제2의 DeepSeek 쇼크? (2) Manus의 두 얼굴 2. OpenAl의 카운터 (1) Agent 관련한 API와 SDK 출시 (2) 2025년은 에이전트의 해	21 21 24 27 27 30
Ⅲ. 미국의 해자인 자율주행 소프트웨어	33
 테슬라 FSD의 중국 진출이 갖는 의미 (1) 자율주행 시장은 과점 시장이 될 것 (2) FSD의 중국 진출 배경 및 데이터 처리 전략 중국판 FSD의 등장과 사용자 경험 (1) "중국판 FSD"의 출시와 특징 (2) 가격 책정 및 시장 잠재력 (3) 중국판 FSD의 실제 성능과 사용자 반응 "모든 것이 컴퓨터다": FSD의 중국 진출이 보여준 비교우위 (1) 중국판 FSD의 기술적 특성과 테슬라의 방법론 (2) End-to-end 방식이 좋은데 왜 빨리 못 따라 하나? (3) 미국 기업 테슬라의 SW 격차 우위와 함의 	33 34 35 35 36 38 40 40 42 43
IV. Paper of the Week	46
Al Scribe 애플리케이션의 효용 조사	46
V. Al Peer Table	50
VI. Appendix: Charts	52

Highlight of the Week

디지털 중상주의와 AI 패권 경쟁

1. 디지털 중상주의의 서막

지난 3월 11일 도널드 트럼프 대통령이 백악관에서 테슬라 차량의 기술을 경험하며 "모든 것이 컴퓨터다!"라고 말해 화제가 되었다. 이는 미국의 국가 정책이 AI와 소프트웨어 중심 으로 전환될 것을 시사하는 발언으로 소셜 미디어에서 바이럴됐다. 안두릴 CEO 팔머 럭키 는 미국 제조 경쟁력의 핵심이 소프트웨어라는 점을 강조했다.

미국은 소프트웨어 및 AI에서 중국과 경쟁을 벌이며 기술 우위를 유지하려고 한다. 중국은 AI 기술을 글로벌 사우스 국가들에게 개방하며 소프트파워 전환을 시도 중이며, 오픈소스 전략을 통해 국제적 영향력을 확대하고 있다. 미국과 중국은 AI를 둘러싼 국제 정치 경쟁에 서 미국은 폐쇄적 프론티어 모델 개발을, 중국은 오픈소스 개발을 전략적으로 강조한다.

2. Agent War: 중국 Manus AI의 등장과 미국 OpenAI의 대응

최근 중국의 Monica가 만든 자율 AI 에이전트 Manus가 큰 화제를 모았다. Manus는 사 용자 지시에 따라 작업을 계획하고 문제를 해결하며, 웹사이트 제작이나 보고서 작성 등에 서 뛰어난 능력을 보였다. 그러나 Manus는 독자 기술 없이 미국 기업 앤트로픽의 Claude 모델을 API 형태로 이용하는 "래퍼(wrapper)"라는 비판을 받고 있다. 이러한 Manus의 성 공은 SaaS 시장을 크게 변화시킬 가능성이 있으며, 일반인들도 손쉽게 소프트웨어를 개발 할 수 있는 시대를 열 것으로 전망된다.

이에 대응해 OpenAI는 AI 에이전트를 쉽게 개발할 수 있는 API와 SDK를 발표했다. 웹 검색, 파일 검색, 컴퓨터 사용 등 실제 업무 수행 능력을 갖춘 "행동 중심"에이전트를 지 원하며 시장 경쟁력을 확보하려고 한다. 2025년은 AI 에이전트 시대의 시작을 알리는 해로 평가된다.

3. 미국의 기술적 우위와 중국의 AI 기술 추격전

미국은 AI와 반도체 분야에서 중국 기업인 DeepSeek와 화웨이를 적극적으로 견제 중이다. OpenAI는 DeepSeek를 중국 정부 통제를 받는 기업으로 규정하며, 중국산 모델 사용 금 지를 요청했다.

미국 전략국제문제연구소(CSIS)는 화웨이가 TSMC에서 AI 칩을 우회 수입하고 있다고 주 장했으며, 이는 미국의 기술 우위를 위협할 수 있다고 경고했다. RAND 코퍼레이션에 따르 면 화웨이의 기술은 엔비디아의 최신 기술 대비 여전히 열세에 있지만, 기술적 격차가 빠르 게 줄고 있어 미국의 우려가 크다.

4. 테슬라 FSD의 중국 진출과 자율주행 소프트웨어의 중요성

최근 중국에서 테슬라의 자율주행 소프트웨어 FSD가 처음으로 승인되었다. 이는 중국이 아직 자율주행 소프트웨어에서 미국 기업에 기술적으로 의존할 필요성을 인정한 사례다. 중 국에서 출시된 "중국판 FSD"는 현지 도로 환경에 맞게 일부 최적화되어 있지만, 테슬라의 미국 모델을 기반으로 한다. 초기 사용자 반응은 예상보다 뛰어난 성능을 보인다고 평가했 으며, 전문가들은 기본 기술력이 매우 높다고 분석했다.

테슬라의 FSD 기술은 End-to-end 방식으로, 차량 센서 데이터를 직접 차량 제어 명령으 로 변환하는 방식이다. 이 방식은 데이터가 방대해야만 가능하기에, 중국 기업들이 따라잡 기 어렵다. 따라서 자율주행 소프트웨어는 미국이 지켜야 할 가장 중요한 기술적 비교우위 로 평가된다.

Paper of the week: Al Scribe 애플리케이션의 효용 조사

AI 기반 임상 문서 작성(Scribe) 기술은 의사의 진료 기록 작성 부담을 줄이고 업무 효율성 을 높일 것으로 기대된다. 실제로 Nuance의 DAX Copilot을 사용한 연구 결과, 진료 당 기록 작성 시간, 당일 진료 마감률, 근무 시간 외 업무 시간이 개선되었고, 임상의들은 환 자와의 대화에 더 집중하고 정신적 부담이 줄었다고 보고했다. 이는 Al Scribe가 진료 중 실시간으로 진료 기록 초안을 생성함으로써 임상의의 업무 효율을 높이고 환자 진료의 질 을 향상시킬 수 있음을 시사한다.

그러나 AI가 생성한 기록의 품질에 대해서는 엇갈린 반응이 있었으며, 오류 수정 및 교정에 추가적인 시간이 소요된다는 문제점이 지적되었다. 특히, 환자의 증상 표현은 그대로 유지 하면서 상황에 맞게 의학 용어를 사용해야 한다는 점에서 AI 모델의 성능 개선 필요하다. 또한, 병원 차원에서의 AI Scribe 도입을 위해서는 비용 대비 효과, 즉 환자 만족도 및 진 단 신뢰도 증가와 같은 수익성 지표와의 상관관계에 대한 추가적인 연구가 필요하다.

표 1. Al 관련 주요 일정

일	월	화	수	목	금	토
16	17	18	19	20	21	22
	· GTC 2025(~21)	· 크래프톤 실적(예)		· 포스 패러다임 실적(예)· ACN 실적· MU 실적	· 루닛 실적(예)	
23	24	25	26	27	28	29
			• 센스타임 실적			
30	31	1	2	3	4	5

자료: Bloomberg, 미래에셋증권 리서치센터

I. G2 Tensions

1. 디지털 중상주의의 서막

(1) "소프트웨어가 세상을 삼키다"

지난 3월 11일, 도널드 트럼프 대통령은 백악관 남쪽 잔디밭에서 테슬라 차량들을 전시하 는 이벤트를 개최했다. 이 자리에서 트럼프 대통령이 빨간색 테슬라 모델 S의 운전석에 앉 아 차량의 첨단 기술을 접하며 "Wow... everything's computer!"라고 말한 것이 화제가 되었다.

이 발언은 그가 전기차/자율주행차의 디지털 제어판과 컴퓨터 기반 기능을 처음 경험하며 느낀 감탄을 반영한 것으로 보인다. 당시 조수석에는 일론 머스크가 앉아 있었다. 과거만 하더라도 트럼프 대통령은 그닥 기술 친화적인 인물은 아니었던 것처럼 보였다. 이메일 대 신 볼펜을 선호하는 경향이 있고, "디지털이 뭐죠?"라는 질문을 한 적도 있다. 이런 맥락에 서 최고 국력을 가진 나라의 최고 권력을 가진 자가 모든 게 컴퓨터라고 한 발언은 그 의미 가 적지 않다.

이 간단한 발언은 소셜 미디어에서 빠르게 바이럴되어 "밈"이 되었다. 간단하지만 그만큼 강력한 한 마디이기 때문이다. 영화 아라비아의 로렌스(Lawrence of Arabia)에서 나온 명 대사와 같이 '위대한 일은 작은 발단에서 비롯'되는 법이다. 참고로 영화 속 로렌스는 제한 된 자원으로 최대의 효과를 내기 위해 전쟁의 본질을 다시 바라봤고, "외국인들은 항상 가 르치러 오지만, 오히려 배우는 것이 훨씬 낫다"고도 말했다. 베이비 부머 시대에 태어난 트 럼프가 머스크라는 "이방인"을 만나, 컴퓨터 소프트웨어가 현대전의 모든 것임을 알게 되었 으리라 사료된다.



그림 1. 영화 프로메테우스에서 AI 로봇인 "데이비드"도 인용한 명대사

자료: 영화 "프로메테우스", 미래에셋증권 리서치센터

> 따라서, 트럼프의 발언은 여러 소프트웨어 개발자들의 어깨를 으쓱하게 하는 데 충분했다. AI와 방산의 융합을 이끄는 업체 중 대표적인 기업인 안두릴의 CEO 팔머 럭키는 트럼프의 말을 이렇게 해석했다.

> "간단히 말해서, 소프트웨어가 세상을 삼키고 있다"는 말이었다. 이 말은 미국에서 가장 유 명한 벤처캐피털인 안드레센 호로위츠(Andreessen Horowitz)의 설립자 마크 안드레센이 2011년에 남긴 말이다. 그리고 이 문구는 이후 실리콘밸리의 일종의 선언문이 되었다.

그림 2. 트럼프의 말을 안두릴의 CEO가 재해석 안두릴은 국방 산업에 AI 기술을 더한 기술 기업… 앞으로 미국의 제조 경쟁력은 SW에서 나온다!



자료: Palmer Luckey, 미래에셋증권 리서치센터

안드레센은 소프트웨어 기업들이 경제의 광범위한 영역을 장악할 것이며, 기존 산업 구조를 뒤엎을 것이라고 예측했다. 그리고 현재 시점에서 샌프란시스코의 철학은 결국 워싱턴 D.C. 까지 성공적으로 다다랐다.

트럼프 행정부의 정책 기조가 앞으로 소프트웨어와 그것을 구동할 컴퓨터에 맞춰 혁신할 것이라는 점에서 이는 의미하는 바가 크다. 그리고 현재의 컴퓨터 소프트웨어를 대표하는 가장 좋은 기술은 바로 AI다. 다시 말해, 트럼프는 그가 존경한다는 윌리엄 맥킨리 대통령 시절 미국의 중상주의적 황금기를 부흥시키기 위해. AI 기술로 각종의 문제의 해법을 찾으 려 부단히 노력할 것이다.

(2) 중국의 소프트파워 전환과 미국의 기술 우위 사수 노력

실제로, AI는 국제 외교의 가장 중요한 화두로 떠올랐다. AI Weekly에서 자주 언급했다시피 핵무기와 같은 파괴력을 지닌 비대칭적 전력이기 때문이다. 앞으로 더 많은 국가 지도자들 이 AI에 대해 지겹도록 언급할 것이다.

또한 그 점증하는 중요성 때문에, 복잡하게 얽힌 국제 외교의 수많은 실타래가 결국 AI라는 한 줄기 논리로 풀어질 가능성이 높다. 한 예시로 프랑스의 대통령 마크롱은 국가 주도의 투자(AI 포함)를 하지 않으면 유럽이 죽을 수도 있다고 말했다. 아프다고 말한 정도가 아니 라, 투자를 하는 것이 말 그대로 생사의 문제라고 한 것이다.

과거 2000년 3월 8일, 빌 클린턴 대통령은 존스 홉킨스 대학에서 중국이 인터넷을 통제하 려는 시도를 언급한 적이 있다. 이때 클린턴은 "행운을 빕니다! 마치 젤리를 벽에 못으로 박는 것과 같습니다!"라고 말했다.

1990년대 말 중국이 인터넷 검열 시스템인 Great Firewall(황금 방패 프로젝트)을 구축하 기 시작한 것을 보고 조소 섞인 발언을 한 것이다. 클린턴의 이 유명한 발언은 인터넷의 자 유로운 특성을 고려할 때 중국이 인터넷에 대한 기술 통제와 사회 실험은 불가능할 것이라 는 서방 세계의 일반적인 견해를 반영했다.

* 1998년에 '황금 방패 프로젝트'가 시작되었으며, 이는 이후 '대화재벽'으로 알려진 시스템의 기초를 형성했다. 이 프로젝트는 인터넷 콘텐츠를 필터링하고, IP 주소와 도메인 이름을 차단하며, 데이터를 검사하는 기술적 인프 라를 구축하는 것을 목표로 했다

그러나 역설적이게도, 약 20년 동안의 집착을 통해 중국은 결국 "젤리를 벽에 못으로 고정 하는"데 거의 완전히 성공했다. 그리고 인터넷 혁명 다음의 AI 혁명의 시대로 접어들자 중 국의 태도는 180도 바뀌었다.

2025년 3월 7일 제14기 전국인민대표대회 제3차 회의 기간 중 열린 기자회견에서 왕이 외교부장은 "과학 기술은 철의 장막을 치는 도구가 되서는 안 되며 서로 보편적으로 공유 해야 할 재산이 돼야 한다"고 말했다.

왕이는 중국의 기술이 사람들의 상상을 돌파하고 있다면서, 봉쇄가 있는 곳에 돌파가 있었 고 압박이 있는 곳에는 혁신이 있다고 했다. 개구리 올챙이 때 생각 못하듯 아주 의기양양 해진 모습이다. 그러나 그만큼 기술과 혁신에 대한 자신감이 있다는 것이고, 그 자신감은 분명히 실체가 있다.

이어서 왕이는 양자컴퓨팅, 5G 등 여러 기술 중에서도 특히 AI에 대한 내용에 대해 많이 언급했다. 시진핑이 내놓은 AI 글로벌 이니셔티브에 대한 것인데, 브라질과 아프리카 등 글 로벌 사우스 국가들에 대해 본인들의 AI 기술들을 개방하기로 했다고 말했다.

> 즉, 중국은 AI 기술을 이제 국제 협력과 외교의 중요한 도구로 활용하려고 하고 있다. 이를 위해 오픈소스 형식의 AI 개발을 옹호하는 것이라 할 수 있다. 최근의 중국은 상대국에 고 함을 치는 듯한 전랑외교의 행태를 벗어 던지고 소프트파워를 강조하고 있다.

> 주변 국가, 더 나아가 BRICS+ 국가들에 대한 AI 기술 수출국가이자 AI 거버넌스의 중심에 서 있기 위한 방식에 있어 전랑외교는 전혀 어울리지 않기 때문이다.

> 그래서 오픈소스의 철학을 중국 공산당 레벨에서 강조하고 홍보하는 것이라 생각된다. 앞으 로도 많은 제2의 DeepSeek 업체들이 나와 연쇄적인 Deep Impact를 줄 수 있다고 생각 한다. 그리고 작고 빠른 AI 스타트업들이 다수 출현해 오픈소스 개발 경쟁을 하는 것도 중 국 정부 입장에서는 사실 컨트롤하기가 더 용이하다고 판단할 수 있다. 물론 알리바바의 마 윈에게 기술 복권을 시켜준 것도 AI로 돈을 많이 벌라는 측면보다는, 하루 바삐 AGI 시대 의 글로벌 주도권을 만들라는 숙제를 던진 것이다. "시진핑 사상" 중 중요한 면을 차지하는 "국진민퇴"는 AI에서도 마찬가지일 것이라 사료된다.

그림 3. 시진핑 중국 국가주석이 주재한 기업인 포럼에 마윈… 약 4년만에 공식석상에 복귀 이번 만남에서 시진핑은 마윈과 악수를 나눴으며, 이는 기업인 중 단 세 명에게만 주어진 예우



자료: CCTV, 미래에셋증권 리서치센터

반대로, 미국은 오픈소스의 위험성을 반복적으로 꺼내 분위기를 가라앉히고 주의를 환기시 킬 것이다. 특히 사실상 중국을 겨냥한 '국가 통제 AI'에 대한 디스토피아로 가는 잠재적인 폭주기관차가 바로 오픈소스임을 강조할 것이다. 물론 내부적으로 프론티어 성능의 폐쇄소 스 개발에 박차를 가해야 한다는 압박에 시달릴 것이다. 이것은 물론 더 많은 컴퓨팅 자원, 더 많은 에너지 활용을 의미하게 된다.

권위주의 진영이 오픈소스를 표방하고 자유주의 진영이 폐쇄소스를 지향하는 것은 정말로 흥미로운 역설이다. 물론 미국과 중국 모두 오픈소스에 대해 신중해야 하고 또 폐쇄 모델의 오용을 모니터링해야 해야 할 것이다. 무엇이 옳고 그른지의 문제는 아니다. 가치 판단을 앞세운 각자의 경쟁 전략이다.

> 그리고 최근 주요 경제 현상 중 하나는 '비교 우위' 개념이 무너지고 있다는 것이다. 중국은 경공업부터 첨단산업까지 모든 분야를 석권하며 미국을 포함한 타국의 경제적 기회를 흡수 했다. 미국이 대만의 TSMC를 압박하는 것도 비교 우위를 얻기 위한 워싱턴의 조급함을 보 여주는 사례다. 이런 냉혹한 현실과 트럼프의 "모든 게 컴퓨터다"라는 말을 종합해 생각해 면, 트럼프의 중상주의적 국가 운영의 핵심이 될 주인공이 소프트웨어와 AI라는 점을 알 수 있다. 중국에 대해 절대적으로 사수해야 할 비교우위가 소프트웨어라는 점은 팔란티어의 CEO 알렉스 카프가 몇 년 전부터 강조해온 말이다.

> AI가 모든 것의 화두가 되고 있고, 모든 물자와 역량이 총동원될 글로벌 패권의 엔드게임이 다. 그리고 닷컴 시기는 미국만이 주도하는 단극 시대였다면 지금은 "얄타 2.0"의 시대다. 인터넷 버블 때와 비슷하지만 엄밀하게 다르다고 우리 팀이 생각하는 이유다. 뒤처짐의 대 가가 너무 가혹할 것이기 때문이다.

- * 1945년 미/영/소 얄타 회담을 현대적으로 비유한 "얄타 2.0"
- 시진핑과 루스벨트(미국): 중국은 일대일로 이니셔티브와 우크라이나 평화유지 역할 제안을 통해 글로벌 영향력 을 확대하려 함. 이는 루스벨트가 유엔과 브레튼우즈 체제를 설립하며 세계 질서를 주도한 것과 유사하게 보이기 도... 시진핑은 트럼프와 푸틴과 함께 새로운 세계 질서를 형성하려는 의도를 내비침.
- 푸틴과 스탈린(소련): 우크라이나 침공과 구소련 국가들에 대한 영향력 강화로 스탈린 시기 소련과 유사한 행보.
- 트럼프과 처칠(영국): 트럼프의 "미국 우선" 정책은 국제적 약속 축소와 리더십 후퇴로 이어질 가능성이 있고, 처칠 시기의 영국처럼 동맹국간 주도적이지 않은 위치로 간주될 수 있음.

(3) 기술주의 "운율": 닷컴 버블의 학습효과와 빅테크의 대규모 투자 행보

History doesn't repeat itself, but it often rhymes. (역사는 반복되지 않지만, 종종 비슷한 방식으로 운율을 이룬다) - 마크 트웨인

닷컴 버블 시기 기술주 붕괴를 연상시키듯 요며칠 증시 하락은 아주 가파랐다. 아마존 (AMZN)은 2000년부터 2003년까지 \$113에서 \$6까지 폭락했었다. 90% 이상 하락한 것 이다. AI는 다를까 하는 의심의 눈초리가 많이 생겼다.

그러나 결국 아마존은 명확한 인터넷 시대의 승자로서, 이후 400배나 성장했다는 사실을 되돌아볼 필요가 있다. 우리 팀은 닷컴 버블 때와 비슷한 방식으로 AI 주식들의 향방이 그 운율을 따를지는 모르겠지만, 시장의 습성인 '학습 효과'를 더 믿고 있다.

게다가 AI의 개발 속도는 점점 압축되고 있고 혁신의 속도가 더 가팔라지고 있다는 점도 감안해야 한다. 다시 말해, 시장의 충격은 있겠지만 지난 시대의 교훈이 있고 AI 시대의 승 자가 가질 왕관의 가치가 너무나 막강하기 때문에 미래를 좀 더 긍정해보게 된다.

> 현 상황에서 가장 유리한 입지를 지닌 빅테크들의 경우에도 투자의 지속적 확대 흐름은 여 전히 관찰된다. 3월 12일에 텍사스 휴스턴에서 열린 CERA Week 2025 컨퍼런스에서 아마 존, 구글, 메타가 2050년까지 전 세계 원자력 에너지 용량을 3배로 늘리는 서약에 서명했다.

> 현재 전 세계 원자력 용량은 417개의 운영 중인 원자로에서 생산되는 377GW로, 2050년 까지 이를 1,131GW로 확대하는 것을 목표로 한다는 것이다. 세계원자력협회가 주도한 이 서약은 원자력 산업 외부의 주요 기업들이 '처음으로' 원자력 발전의 대규모 확장을 공개적 으로 지지한다는 점에서 의미가 깊다.

> 원전뿐 아니다. AI를 위한 에너지 수요는 조금 더 "친환경적이지 않은" 천연가스에 대한 관 심도 다시금 일으켰다. 마이크로소프트는 에너지 수요를 따라잡기 위해 AI 데이터센터에 전력을 공급하기 위해 천연 가스를 사용할 수 있다고 말했다. AI 때문에 ESG 정책을 사실 상 폐기한 것이다.

> 한편, 일론 머스크의 xAI는 최근 AI 데이터센터 입지를 강화하기 위해 멤피스에 100만 평 방피트 시설을 인수하기도 했다.

> 또한 아마존의 CEO 앤디 재시는 "AWS 사업에서 생성 AI가 연간 수십억 달러 규모의 사 업이고 전년 대비 세 자릿수 퍼센트로 성장하고 있다. 하지만, 만약 더 많은 용량이 있다면 그것을 활용해 수익화할 수 있을 것이다"라고 했다. 연간 2배 이상씩 성장하는 사업이니 기 회가 있을 때마다 AI 칩을 구매하겠다는 말이다.

> 그리고 오라클의 창업자이자 CTO인 래리 엘리슨은 "CSP 3사의 오라클 데이터베이스 멀티 클라우드 수익은 지난 3개월 동안만 92% 증가했다. AI 훈련을 위한 GPU 소비는 지난 12 개월 동안 244% 증가했다."고 말했다. 마이크로소프트의 약간이나마 소극적으로 바뀐 최 근의 톤 변화와 달리 현실은 투자 감속의 여지를 찾아보기가 힘들다.

> * "Database MultiCloud"는 오라클이 자사의 데이터베이스 기술을 다른 주요 클라우드 제공업체(마이크로소프트, 구글, 아마존 등)에 제공하여, 이들이 오라클의 테이터베이스 서비스를 자신들의 클라우드 플랫폼에서 사용할 수 있게 하는 것. 예를 들어, 아마존 AWS나 구글 클라우드에서 오라클 DB를 실행할 수 있는 환경을 제공하는 식.

> TSMC의 미국 애리조나 칩 공장은 2027년 후반까지의 생산량이 벌써 모두 판매된 것으로 최근 알려졌다. 물론 애플, AMD, 퀄컴, 엔비디아, 브로드컴과 같은 주요 미국 고객들의 수 요를 충당하기 위해서다.

> 빅테크들 중에서 최근 가장 놀라운 베팅을 하고 있는 업체 중 하나는 구글이다. 구글은 현 재 앤트로픽의 지분 14%를 소유하고 있으며, 총 투자액은 30억 달러를 초과했다는 소식이 유출됐다. 참고로 앤트로픽은 최근 자금 조달 라운드에서 35억 달러를 모금했으며, 이를 통해 회사의 가치는 615억 달러로 평가되었고 총 148억 달러를 모금한 상태다. 이 중에서 구글의 총 투자액은 30억 달러, 아마존의 투자금액은 약 80억 달러다.

> 이는 뉴욕타임스가 입수한 법원 문서에 따른 것인데, 정확한 지분 비율(14%)이 공개된 것 은 이번이 처음이었다. 해당 문서에 따르면, 구글이 앤트로픽에 대한 지분을 최대 15%까 지만 보유할 수 있다는 제한이 있기 때문에 사실상 할 수 있는 최선을 다한 것이다.

> 자체적으로 Gemini라는 언어모델도 있고 OpenAI 및 xAI와 프론티어급 성능 경쟁을 하는 구글이 이러한 베팅을 하는 것은 상당한 의미가 있다. 신생 AI 기업들에 투자함으로써 어떤 방향으로든 이익을 취할 수 있는 상황으로 만든 것이다. 직접 AI 기술을 개발하든, 유망한 스타트업에 투자하든 결국 승자가 되기 위한 절박함이라고 풀이된다.

2. 미국의 전방위적 견제

(1) OpenAl의 읍소 "DeepSeek 쓰지 말자"

미국의 조급함은 투자뿐만 아니라 견제 측면에서도 활발히 이뤄지고 있다. 2025년 3월 13 일, OpenAI의 글로벌 업무 부사장 크리스 리헤인(Chris Lehane)은 백악관 과학기술정책실 에 15페이지 분량의 서한을 제출했다. 자료에 따르면, DeepSeek의 모델(R1 추론 모델 포 함)이 중국 법률에 따라 사용자 데이터를 중국 정부에 제공해야 할 의무가 있어 안전하지 않다는 것이다.

그림 4. OpenAI가 DeepSeek를 '국가 통제 기업'이라 부르고 중국이 생산한 모델의 사용 금지 촉구



자료: TechCrunch, 미래에셋증권 리서치센터

실제로, DeepSeek의 개인정보 보호정책에 따르면 사용자의 개인 정보가 "중화인민공화국 에 위치한 보안 서버"에 저장된다고 명시되어 있다. 또한, OpenAI는 중국의 AI 개발자들 (DeepSeek 포함)은 저작권이 있는 데이터를 포함한 모든 데이터에 제한 없이 접근할 수 있다고 강조했다. 반면, 미국 기업들은 공정 사용(fair use) 원칙에 의해 제약을 받는다면서, 이러한 불균형이 지속된다면 "AI 경쟁은 사실상 끝난 것"이라고까지 주장했다. 중국의 AI 발전이 지정학적 위협으로 인식되고 있음을 분명히 말하는 것이다.

> 한편, 중국 정부는 DeepSeek을 "국가의 보물(national treasure)"로 간주하기 시작했으며, 이에 따라 회사에 대한 통제와 감독을 크게 강화하고 있다고 한다. 일례로 중국 정부는 일 부 DeepSeek 직원들에게 여권을 제출하도록 요구했는데, 핵심 기술 인력의 국외 이동을 물리적으로 차단하고 있음을 의미한다.

> 이런 사실을 들어. OpenAI는 DeepSeek가 "국가 보조금을 받고 국가가 통제하는" 기업이 라고 주장하면서 그들에 대한 미국 정부의 제재를 요청하기도 했다.

> 또한 지난 1월에 있었던 바이든 행정부의 수출 규칙에서 "티어 1"로 분류된 모든 국가에서 는 IP 도난 위험을 방지하기 위해 중국 제작 모델 사용을 금지해야 한다고 말했다. 사실상 "미국 진영 vs 중국 진영"의 구도로 AI 업계를 양분하자는 것이다. 이에 대해서는 AI Weekly에서도 몇 차례 전망해 온 것이다. AI와 국제 관계는 이렇듯 매우 "정렬"되어 있다.

표 2. 바이든 정부의 AI 칩 수출 통제 정책의 3등급 체계. 'OpenAI의 제안이 받아들여지면 1등급에 속한 국가들은 중국 AI를 못 쓰게 됨'

1등급	2등급	3등급
호주	브라질	ਨ ੋ국
벨기에	인도	러시아
캐나다	이스라엘	아프가니스탄
덴마크	말레이시아	벨라루스
핀란드	싱가포르	미얀마
프랑스	스위스	캄보디아
독일	오스트리아	중앙아프리카공화국
아일랜드	체코	콩고
이탈리아	그리스	쿠바
일본	인도네시아	에리트레아
네덜란드	케냐	아이티
뉴질랜드	룩셈부르크	이란
노르웨이	멕시코	이라크
대한민국	모나코	레바논
스페인	-	리비아

자료: Data Gravity, 미래에셋증권 리서치센터

(2) CSIS의 화웨이 보고서

지난 3월 7일, 이번에는 미국의 저명한 싱크탱크인 전략국제문제연구소(CSIS)는 화웨이에 대한 보고서를 발표했다. 이 보고서를 작성한 Gregory C. Allen은 화웨이, DeepSeek 등 중국의 AI 굴기에 대한 인사이트 있는 정보를 다수 열거했다. 먼저 DeepSeek는 분명히 심리전으로 생각할 게 아니라 유의미하고 무시할 수 없는 발전이라고 했다.

그러나 더 중요한 것은 화웨이의 최근 두 가지 성공이라고 강조했다. 하나는 TSMC가 대량 의 Ascend(화웨이의 AI 가속기) 칩을 제조하도록 하고 우회 수입한 것, 그리고 다른 하나 는 SMIC가 첨단 칩을 제조할 수 있도록 반도체 장비를 밀수한 것이라고 했다.

주: TPP는 Total Processing Performance의 약자. 엔비디아 GPU와 같은 AI 칩들의 TFLOPS 성능에다가 비트길이를 곱한 값이 특정 TPP를 넘으면 규제 대상으로 잡히게 됨

> 특히, 장기적인 위험 요소는 중국 오픈소스 커뮤니티가 DeepSeek 모델을 화웨이의 CANN(Compute Architecture for Neural Networks) 플랫폼에서 실행하는 데 관심을 갖 게 될 가능성이라고 했다. 이는 엔비디아의 CUDA 플랫폼 독점 상태에 위협이 될 수 있다 는 것이다.

> 우리 팀 역시 엔비디아의 진정한 해자는 실리콘이라기보다는 CUDA라고 강조한 바 있다. 미국이 중국에 대항해 절대적으로 지켜야 하는 비교우위가 소프트웨어라는 점에서 이를 면 밀하게 주시할 것으로 사료된다.

> 다만, DeepSeek는 현재 Ascend 칩에 관심이 없다. CSIS에 따르면, Ascend는 아직 "형 편없는 칩과 끔찍한 지원 소프트웨어"를 가지고 있기 때문이다. CUDA에 대항하는 화웨이 의 소프트웨어 생태계 구축은 앞으로 몇 년이 더 걸릴 것이나, 하드웨어 격차는 중국이 지 속적으로 추격 중인 것도 사실이다.

> 그리고 해당 보고서에서 가장 길게 강조된 내용은, "화웨이가 페이퍼 컴퍼니를 통해 미국 제재를 우회하여 TSMC의 AI 반도체 200만개 이상을 확보했다"고 대만 당국자들을 인용 하여 밝힌 점이었다. TSMC가 어센드 910B 로직 다이를 제조했는데, 화웨이는 페이퍼 컴 퍼니를 통해 TSMC의 첨단 노드 제조 능력에 접근할 수 있었다는 내용이다. 물론 화웨이는 "2020년 미국 상무부의 FDPR 개정안 시행 이후 TSMC를 통해 어떤 칩도 생산하지 않았 다"고 반박했다.

> 본인들이 주문하지 않았다는 것과 제3자를 통해 수입했다는 것은 엄밀하게 말하면 다른 내 용이기는 하다. 제3자 페이퍼 컴퍼니의 역할을 기꺼이 맡은 곳으로는 중국 기반 칩 설계업 체인 소프고(Sophgo)와 비트메인(Bitmain)과 같은 제3의 회사들이 관여했을 가능성이 제 기되었다.

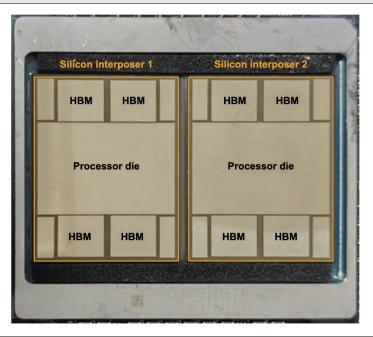
> 2024년 10월, TSMC는 자사가 생산한 칩이 화웨이 제품에서 발견되었다는 사실을 대만 정부와 미국 상무부에 통보했다. 이런 일이 발생할 수 있는 이유는, 미국의 해당 제재가 특 정 업체들만 거부하는 네거티브 방식인데다, 물리적 검증이 아니라 쉽게 속일 수 있는 '고 객 증명'에 의존해 제재에 구멍이 많았기 때문이다.

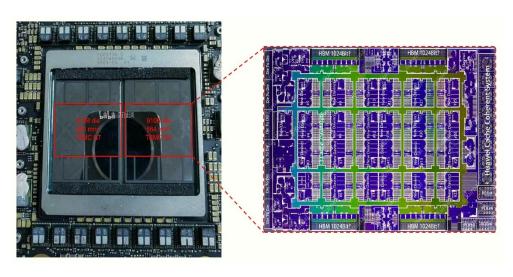
> 결국. 작년 11월부터 TSMC는 소프고를 포함해 중국 고객들에게 7나노 이하의 첨단 칩을 제공하는 것 자체를 중단한 상황이다. 중요한 점은 화웨이는 Ascend 910B 칩의 다이 200만개를 TSMC로부터 불법이지만 어찌됐건 확보했다는 점이다.

> 거기에 지지난 AI Weekly에서 밝힌 것처럼, 그들의 든든한 조력자 SMIC가 본인들만의 7 나노 이하 미세공정을 동원해 910B와 910C를 각각 30만 개, 10만 개를 생산할 계획임을 이미 밝힌 상태다. 작년에 생산된 910B 물량까지 합하면 SMIC로부터 생산되어 확보된 910C 칩의 수량은 40만 개가 된다. 그리고 이에 필요한 HBM2E 메모리도 삼성전자로부 터 비축해 놓은 상태다. 물론 이것도 통제 대상이지만 그 전에 비축한 것이다.

> TSMC로부터 받은 것과 SMIC에서 생산된 것을 합치면 결국 약 140만 개의 910C 가속기 를 생산하기에 충분해진다. 절대 무시 못할 양으로 보인다.

그림 5. 화웨이의 Ascend 910C는 두 개의 Ascend 910B를 연결한 것 엔비디아의 CoWoS-L과는 다름! 별도 실리콘 인터포저를 유기물 기판으로 연결하는 열위의 방식





자료: Lennart Heim, CSIS, 미래에셋증권 리서치센터

> 그러면 여기서 궁금한 점은 이들의 연산 능력이 미국에 비해서 얼마나 되는지다. 이에 대해 서 또 다른 미국의 유명 싱크탱크인 RAND 코퍼레이션의 정보 과학자 Lennart Heim이 밝힌 정보들을 살펴볼 필요가 있다.

> 910C는 FP16에서 약 800 TFLOP/s와 약 3.2 TB/s의 메모리 대역폭을 달성할 것으로 예 상된다. 이는 60% 더 많은 '로직 다이'의 면적을 사용하면서도 NVIDIA의 이전 세대 H100(2022년 출시) 대비 약 80% 수준의 성능에 불과한 것이다. 참고로 H100의 경우 FP16에서 약 1000 TFLOP/s와 3.9TB/s 메모리 대역폭을 가지고, 다이 면적은 814 mm²였다.

> NVIDIA의 B100/200 시리즈의 고급 패키징과 달리, Ascend 910C는 유기물 기판으로 연 결된 두 개의 별도 실리콘 인터포저를 사용하는 기술적으로 덜 정교한 접근 방식을 사용할 가능성이 높다. 위 그림에서 보듯이 각 인터포저에는, 하나의 프로세서 다이와 4개의 HBM 스택이 탑재되어 있다. 910B 다이는 각각 664mm²의 크기를 가지고, 두 인터포저는 물리 적으로 분리된 채로 유기물 기판 위에 나란히 배치되어 있다. 문제는, 다이 간의 통신 경로 가 길다는 점이다. 신호가 첫 번째 다이에서 인터포저를 통해 유기물 기판으로, 다시 두 번 째 인터포저를 거쳐 두 번째 다이로 이동해야 한다.

> 또한 유기물 기판의 배선 밀도는 실리콘 인터포저보다 훨씬 낮다. 실리콘 인터포저는 수 마 이크로미터 수준의 미세한 배선이 가능하지만, 유기물 기판은 수십 마이크로미터 수준의 더 굵은 배선으로 제한된다.

> 이러한 물리적 차이로 인해 NVIDIA 솔루션에 비해 다이 간 대역폭이 10-20배 낮아질 수 있다. 대역폭이 그렇게 낮다면, 실질적으로 하나의 칩이 아니다. 즉, 하나의 논리적 장치로 취급하는 대신 두 개의 연결된 가속기로 취급해야 하는 상황이 된다.

> 결론적으로는, 화웨이의 추격이 무섭기는 해도 기술적 격차가 생각보다 아직 상당해 보인다 는 점이었다. 올해 데이터센터에 투입될 NVIDIA의 B200과 비교했을 때, 910C는 계산 성 능이 약 3배 낮고, 메모리 대역폭이 약 2.5배 낮으며(HBM2E 기준; HBM3도 가능성 있 음), 전력 효율성이 훨씬 떨어진다고 RAND 코퍼레이션은 꼬집었다. 그리고 미국을 포함한 서방 진영은 올해 적어도 5배 많은 칩과 10-20배의 컴퓨팅 파워를 보유할 것으로 전망되 기 때문에, 총 컴퓨팅 면에서 미국의 우위는 여전히 강하다고 볼 수 있다.

> Ascend 910C 칩 140만 개의 성능은 개당 H100의 80%를 가정할 때 약 110만 개의 H100 GPU를 의미한다. 만약에 중국이 이 칩들을 중앙 집중화하게 된다면 어떤 일이 벌어 질까?

> xAI가 꿈꾸는 Colossus의 다음 버전이나 OpenAI 컨소시움의 Stargate 프로젝트에 버금 갈 정도의 무서운 대형 클러스터가 될 수 있다. 절대로 안심할 수 없다. 그리고 중국은 권 위주의 체제로 정책 집행 속도에서 타의 추종을 불허하는 일사불란함을 지닌다.

> 이것이 주시해야 할 핵심 사항이다. 중국이 많은 영역에서 경쟁력을 갖게 될 것임을 의미한 다. 이들이 정말로 중앙 집중화된 클러스터링에 성공한다면 GPT-6와 대결할 차세대 모델 을 만들어낼 가능성도 배제할 수 없다. 생각보다는 미국과 중국의 AI 경쟁이 지금보다 더 불이 붙을 것이라는 이야기다.

> 그래도 지난 1월 강화된 미국의 수출 통제로 그 몇 배가 되는 칩을 싹싹 긁어모으기가 매 우 어려워진 것은 사실이다. 이 측면에서 개별 AI 모델을 개발하는 것은 가능하더라도 그것 을 수백만 사용자나 수많은 AI 에이전트에게 동시에 서비스하는 대규모 인프라를 구축하는 데는 한계가 있을 것이라 보인다.

> 여전히 중국은 칩이 상당히 모자라다. 화웨이의 회장 런정페이든 DeepSeek CEO 량원평 이든 하나같이 컴퓨팅이 부족하다고 말하는 것은 진심이다. 아직까지 확보할 수 있는 엔비 디아의 칩에 대해 중국의 '패닉 바잉'은 당분간 계속되리라 생각된다.

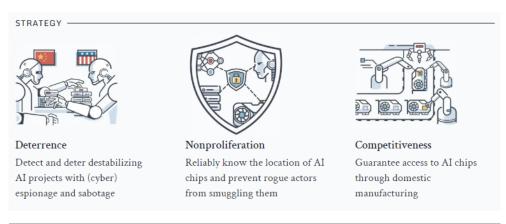
(3) 에릭 슈미트의 새로운 논문

미국의 견제는 수출 통제만으로 끝날 것은 아닌 것 같다. 정부 차원을 넘어 민간에서도 AI 의 발전에 대해 우려/경고하는 목소리가 점점 더 커지고, 또 많아지고 있다. 이번에는 구글 의 전성기를 이끈 CEO 에릭 슈미트가 책임을 맡았다.

며칠 전, 슈미트는 알렉산더 왕(Scale AI의 CEO)과 댄 헨드릭스(AI 안전센터 이사)와 함께 "Superintelligence Strategy"라는 논문을 게재했다. 초지능의 개발은 국가 간 패권 경쟁 을 심화시키며, 대국 간 갈등(제3차 세계대전)의 위험을 높일 수 있고, 이를 막기 위해 전 략적 접근이 필요하다는 것이다. 해당 논문은 초지능을 "거의 모든 인지 작업에서 인간보다 훨씬 뛰어난 AI"로 정의한다. AI 연구자들은 초지능의 등장이 멀지 않았다고 예상했다.

특히 그는 세 가지 전략(억제, 경쟁력, 비확산)을 제안했다. 이 중에서 억제는, 과거 핵무기 시대에 상호확증파괴(MAD, Mutually Assured Destruction) 전략이 있었던 것처럼 초지능 시대에도 유사한 전략이 필요하다는 논문의 핵심 주장이다.

그림 6. Superintelligence Strategy 논문에서 강조한 3가지 전략



자료: 논문 "Superintelligence Strategy", 미래에셋증권 리서치센터

- 억제(deterrence): 다른 국가의 과도한 AI 개발을 사이버 스파이, 사보타지 등을 통해 억제.
- 경쟁력 유지(competitiveness): 안정적인 AI칩 공급망 확보 및 드론·사이버 공격력 증강 등.
- 비확산(nonproliferation): 초지능이 위험한 행위자(테러집단 등)에 넘어가지 않도록 엄격히 관리.

에릭 슈미트를 포함한 저자들은 MAIM(=Mutual Assured Al Malfunction)이라는 새로운 개념을 제시했다. 한 국가가 AI 초지능을 독점해 군사 목적 등 위험하게 이용하려 할 경우, 다른 국가들이 이를 저지하기 위해 예방적인 '사보타주(sabotage)'를 감행해야 한다고 설 파했다. 예를 들어, 한 국가가 AI를 위험하게 독점하려 하면, 경쟁국들은 데이터센터 폭파, 연구자 방해, 사이버 공격 등 다양한 방식으로 AI 개발을 방해할 수 있어야 한다는 놀라운 아이디어다.

물론, 사보타주로 인한 전면전으로 비화되지 않도록 명확한 커뮤니케이션이 있어야 한다고 말했는데, 마치 러시아가 전쟁이 아니라 "특수군사작전"이라고 주장하는 것과 유사하게 들 린다. 결국 힘에 의한 논리가 될 가능성이 높고, 미국 진영과 중국 진영의 대결을 암시한 게 아닌가 하는 생각이 든다.

MAIM의 존재 의의는 슈퍼인텔리전스의 출현을 지연시켜 안전성을 확보할 시간을 벌기 위 함이다. 물론, MAIM은 실제 군사력(kinetic attacks)을 사용하는 대신, 사이버 공격 능력을 강화하여 AI 개발 경쟁을 억제하는 것을 목표로 한다고 밝혔다. MAIM 체제 하에서는 사이 버 공격이 일상적인 견제 수단으로 간주되어야 한다고도 했다. MAIM의 체제가 국제 사회 에 받아들여지든 그렇지 않든 사이버 공격 및 보안에 대한 관심은 지금보다도 훨씬 더 극 적으로 변화할 수도 있겠다는 생각이 들었다.

그리고 MAIM 전략은 마치 류츠신(刘慈欣) 작가의 SF 소설 "삼체" 3부작에 등장하는 "어 두운 숨" 이론과의 유사점을 생각하게 했다.

> 어두운 숲 이론은 우주는 수많은 문명으로 가득한 '어두운 숲'과 같으며, 그 속에서 생존하 기 위해 모든 문명은 다음의 두 가지 가정을 기반으로 행동한다는 논리다.

> 1. 생존은 모든 문명의 최우선 목표. (모든 문명은 스스로를 보호하기 위해 최대한 은폐하고, 조용히 숨어 지내며 위치를 노출하지 않으려 함)

> 2. 모든 문명은 다른 문명의 의도를 알 수 없으므로, 다른 문명을 발견하면 가장 합리적인 전략은 선제적 공격으로 제거하는 것. (자신의 존재가 노출되었을 경우, 상대방으로부터 공 격받을 가능성을 없애기 위해 반드시 먼저 공격해야 한다는 의무감에 시달림)

> 초지능은 '궁극의 무기'로 여겨질 수 있기 때문에, 상대 국가가 자신보다 먼저 개발할 경우 그 상대국의 의도가 악의적일 가능성에 대비하는 것이 전략적으로 합리적이라는 게 슈미트 의 생각이다. 즉, 두 가정 모두 상대가 우위에 서지 못하도록 하는 '선제적이고 암묵적 형 태'의 공격을 정당화하고 안정적 평화가 아닌 '불안한 균형' 상태를 전제로 한다. 이상적인 해결책은 당연히 아니지만 현재로서는 불안정한 균형이 최선의 전략 중 하나일 수도 있겠 다는 생각이 든다.

(4) 왜 이런 견제와 우려가 계속될까?

최근 들어 초지능에 대한 경계 목소리가 Al guru들 사이에서 계속 출현하는 것은 정말로 지능 폭발의 시대가 얼마 남지 않았다는 평가가 계속 나오고 있기 때문이다. (필자가 언급 한 최근은 지난 한 달 이내의 사실이고, 긴 호흡의 사건이 아니다)

Open Philanthropy(OpenAI의 후원자)가 지원하는 Forethought라는 비영리 연구기관은 지난 3월 11일, "지능 폭발에 대비하기"라는 글을 게시했다. 향후 10년 내에 초지능이 현 실화될 가능성이 높다면서 지금부터라도 사전 준비를 해야 된다고 강조하는 글이다.

Forethought에서는 1925년부터 2025년까지 100년 동안 발생했던 과학, 기술, 정치, 철 학 분야의 발전들이 이제 10년 단위로 압축해서 나올 것이라는 것으로 운을 뗐다. 하지만 이에 반해 인간의 의사 결정 능력은 그 속도를 따라가지 못하고, 기존 제도는 제대로 작동 하지 않을 가능성이 높다. 인간이 병목인 셈이다.

AI의 발전 속도가 압축 가속된다는 그들의 근거는 명확했다. 훈련에 사용되는 연산량은 연 간 약 4.5배 증가했고, 새로운 알고리즘은 동일한 연산량으로 더 높은 성능을 달성하도록 훈련 효율성을 매년 약 3배 높여왔다. 그리고 도구 사용, 프롬프팅 개선, 합성 데이터 등 훈련된 모델의 능력을 향상시키는 스캐폴딩이 연간 3배의 개선을 이끌어 왔고, 추론에 필 요한 연산량은 매년 1/10로 줄었는데 추론 작업을 위한 컴퓨팅 양은 연간 2.5배씩 커지고 있다.

그림 7. AI에서 발전을 이끄는 요인으로 AI 시스템의 인지적 노동이 급격히 증가 중임을 지적 알고리즘 효율성(연간 3배), 훈련 컴퓨팅 능력(연간 4.5배), 훈련 후 개선 기술(연간 3배)

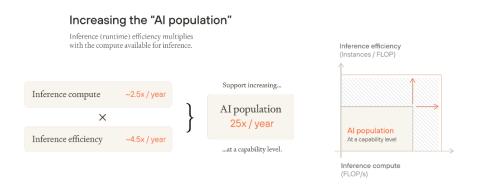
What is driving progress in AI? Cognitive labour from AI systems is increasing dramatically. Models get better and more efficient ... Algorithmic efficiency Capabilities ~3x / year in training Performance (e.g. accuracy and reliability) improves on established measures. Plus, qualitatively **new abilities** emerg Training compute ~4.5x / year Inference efficiency Post-training enhancements Training and development .. and we have more compute to run them Inference compute

자료: Forethought 논문 "Preparing for the Intelligence Explosion", 미래에셋증권 리서치센터

이러한 요인들의 결합으로, AI 시스템의 총 인지 능력은 매년 약 25배 이상 증가할 수 있다. 이는 마치 인구가 매년 25배씩 증가하는 것과 같은 엄청난 변화다. 게다가 Forethought에 서는 AI 연구 역량이 인간 연구 역량과 동등해지는 시점부터, AI 연구 역량은 매년 "최소 25배"이상 증가할 수 있다고 했다.

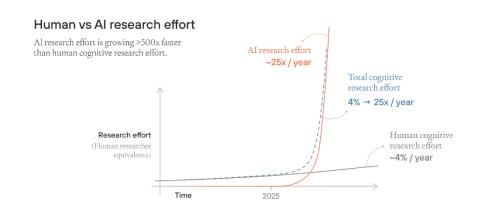
다시 말해, 1년 동안의 AI 연구 성과가 1년 후에는 25년, 2년 후에는 625년의 연구 성과와 맞먹는 것과 같은 엄청난 가속화가 된다. 물론 이것은 자체적인 피드백 루프를 지닌 "자가 개선 AI"를 고려하지 않은 추측이다. 인류 역사상 유례없는 지능의 폭발이 다가오고 있다.

그림 8. AI 인구(특정 능력 수준을 가진 AI 시스템의 수)가 폭발적으로 증가하고 있음 추론 컴퓨팅 자원(연간 2.5배)과 추론 효율성(연간 4.5배)의 곱으로 AI 인구가 연간 25배씩 증가



자료: Forethought 논문 "Preparing for the Intelligence Explosion", 미래에셋증권 리서치센터

그림 9. 인간의 인지적 연구 능력과 AI 연구 능력 간의 급격히 벌어지는 격차 AI 연구 능력이 인간보다 500배 이상 빠르게 성장… 2025년은 일종의 변곡점으로 강조됨



자료: Forethought 논문 "Preparing for the Intelligence Explosion", 미래에셋증권 리서치센터

그러나 초지능에 모든 문제를 맡길 수 없다. 따라서, 현재만 가능한 해결책들은 시간이 지 나면 실행하기 어려우니 지금 당장 행동해야 한다는 것이다. 이들이 내놓은 AGI를 대비한 구체적인 실천 방안 중 특히 흥미로운 것은 "책임감 있는 행위자"를 선정해 권한을 부여해 야 한다는 점이었다. "책임감 있는 행위자"는 다음과 같은 특성을 가진 개인 또는 집단을 의미하다.

- 1. Al 기술의 잠재적 위험과 사회적 영향에 대한 깊이 있는 이해를 바탕으로 윤리적인 의사 결정을 내릴 수 있는 능력을 갖춘 사람.
- 2. 단기적인 이익이나 경쟁 우위보다는 인류 전체의 장기적인 이익과 안전을 우선시하는 사람.
- 3. 다른 연구자, 기업, 정부, 국제기구 등과 협력하여 AI 기술의 안전하고 책임감 있는 개발을 추구하는 사람.
- 4. 자신의 연구와 의사 결정 과정을 투명하게 공개하고, 그 결과에 대해 책임을 지는 사람.

"책임감 있는 행위자"는 해당 소설에 등장하는 "면벽자(面壁者, Wallfacer)"와 상당히 유사 한 면모를 보인다. 두 개념 모두 인류의 미래를 좌우할 중대한 위기 상황에서, 특별한 권한 과 책임을 부여받은 소수의 개인을 지칭한다는 점에서 공통점을 갖는다.

*《삼체》에서 인류는 지구보다 훨씬 발전된 기술을 가진 외계 문명인 삼체 문명의 침공 위협에 직면. 삼체 문명은 "지자(智子, Sophon)"라는 양자 컴퓨터를 통해 지구의 모든 통신을 감청하고, 인간의 생각을 읽을 수 있음. 인류 는 "면벽 계획(Wallfacer Project)"을 수립하고, 4명의 면벽자를 선정. 면벽자는 삼체의 감시를 피하기 위해 자신. 의 진정한 의도와 계획을 숨기고, 오직 자신의 머릿속으로만 전략을 구상해야 함. 이들은 인류를 구원할 계획을 수립하고 실행하기 위해 막대한 권한과 자원을 부여받음.

AI 기술 발전이 가속화되면서. 우리는 점점 더 류츠신의 《삼체》 속 면벽자와 어두운 숲 이 론과 같은 상상 속 개념들이 현실에서 논의될 가능성에 직면하고 있다. 참으로 격변의 시대 를 앞두고 있다는 점을 실감할 수 있다. 그리고 이 시대를 이끌어 가고 있는 두 나라의 경 쟁이 격화될수록 상상은 현실이 될 수도 있지 않을까 하는 생각에 빠진다.

II. Agent War

1. 중국의 Manus Al

(1) 제2의 DeepSeek 쇼크?

미국의 대표적인 폐쇄소스 업체인 OpenAI에 대항해 중국의 오픈소스 진영의 새로운 플레 이어가 또 등장했다. 세계 최초의 완전 자율 AI 에이전트인 "Manus AI"를 만들어 낸 Monica라는 기업이다. 벌써부터 제2의 DeepSeek라는 별명을 얻어 미국 AI 업계에 큰 영향을 줄 것으로 각광을 받아 각종 hype가 발생했다.

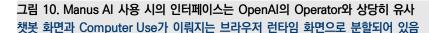
Monica는 "ChatGPT for Google" 브라우저 플러그인을 설계했고 1,000만 사용자를 빠르 게 확보하며 유명세를 탄 업체로, 설립자는 우한 출신의 1992년생 샤오 홍(肖弘)이다. 최고 과학자 지 이차오(季逸超) 등 젊은 남자들로 구성된 이 팀은 지난 3월 6일 Manus의 프리 뷰를 출시하자마자 미국인들을 충격에 빠트렸다. 중국에는 젊은 천재 개발자들이 참 많은 것 같다.

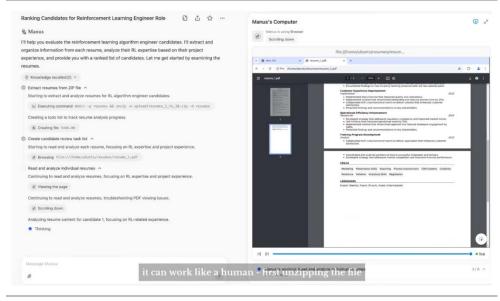
Manus AI는 사실 출시 전부터 "잠재적으로 AGI를 엿보는 것", "세계 최초의 진정한 범용 AI 에이전트", "Universal Hand" 등 다소 과장된 표현을 사용하여 대중의 기대치를 높였다. 어쨌든 이 홍보 전략은 먹혔고, 글로벌하게 엄청난 이목을 끌었다. 특히 Global Times, China Daily 등 중국의 언론들은 앞다퉈 "또 다른 스푸트니크 모멘트"가 왔다고 칭찬 일색 이었다.

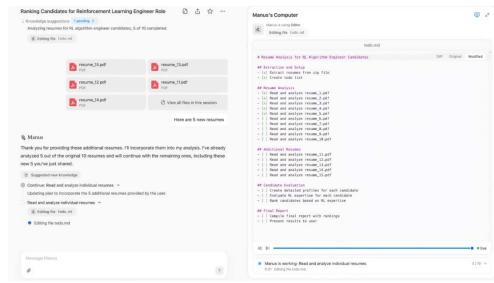
Monica측은 Manus AI의 대기자 명단을 숨기고 일부 인플루언서에게만 초기 접근 권한을 제공하여 희소성과 기대감을 조성했다. 다시 말해, 프리뷰 버전의 초대 코드를 받아야만 Manus 사용이 가능했다. 그런데 이 사용권이 중고 플랫폼에서 무려 최대 천만원 정도에 거래됐을 정도였다. 올해 안에 일부 모델을 오픈소스로 공개할 예정이라고 한다.

Manus는 사용자의 지시에 따라 스스로 작업을 계획하고, 필요한 도구를 가져오며, 문제를 해결하는 능력을 갖추고 있어 진정한 의미의 자율 AI 에이전트로 평가받고 있다. 특히 멀티 에이전트 시스템을 기반으로 하여 여러 작업을 동시에 처리할 수 있다는 것이 Manus의 핵 심이다.

텍스트 기반 인터페이스로 명령어를 입력하여 컴퓨터를 기본적으로 알아서 제어한다. 마치 OpenAl의 Operator나 앤트로픽의 Computer Use와 상당히 닮아 있고 파일 관리 및 프 로그램 실행 등을 스스로 수행할 수 있다.







자료: Manus AI, 미래에셋증권 리서치센터

뿐만 아니라 OpenAI의 Deep Research와 같이 높은 수준의 보고서를 작성하고 일부 영 역에서는 Manus가 더 나은 성능을 보인다고 보고하는 사용자도 꽤 있었다. GAIA 벤치마 크에서 Manus AI가 최고 성능을 달성하여 OpenAI에게 경쟁적인 위협이 될 수 있음을 시 사하기도 했다.

그러나, 벤치마크 중에서 GAIA만을 결과로 제시한 곳은 Manus가 처음이다. 다양한 벤치 마크를 통해 성능을 객관적으로 평가하지 않고, 특정 벤치마크 결과만 강조한 것이다. Manus AI가 해당 벤치마크에 과도하게 최적화되었을 가능성을 시사한다. 즉, 실제 다양한 작업에서는 벤치마크만큼 좋은 성능을 보이지 못할 수 있다. 특히, 대규모 언어 모델(LLM) 에 대해 꽤 오랜 시간 동안 회의적인 시각을 가진 얀 르쿤이 GAIA 벤치마크 작성에 참여 했다는 사실은 애초에 이 벤치마크 결과의 신뢰성에 대한 의문을 제기하기도 한다.

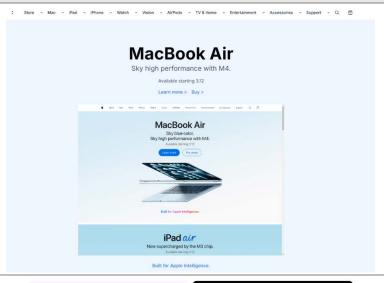
GAIA Benchmark OpenAI DeepResearch (pass@l) Level 1 67.9% 70.1% Level 2 57.7% 42.3%

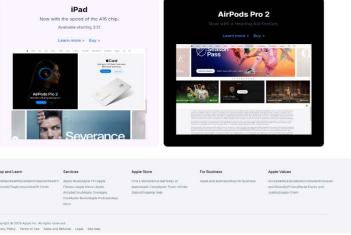
그림 11. GAIA 벤치마크 결과를 보여주는 차트… GAIA는 AI 모델의 능력을 평가하는 테스트

자료: Manus AI, 미래에셋증권 리서치센터

그래도, Manus의 활용 사례에 대해서는 이미 바이럴이 상당히 많이 되었다. 우리 팀이 살 펴본 결과, 연구 보고서 및 웹사이트 등을 자동으로 생성하는 기능이 굉장히 뛰어나다는 점 을 알 수 있었다.

그림 12. Manus AI를 사용해 애플 웹사이트와 거의 동일한 복제본을 약 1시간 만에 생성





자료: Manus AI, 미래에셋증권 리서치센터

> 특히. 아래와 같이 애플의 공식 홈페이지를 복제해 그대로 웹 페이지를 따라 만드는 능력은 머리를 한 대 맞은 듯한 느낌마저 들었다. 애니메이션까지 포함되어 있으며, 이미지를 제외 하고는 거의 같은 수준이었기 때문이다.

- * Manus AI로 가능한 사례들 예시
- * 이력서 검토 및 순위 제안: CEO 직무 지원자 20명의 파일(zip)을 제공했더니, 스스로 웹을 검색하고 메모를 작 성하며 각 지원자를 심층 분석한 후 요구사항에 따라 순위를 매김
- * 부동산 조사 및 보고서 작성 (다양한 기준에 따른 필터링)
- * 주식 상관관계 분석, 시각화 및 웹사이트 배포
- * 여행 보험 비교 분석
- * B2B 리드 생성 및 정보 제공
- * 생물 의학 등 다양한 분야 연구 및 분석
- * 맞춤형 소프트웨어 생성: 사용자가 원하는 앱 아키텍처를 설계하고, 구현 가이드를 작성하며, 실제 코드를 생성
- * SEO 보고서: 사용자가 URL만 제공하면 웹사이트를 자동으로 분석하여 포괄적인 SEO 보고서(웹사이트가 검색 엔진에서 더 잘 노출되도록 개선할 수 있는 요소들을 분석하고 제안하는 문서)를 생성

(2) Manus의 두 얼굴

Manus AI는 자체 클라우드 컴퓨터(리눅스 우분투) 환경에서 커맨드 라인을 활용한다. 사용 자가 요청한 작업을 여러 단계로 분해하고 Markdown 파일로 기록하며, Cloudflare Workers, Git 저장소 등 외부 도구/서비스를 활용하여 웹사이트를 완성한다.

오픈 소스 모델을 기반으로 하지만 놀라운 에이전트의 능력 덕에, 어떤 회사는 연간 \$6,000를 지불하고 있는 SaaS 사용을 종료하고, 그냥 Manus에게 웹 애플리케이션을 만 들어야겠다고 말하기도 했다.

즉. Manus AI가 중요한 이유는 사용자가 직접 소프트웨어를 만들 수 있게 하여, 기존 SaaS 시장에 큰 변화를 가져올 수 있기 때문이다. 개발 패러다임을 바꿀 수 있으며, 특히 SaaS 시장에 큰 영향을 미칠 수 있다고 사료된다. 금융, 부동산, 보험, 마케팅, 연구 등 다 양한 분야에서 각자의 업무 효율성을 높이기 위해 개인용 애플리케이션을 스스로 만들 수 있는 시대가 곧 오는 것이다. 특히 반복적인 작업을 자동화하는데 큰 도움을 주는 것이 이 러한 에이전트의 본질이다.

이처럼 Manus의 등장에 "세상이 바뀌었다", "AGI를 경험하는 것 같다"며 극찬하는 사람들 이 많다. 특히 정식 출시가 될 때 가격이 \$200 미만으로 유지된다면 ChatGPT Pro 구독에 위협이 될 것이다. \$2,000 이상의 에이전트 구독 모델을 출시하려고 하는 OpenAI에게는 부정적인 센티멘트다.

> 심지어는 처음에 vaporware일 것이라고 회의적이었던 사람들도 실제 사용 후 '존재론적 위기'를 느낄 정도로 놀랐다고 표현하기도 했다. 스스로 작업을 분해하고 필요한 도구를 스 스로 탐색하는 자율성과 지능이 높이 평가되었다. 특히 관련 자료가 거의 없는 분야에 대해 서도 스스로 학습하기 위해 자료를 찾는 '지식 습득 능력'과 적용 능력은 상당히 놀라운 부 분이었다.

표 3. Manus Al를 활용한 웹사이트 제작 프로세스

단계	주체	행동	세부 내용
1. 작업 지시	사용자 → Manus Al	작업 요청	예: "AI 관련 최신 뉴스 웹사이트 제작"
2. 작업 분해	Manus Al	작업 단계 세분화	콘텐츠 수집 (AI 뉴스 검색) 구조 설계 (HTML, CSS) 기능 구현 (JavaScript) 디자인(UI/UX) 배포 (Cloudflare Workers) 버전 관리 (Git, 필요 시)
3. 리눅스 커맨드 라인 실행	Manus Al	각 단계별 명령어 실행	예: mkdir, cd, touch, nano/vim, git, curl/wget
4. Markdown 파일 기록	Manus Al	작업 과정 및 결과 문서화	진행 상황, 실행 결과, 사용된 명령어, 파일 목록, 작업 로그, 결과 요약, 코드 저장
5. 외부 도구/서비스 활용	Manus Al	다양한 서비스 통합	Cloudflare Workers: 서버리스 컴퓨팅 플랫폼 Git 저장소: 코드 저장/버전 관리 기타 API(필요 시): OpenAl, ElevenLabs 등
6. 웹사이트 완성	Manus Al → 사용자	최종 결과물 제공	모든 단계 완료 후 완성된 웹사이트 제공

자료: 미래에셋증권 리서치센터

그러나 긍정적인 평가만 있었던 것은 아니다. 오히려 실망스럽게 들리는 것은, 기본적으로 Manus가 미국 업체인 앤트로픽의 Claude 3.7 Sonnet 모델을 기반으로 했다는 점이다. 특히 Manus AI는 Claude Code CLI(Claude 모델을 터미널 환경, 즉 명령줄 인터페이스에 서 실행할 수 있게 하는 도구)를 모방하는 것이 아니라, 단순히 API를 호출해서 Claude Code CLI의 기능을 사용했다.

모방이나 흉내를 낸 것이 아니라 그저 Claude의 기능을 가져다가 쓴 것에 불과하다. Manus AI는 Claude Code CLI로부터 받은 결과를 사용자에게 적절한 형태(예: 웹사이트, 텍스트, 코드 등)로 제시하는 것뿐이다. 심지어, 미국의 Claude 모델은 중국에서 차단되어 있기 때문에, Manus AI 웹사이트는 중국에서 접속할 수조차 없다. 이는 Manus AI가 중국 시장을 목표로 개발되었음에도 불구하고, 자체적인 기술적 한계로 인해 시장 접근성이 제한 된다는 것을 의미한다.

어떤 X 유저는 Manus AI의 핵심인 샌드박스의(Manus AI가 웹 브라우저를 제어하고, 웹사 이트와 상호작용하며, 작업을 수행하는 원격 환경) 런타임 코드를 추출한 결과를 공개했는 데, Manus가 Claude Sonnet 모델을 기반으로 하고, 29개의 도구로 이뤄져 있다고 밝혔다.

기본적으로 새로운 기술이 아닌, 기존 기술들을 짜깁기한 것에 불과하다는 뉘앙스였다. 심 지어 해당 유저는 본인도 Claude Sonnet 3.7 모델을 사용하여 Manus의 코드를 매우 쉽 게 복제했다며 그 증거물을 제시하기도 했다. 누구나 쉽게 구현할 수 있다는 것이다. 그리 고 이러한 유출 사태는 그만큼 Manus의 보안 취약점을 드러낸 것이기도 하다.

> 또한 Manus 측은 Claude Sonnet 모델을 기반으로 사후 훈련(post-train)했다고 주장했 었는데, 사실 Sonnet 모델을 그대로 사용하고 있을 뿐 추가적인 훈련을 하지 않았다는 분 석도 있었다. 즉, 이러한 것들로 보아 Manus AI가 "래퍼(wrapper)"일 뿐이라는 주장에 조 금 힘이 실리기도 한다. 여기서 래퍼는 기존 기술(예: LLM)을 활용하여 새로운 기능이나 서 비스를 제공하는 소프트웨어 구성 요소를 말한다.

> 그리고 Manus 사용을 위해 필요한 초대 코드 발송이 제한되기도 했다. 물론 높은 인기 때 문에 생기는 서버 과부하일 가능성도 있다. 공격적인 마케팅을 펼친 탓에, 감당하기 어려울 정도로 많은 사람이 몰렸을 것이기 때문이다. 그러나, 애초에 운영 비용이 매우 비싸기 때 문이 아닌가 하는 생각이 든다. AI 개발자 커뮤니티에 따르면 Manus AI의 작업 당 비용은 약 2달러로 추정된다.

> 높은 가격의 원인으로 지목되는 것은 바로 기반이 되는 Claude 3.7 Sonnet 모델의 존재 때문이다. 애초에 이 모델은 매우 비싸고, 사용량 제한이 걸려 있는 모델이다. 즉, Manus AI가 동시에 처리할 수 있는 작업 수에 제한을 가하게 된다. 또한 Manus AI가 각 작업을 처리하는 데 드는 비용을 증가시키고, 작업에 더 많은 시간이 걸리게 하는 요인이다.

그림 13. 앤트로픽의 Claude 모델을 사용하고 있음을 인정한 Manus의 CTO 물론 이러한 조합 능력도 애플리케이션 레벨에서의 성공으로 봐야 함



We use Claude and different Qwen-finetunes. Back when we started building Manus, we only got Claude 3.5 Sonnet v1 (not long-CoT, aka reasoning tokens), so we need a lot of auxiliary models. Now Claude 3.7 looks really promising, we are testing internally, will post updates!

12:20 AM · 3/10/25 · **11K** Views

자료: Yichao Ji – Manus Al를 만든 CTO, 미래에셋증권 리서치센터

그래서, 세상을 바꿀만 한 기술이라는 hype에 동조하는 것에는 반대한다. 적어도 DeepSeek는 자체 개발한 언어 모델을 기반으로 하며, 저렴한 비용과 광범위한 접근성을 제공해 큰 주목을 받았다. 반면, Manus AI는 타사 모델(Claude)을 기반으로 하고, 높은 비 용과 제한된 접근성을 가지고 있다. 2번째 DeepSeek 모멘트라고 부르기에는 부적절하다.

> 그럼에도 아예 무의미한 것이라 평가 절하할 필요는 없다고 본다. Cursor와 퍼플렉시티도 "래퍼"이기 때문이다. 이들은 자체 모델 없이 래퍼 형태로 서비스를 제공하지만 매우 성공 적인 기업들로 가치를 평가받고 있다.

> Manus의 성공은 DeepSeek처럼 모델 구축 레벨에서의 성공이 아니라 애플리케이션 구축 레벨의 성공이라 할 수 있다. 성공은 성공이고, 그것을 중국이 또 한 번 해냈기에 더 큰 이 슈가 되고 있다는 정도로 정리할 수 있겠다.

2. OpenAl의 카운터

(1) Agent 관련한 API와 SDK 출시

에이전트계의 또 다른 차이나 쇼크가 출현하자, 일주일도 안 되어 미국의 대답이 나왔다. OpenAI에서 드디어 에이전트의 대대적인 상업화를 위한 API를 제공한다고 선포했다. OpenAI는 이미 Deep Research와 Operator라는 두 가지 강력한 에이전트 솔루션을 출 시한 적 있는데 한화 30만원이 넘는 유저들만 쓸 수 있었다.

그런데 이제 개발자들이 자신만의 에이전트를 만들 수 있도록, 이 도구들을 API 형태로 내 놓은 것이다. 여기서 AI 에이전트란 특정 목적을 위해 설계된 AI 시스템으로, 단순히 질문 에 답하는 것을 넘어 웹을 검색하고, 파일을 분석하며, 심지어 컴퓨터를 제어하여 실제 작 업을 수행하는 것이다. 예를 들어, OpenAI가 시연한 "개인 스타일리스트 에이전트"는 사용 자의 취향을 이해하고, 최신 트렌드를 검색하며, 실제로 쇼핑 사이트에서 옷을 찾아 구매까 지 도와줄 수 있었다.



자료: OpenAI, 미래에셋증권 리서치센터

OpenAI가 지난 11일 내놓은 것들은 AI 에이전트 개발을 훨씬 쉽고 효율적으로 만들어주 는 것들이다. 새로운 Built-in 도구 3가지, 그리고 새로운 API인 Responses API, 마지막 으로 오픈소스 SDK였다. 여기서 오픈소스 SDK의 알맹이는 Responses API이고 이 API 의 핵심은 3가지 도구로 이뤄진다.

표 4. 에이전트 구축을 위한 세 가지 내장 도구 (Built-in tools)

도구 이름	주요 기능	세부 사항
웹 검색 도구	인터넷 정보 접근 및 검색 (공개 데이터 검색)	 ChatGPT Search와 동일 내부적으로 fine-tuned GPT-4o 모델 사용 검색된 정보의 출처를 명확히 제시
파일 검색 도구	업로드된 문서 기반 RAG 수행 (개인 데이터 검색)	 문서를 청크 단위로 임베딩 메타데이터 필터링: 파일 속성 기반 관련 정보 필터링 직접 검색 엔드포인트: 모델 없이 벡터 저장소 직접 검색
컴퓨터 사용 도구	작업 자동화 및 애플리케이션 구축	 API를 통해 Operator 사용 가상 머신 제어 GUI 기반 레거시 애플리케이션 제어 ChatGPT의 Operator와 동일한 모델 사용

자료: 미래에셋증권 리서치센터

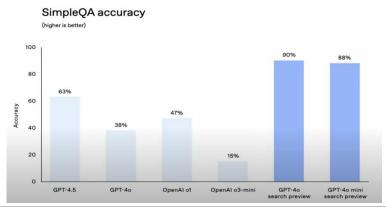
이 세 도구들은 에이전트가 필요로 하는 핵심 능력(정보 접근, 지식 처리, 행동 실행)을 포 괄한다. 외부 정보 접근과 실제 행동 실행 능력을 갖춤으로써 에이전트의 자율성이 크게 높 아지게 되고, 질의응답을 넘어 실제 업무 환경에서 가치를 창출할 수 있는 에이전트를 구축 할 수 있게 된다.

표 5. OpenAI가 예시로 든 AI 에이전트(개인 스타일리스트)를 구현하는 단계 및 필요한 내장 도구

단계	도구	기능	작동 방식	예시
1. 기본 설정	-	Al 비서 역할 및 모델 설정	• AI 비서에게 역할 부여 • 질문(컨텍스트) 제공 • AI 모델 선택	"너는 개인 스타일리스트야" "최신 트렌드는 뭐야?" GPT-4o 선택
2. 사용자 취향 파악	파일 검색 도구	사용자 정보 활용	벡터 저장소에서 사용자 정보 검색개인화된 스타일 정보 분석	"Elon은 평소에 어떤 옷을 입지?" → 저장소에서 Elon의 스타일 정보 검색
3. 최신 정보 파악	웹 검색 도구	인터넷 정보 검색	인터넷에서 최신 트렌드 검색사용자 위치 기반 매장 정보 탐색벡터 저장소 정보와 조합	Kevin이 도쿄에 있을 때 "나한테 어울리는 재킷 찾아줘" → 취향과 도쿄 매장 정보 조합 추천
4. 구매 대행	컴퓨터 사용 도구	온라인 쇼핑 실행	도커 컨테이너에서 웹 브라우저 조작마우스 클릭, 키보드 입력 수행온라인 쇼핑 과정 자동화	"Kevin에게 검은색 파타고니아 재킷 찾아줘" → 웹사이트에서 검색, 장바구니 담기, 결제 (사용자 최종 동의 필요)

자료: 미래에셋증권 리서치센터

그림 15. SimpleQA 벤치마크 정확도로 알아보는 웹 검색 도구의 효과 웹 검색 기능이 통합된 GPT-4o 및 GPT-4o mini가 일반 모델들보다 훨씬 높은 정확도



자료: OpenAI, 미래에셋증권 리서치센터

그림 16. 컴퓨터 사용 도구(CUA)가 다양한 벤치마크에서 이전 최고 성능(SOTA)보다 우수한 결과 OpenAI의 컴퓨터 사용 도구가 컴퓨터 인터페이스를 조작하는 능력이 크게 향상되었음을 증명

Benchmark	Computer use (universal in	terface)	Web browsing agents	Human
	OpenAl CUA	Previous SOTA	Previous SOTA	
Computer use				
OSWorld	38.1%	22.0%	-	72.4%
Browser use				
WebArena	58.1%	36.2%	57.1%	78.2%
WebVoyager	87.0%	56.0%	87.0%	

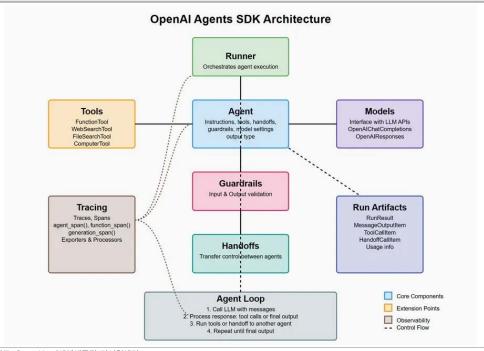
자료: OpenAI, 미래에셋증권 리서치센터

즉, 3가지 핵심 도구들로 무장한 Responses API는 개발자들이 똑똑한 AI 비서를 만들 수 있게 해주는 도구 모음이다. "벡터 저장소"에서 사용자 정보를 탐색하고 인터넷 정보도 탐 색한 다음 "도커 컨테이너" 안에서 컴퓨터를 조작해 일을 수행하는 것이다.

그런데 고객 지원 에이전트와 같이 더 복잡한 애플리케이션을 구축하는 경우, 개인 스타일 리스트와 같은 하나의 에이전트만으로는 충분하지 않을 수 있다. 환불을 처리하는 에이전트, 고객 지원 FAQ 쿼리에 응답하는 에이전트, 주문 및 청구를 처리하는 에이전트 등 여러 에 이전트가 필요할 수 있다.

에이전트들의 오케이스트레이션 작업을 하기 위한 기능으로 재탄생한 것이 Agents SDK다.

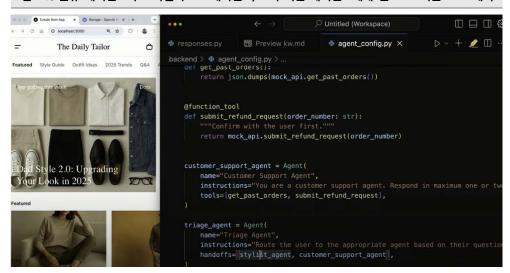
그림 17. OpenAl의 Agents SDK의 전체 구조 에이전트는 핵심 구성 요소로 지시사항, 도구, 핸드오프, 가드레일, 모델 설정, 출력 유형 등을 관리



자료: OpenAl, 미래에셋증권 리서치센터

> Responses API(블록)만으로도 에이전트를 만들 수 있지만, Agents SDK(설명서)를 사용하 면 더 쉽고 빠르게 만들 수 있다. Responses API가 레고 블록이라면 Agents SDK는 레고 설명서다. 여러 에이전트를 만들고 연결, 관리하는 방법을 제공한다. 예를 들어, 쇼핑 지원 에이전트가 반품 요청을 받으면 자동으로 고객 지원 에이전트에게 작업을 넘기는 "핸드오 프"기능을 구현할 수 있다.

그림 18. 분류 에이전트가 스타일리스트 에이전트/고객 지원 에이전트에게 핸드오프하는 코드 예시



자료: OpenAI, 미래에셋증권 리서치센터

핸드오프란 마치 콜센터에서 상담원이 전화를 다른 부서로 연결해주는 것과 비슷하다.

각 에이전트는 특정 작업을 전문으로 처리하고, 필요한 경우 다른 에이전트에게 도움을 요 청한다. 이 때 전체 대화는 동일하게 유지되지만, 백그라운드에서는 도구가 교체되는 형태 다. 이를 통해 대화를 분류하고 대화의 각 부분에 맞는 컨텍스트를 로드할 수 있다고 했다.

* 분류 에이전트(Triage Agent): 사용자의 요청을 받고 어떤 에이전트에게 보내야 할지 결정. 여기서는 "반품" 요 청이므로 "고객 지원 에이전트"에게 넘김.

* 고객 지원 에이전트(Customer Support Agent): 분류 에이전트로부터 요청을 받아 사용자의 과거 주문 내역을 확인하고 반품 절차를 안내.

(2) 2025년은 에이전트의 해

그런데 OpenAI는 이 SDK를 오픈 소스 프레임워크로 출시한다고 했다. 지금 바로 'pip install openai-middle-agents'를 실행하여 설치할 수 있다는 것이다.

그림 19. 파이썬에서 간단하게 설치 가능한 OpenAI의 Agents SDK

```
pip install openai-agents
Hello world example:
  from agents import Agent, Runner
  agent = Agent(name="Assistant", instructions="You are a helpful assistant")
result = Runner.run_sync(agent, "Write a haiku about recursion in programming.")
  print(result.final_output)
```

자료: OpenAl, 미래에셋증권 리서치센터

하지만 여기서 헷갈려서는 안 되는 것이, 핵심 기술(이 경우 AI 모델)은 비공개로 유지하고 유료 API로 제공된다는 점은 변함없다. 에이전트를 구축하는 데 필요한 소프트웨어 개발 도구(SDK)를 오픈소스로 제공한다는 뜻이다.

그림 20. Responses API로 기본 에이전트를 만들고 실행하는 최소 코드는 단 4줄이면 됨 파일/웹 검색, 함수 호출, 구조화된 출력 등 기능이 매개변수 하나로 제공

```
1 curl https://api.openai.com/v1/responses \
  -u :$OPENAI_API_KEY \
3 -d model="gpt-40" \
4 -d input="What is the capital of France?"
```

자료: OpenAI, 미래에셋증권 리서치센터

한편, OpenAI는 Agents SDK를 공개하기 전에 샌프란시스코의 AI 개발자들을 초청해서 밤샘 해커톤(Hack Night)을 개최했었다. 아래 표에서 보듯, 거기서 개발자들이 만든 흥미로 운 AI 에이전트 사례들로 어떤 미래가 펼쳐질지 잠깐 상상해 볼 수 있다.

표 6. OpenAl의 "Agents SDK" 해커톤: Al 에이전트에 관한 실리콘 밸리 개발자들의 밤샘 '전야제'

에이전트 이름	유형	주요 기능
Bansky Al	컴퓨터 사용 에이전트	HTML 캔버스(웹페이지 그림판)에 그림을 그리는 AI 화가
Touch Grass	컴퓨터 사용 에이전트	현재 시간과 위치를 확인해서 주변의 자연 명소를 추천
Real Estate Agent	웹 검색 에이전트	인터넷에서 임대 아파트를 검색하여 부동산 중개인 역할 수행
Memory Agent	Agents SDK + 웨어러블 기기	Al Compass 웨어러블 기기와 연결하여 사용자의 일상생활 습관이나 고민 파악
Agent Polling Simulator	Swarm(다중 에이전트)	다양한 정치적 성향의 AI 에이전트로 정책/뉴스에 대한 반응 시뮬레이션
Salary Seeker	RapidAPI 활용	Glassdoor에서 특정 직업의 연봉 정보를 검색하는 AI 헤드헌터
Multi-Agent Brainstormer	브레인스토밍+평가 에이전트	여러 에이전트가 아이디어를 제안하고 평가하여 최종 전략 도출

자료: OpenAI, 미래에셋증권 리서치센터

위 표에서 적힌 것보다 훨씬 많은 개발 사례들이 앞으로 쏟아질 것이다. OpenAI에서도 '모 든 사람이 소프트웨어를 만들 수 있게 되는' 대중화 효과의 잠재력을 강조하고 있다.

> 그럼에도 우리 팀이 강조하고 싶은 것은 이러한 앱 개발을 모든 사람이 하지는 않을 것이 라는 점이다. 거의 모든 사람이 요리를 할 줄은 알지만 여전히 외식 산업은 성행하는 것처 럼 말이다.

> 그래도 기존의 앱 개발의 어려움이 파인 다이닝이라면 이제는 김밥 프랜차이즈 정도로 급 격하게 내려오고 있는 것 같아 보인다. 물론, 김밥집들이 문을 닫을 때까지 김밥은 잘 팔릴 것이나 요리 인구는 급격하게 늘어나 현재의 앱 생태계 지형이 완전히 변할 것이다.

- * 글로벌 온라인 간편결제 시스템 업체인 Stripe가 공식 발표한 에이전트가 일하는 프로세스. Agents SDK로 여러 개의 AI 에이전트(자동 청구서 작성 에이전트 + 상시 이메일 고객지원 에이전트)를 만들어 일상 업무(청구서 발급, 이메일 고객지원)를 자동화해 사람의 업무를 크게 줄임
- ① 이메일 읽기: AI 에이전트가 고객의 이메일 내용을 이해 (예시: 카드 정보 업데이트 요청임을 파악)
- ② 고객 요청 구체화: 이메일의 요청이 정확히 무엇인지 파악(예시: 카드 정보 변경 요청이 "유효기간 변경"인지, "카드를 잃어버려 재발급"인지 판단)
- ③ 회사 정책 확인: 요청 사항을 처리할 때, 회사 내부 규정을 확인 (예시: 요금제 변경을 요청했을 때 요금제 변 경 가능 조건을 확인)
- ④ 처리 방법 결정: 고객의 요청과 회사 규정에 따라 어떻게 처리할지 결정 (예시: 카드 정보 업데이트를 진행하 거나, 추가적인 정보를 고객에게 요청하거나, 직원에게 연결할지 결정)
- ⑤ 자동 처리 혹은 직원 연결: 간단한 요청(카드 유효기간 업데이트 등)은 에이전트가 직접 자동 처리, 복잡한 요 청(새로운 요금제 상담 등)은 실제 직원에게 연결하여 처리
- ⑥ Stripe 결제 시스템 연동: 결정한 조치를 실행하기 위해 Stripe 시스템과 연동하여 카드정보를 바꾸거나, 요금 제를 바꾸는 등의 작업을 진행
- ② 고객에게 답장: 처리된 결과를 다시 이메일로 송부 (변경된 카드 정보, 요금제 내용 등을 고객에게 알림)

2025년은 에이전트의 해가 될 것이다. 만약 ChatGPT 단계에서 AI 혁명이 멈춘다면 투자 자들이 관심 갖고 있는 AI 관련 주식들은 폭락을 면치 못할 것이다. AI 주식들의 높은 밸류 에이션은 AI의 대중화를 감안하고 있고, 이것이 보장되려면 현실의 세상에서 사람들의 일 상을 대신해주는 에이전트가 있어야 한다.

실제로 며칠 전, OpenAI의 최고 제품 책임자(CPO) Kevin Weil은 "OpenAI는 2025년을 AI가 프로그래밍에서 인간을 영구적으로 능가하는 해로 전망하고 있다"라고 밝혔다.

이처럼 Manus의 등장으로 OpenAI가 원래 의도한 타이밍은 아니었을지 모르지만, 에이전 트 구축에 대한 API는 출시됐다. 새로운 시대가 열리려고 한다. 지금까지의 AI는 주로 채 팅 위주였지만(chat in, chat out), 앞으로는 행동 중심(data in, action out)으로 변화할 것 이다.

III. 미국의 해자인 자율주행 소프트웨어

1. 테슬라 FSD의 중국 진출이 갖는 의미

(1) 자율주행 시장은 과점 시장이 될 것

AI 가속기와 언어 모델을 포함한 여러 측면에서 중국의 추격이 굉장히 괄목할 만하고 곧 따라 잡힐 수도 있는 우려가 생기기도 한다. 그렇다면 반대로 중국이 당장 따라잡기 힘든 미국 진영의 기술은 무엇이 있을까? 그러한 해자를 지닌 기업을 찾는 것이 AI 투자의 나침 반 설정에 도움이 되리라 판단한다.

중국은 기본적으로 기술의 난이도 측면에서 일종의 보호 무역주의의 패턴을 따라왔다. 과거 스마트폰, 2차 전지 및 전기차 모두 비슷했다. 기술 구축 과정에서 서방 진영에게 시장을 열어주어 기술을 먼저 흡수하고 이를 고도화한 뒤 대체가 가능한 상황이 되면 자연스럽게 국산 기업을 밀어주는 형태가 되었다. 약 10년 전부터 LG화학이 번번이 중국 정부의 EV 배터리 보조금 목록에서 제외됐던 사례를 쉽게 떠올릴 수 있다. 자국의 배터리 산업을 성장 시키기 위한 정책의 일환이었고, 또한 스스로 밸류체인을 형성하더라도 글로벌 경쟁력이 있 다고 판단한 시점이었을 것이다.

약 3주 전 테슬라의 자율주행 시스템인 FSD가 중국에 부분적으로 오픈되었다. 그런데 시 장의 관심은 놀라울 정도로 잠잠하다. 그러나 중국 규제 당국이 이를 승인했다는 점과 테슬 라의 실적/경제적 측면에서 이것은 의미가 깊다. 규제 당국이 자율주행 기술에 있어서는 여 전히 "미국 기업"에게 배울 것이 있다고 판단해 테슬라의 요청에 응한 것일 가능성이 있다.

또한 테슬라로 하여금 자율주행 기술에 관한 대중들의 인식을 개선시키고, 새롭게 파생되는 시장 및 산업에 관한 테스팅 베드로 활용할 수도 있을 것이다. 물론, 테슬라에게 어느 정도 과실을 내어주고 일종의 사회 실험을 하게 하는 동안 자국 기업들이 빠르게 그들의 방식을 모방하고 기술을 갈고 닦도록 권장할 것이다.

다시 말해, 아직까지는 자율주행 소프트웨어의 격차가 있고 테슬라에게는 시간이 있다. 왜 그런 것일까? 자율주행 모델은 언어 모델보다 복잡성이 훨씬 높다. 첫째, 운전 시나리오가 극도로 복잡하며, 둘째, 사고 발생 시 경제적 손실이 매우 클 수 있어 안전 요구사항이 매 우 엄격하기 때문이다. 자율주행은 인공지능과 현실 세계의 인터페이스를 대표하는 최전선 기술로 볼 수 있다.

> 편리하고 완벽한 자율주행을 가진 고객들은 안전성을 희생하면서 굳이 다른 회사의 불완전 한 제품을 구매하고 싶지 않을 가능성이 높다. 중국의 자율주행 기술을 선도하는 한 업체인 Momenta의 CEO가 "전세계 자율주행 시장은 과점 시장이 될 것이다"라고 말한 이유다.

> 즉, 편리함도 중요하지만 생명과 관련한 기술이기 때문에 그만큼 최고 성능을 가진 소수의 업체들이 차지할 시장이라는 것이다. 이는 테슬라가 FSD를 통해 구축하고 있는 경쟁 장벽 의 핵심이기도 하다.

(2) FSD의 중국 진출 배경 및 데이터 처리 전략

테슬라가 중국에서 FSD를 운영하기 위해서는 역시 현지 데이터 수집이 핵심이다. 테슬라 의 차량 시스템에서는 '그림자 모드'로 데이터 수집 및 타당성 검증을 시작한다. 그림자 모 드는 사용자가 운전하는 동안 FSD 시스템이 어떤 결정을 내릴지 추적하지만 실제로 차량 을 제어하지는 않는 방식으로, 시스템 검증에 중요한 데이터를 수집한다.

또한 테슬라는 데이터 저장소 현지화를 위한 데이터센터를 중국에 건설하고 있고, 중국 본 토 시장에서 판매되는 차량에서 발생하는 모든 데이터는 중국 내에 저장할 계획이다. 즉, 전체 "데이터 폐쇄 루프"가 중국 대륙에 있어야 한다는 말이다. 이러한 접근은 중국의 데이 터 주권 정책을 준수하기 위한 필수적인 조치다.

실제로 테슬라는 몇 년 전부터 FSD의 중국 시장 정착을 촉진하기 위해 약 20명의 현지 운 영 팀을 중국에 구성했고, FSD 알고리즘 학습을 위해 수백 명의 데이터 라벨링 팀을 중국 에 설립한 적이 있다. 데이터 라벨링은 자율주행 시스템 개발에 있어 핵심적인 과정 중 하 나로, 중국인 라벨러들은 현지 도로 환경과 교통 규칙에 대한 이해를 바탕으로 더 정확한 데이터 주석을 제공할 수 있다.

그런데, 테슬라가 중국에서 데이터 수집 자격을 취득하더라도 모델 훈련 작업도 현지에서 수행해야 한다는 점은 큰 병목 사항으로 보였다. 미국과 동등한 FSD 훈련 효율을 얻으려 면 중국에 Dojo와 유사한 슈퍼컴퓨팅 센터를 구축해야 하는데, 이는 상당한 개발 주기와 비용이 필요하기 때문이다. 애초에 중국에 수만 개의 GPU를 수입해 올 수도 없는 처지다.

다만, Dojo급 슈퍼컴퓨팅 센터가 중국에 필요하지는 않다. 미국 내 테슬라의 슈퍼컴퓨터 클 러스터인 Cortex의 Dojo에서 FSD의 일반 세계 모델(General World Model)을 위한 기본 모델을 만들고, 중국 현지 데이터로 파인튜닝해 이를 극복하는 방식일 것이기 때문이다. 이 접근법은 기본 주행 지능을 유지하면서 중국 특유의 도로 환경에 맞게 최적화하는 효율적 인 방식이다.

그리고 테슬라는 이번에 그 방식이 가능함을 증명해냈다.

2. 중국판 FSD의 등장과 사용자 경험

(1) "중국판 FSD"의 출시와 특징

2025년 2월 25일, 테슬라는 중국 고객을 위해 FSD 기능과 유사한 소프트웨어 업데이트 를 단행했다. 흥미롭게도 테슬라는 중국에서 FSD라는 용어를 직접 사용하지 않고 "Autopilot Automatic Assisted Driving on Urban Roads"(도시 도로 Autopilot 자동 보 조 운전)이라는 다소 복잡한 네이밍으로 표현했다. 중국 규제 당국의 승인 과정을 거치며 발생한 시장 전략과 관련된 것으로 보인다.

그림 21. 테슬라의 중국 내 자율주행 SW(2024.45.32.12) 버전이 단계적으로 배포되기 시작



자료: Tesla, 미래에셋증권 리서치센터

그러나, FSD라는 세 글자는 테슬라 차이나 공식 홈페이지 어디에도 언급돼 있지 않지만, 기능 설명은 북미 FSD와 매우 유사하다. 실제로 테슬라 차이나의 고객 서비스 담당자가 최근 업데이트된 기능을 "중국 내 FSD로 이해하면 된다"고 직접 말했다는 한 중국인 테슬 라 오너의 경험담도 있었다. 따라서 지금부터는 도시 도로 Autopilot 자동 보조 운전라는 공식 명칭을 편의상 "중국판 FSD"라고 하겠다.

현재 중국에서 출시되는 FSD 버전은 v13.2.6으로 보이며, 이는 북미 버전인 V13.2.7과 거의 동일한 수준이다. (물론 지금 미국에서는 V13.2.8이 순차적으로 배포되고 있는 중) 이 는 테슬라가 중국 시장에 제공하는 기술이 최신 버전에 가깝다는 것을 보여준다. 물론, 중 국 도로 데이터로 훈련된 버전으로 미국 버전과 일부 차이가 있을 수 있다.

중국 버전은 현지 도로와 교통 규칙에 맞게 최적화되었으나, 미국 FSD보다 데이터 훈련 수준이 낮을 가능성이 있다. 이러한 차이는 사용 가능한 훈련 데이터의 양과 질, 그리고 현 지 교통 법규 및 운전 문화의 차이에서 비롯된다. 이번 중국 업데이트의 주요 기능으로는 다음과 같은 요소들이 포함된다:

> 1. 기존 NOA 자동 보조 운전 기능: 도로 제한 속도에 맞춰 주행. 사용자가 진입/진출 램프(고속도로 진출입로)를 통과해야 하는 경우와 시내 주행에서 사용할 때 시스템은 도로 표지판과 신호등을 기반으로 차량을 유도. 또한 차 선을 변경하며, 좌/우회전, 정지 등을 수행. 차선 변경 시 속도와 차선을 자동으로 조절하여 최적의 경로를 선택.

> 2. 운전석 쪽 카메라: 운전석 쪽 카메라를 통해 운전자의 주의 집중 여부를 판단하고 필요시 경고를 보냄. 운전석 쪽 카메라 영상은 차량 내부에만 저장되며, 테슬라사는 해당 영상에 접근 권한이 없음. 이는 사용자 프라이버시를 보호하기 위한 조치.

> 3. 지도 버전 업데이트: CN-2025.8-15218 버전이 적용되었으며, 이는 중국 도로 환경에 맞게 최적화된 지도 데이터를 포함. 3D 차선 수준 내비게이션(Baidu Maps)과 통합.

표 7. 테슬라 소프트웨어 업데이트 주요 기능 및 중국 현지화 내용

카테고리	기능 세부 사항	비고 및 추가설명
	 도시 도로에서 차량 안내, 출구 및 교차로 인식·안내 교통 신호 인식 및 자동 속도 조절 좌회전, 우회전, U턴 자동 실행 교통 상황 기반 차선 변경과 속도 조절 운전자의 지속적인 감독 필수 완전 자율 주행 아님, 항상 운전자 주의 필요 	 중국 내 공식 명칭: 도시 도로 Autopilot 자동 보조 운전 HW4.0 차량 우선 적용 중국 현지 데이터 기반으로 기존 미국 FSD 모델 미세조정(Fine-tuning)
네비게이션 개선	 4가지 선호경로 옵션: 가장 빠른 시간, 교통 체증 최소화, 고속도로 우선, 통행료 최소화 서비스 지역(휴게소, 충전소 등) 접근 시 자동 팝업 안내 차량 내부에서 슈퍼차저 볼라드 잠금 해제 가능 	중국 버전 지도 데이터 업데이트 적용 버전: CN-2025.8-15218
음성 명령 향상	 사용자 요청을 더 정확히 이해하고 응답하는 개선된 음성 명령 시스템 음성으로 내비게이션 시작 및 음량 조절 가능 	운전자 안전성 및 편의성 향상
운전자 모니터링 강화 (중국 현지화 추가 사항)	 운전석 쪽 카메라로 운전자의 주의 집중 여부 판단 및 경고 제공 운전석 쪽 카메라 영상은 차량 내부에만 저장됨(테슬라는 접근 불가능) 	운전자 책임과 프라이버시 강화
기타 기능	 업데이트된 중국 현지 지도 데이터로 경로 정확도 향상 예상 도착 시간(ETA) 계산 정확도 개선 상세한 차선 안내 기능 개선 비상 차량 사이렌 인식 데이터 공유 옵션(선택 시 적용) 서비스 지역 접근 시 자동 팝업 안내 차량 내부에서 슈퍼차저 볼라드 잠금 해제 가능 	중국 사용자 현지화
적용 가능 차량 및 조건	HW4(하드웨어 4.0)를 탑재한 차량 우선 적용 FSD 패키지(元64,000, 약 \$8,800) 구매 차량에 적용 가능	단계적 적용 및 하드웨어 제한 존재
중요 참고 사항	 미국 버전(FSD v13.2.7)과 일부 기능 차이 존재 가능 완전 자율 주행 아니며, 운전자의 지속적인 주의 필수 중국 내 데이터 저장 및 처리 필수, 데이터 국외 유출 금지 	현지 규제 준수 및 안전성 강조

자료: Tesla, 미래에셋증권 리서치센터

(2) 가격 책정 및 시장 잠재력

당연하게도 이번 업데이트는 FSD 소프트웨어를 구입한 차량에만 적용된다. 중국판 FSD는 일회성 요금 RMB 64.000(약 \$8.800)으로 구매 가능하다. 참고로 "향상된 자율 주행(EAP, 한국에서도 EAP는 구매 가능)"과는 다르다. EAP는 FSD의 절반 가격으로 제공되지만 FSD보다 기능이 크게 제한된다.

그리고 중국판 FSD는 적용 범위가 현재로서는 많이 제한적이다. "HW 4.0" 하드웨어를 탑 재한 차량이어야 하고, "구독이 아니라"일시불로 FSD를 구매해야 사용할 수 있기 때문이 다. 테슬라 차이나는 2024년 2월 1일 공식 소셜 미디어 계정(Weibo, WeChat)을 통해 모 델 Y의 HW 4.0 업그레이드를 발표한 바 있다.

> 즉, 이날 이후의 모델 Y의 전체 라인업이 HW 4.0으로 업그레이드된 것이고 중국판 FSD 를 구동할 수 있는 차량들로 봐야 한다.

그림 22. 중국판 FSD의 가격은 64,000위안

"미래에 당신의 차량은 운전자의 최소한의 개입 하에 대부분의 운전 임무를 완수할 수 있을 것"

Autopilot 自动辅助驾驶套件

FSD 智能辅助驾驶功能

¥64,000

包括基础版辅助驾驶和增强版自动辅助驾驶的全部功能。 稍后推出受控道路和城市道路 Autopilot 自动辅助驾驶功 能。未来您的车辆将能够在驾驶员极少干预的情况下完成 绝大多数的驾驶任务。

자료: Tesla, 미래에셋증권 리서치센터

따라서 2024년 3월을 기준으로 했을 때. 2025년 1월까지의 중국 내 테슬라 차량 누적 판 매량을 계산해보면 약 62만 대다. 그 중 20%의 오너인 12만 명이 FSD를 구매한다고 가 정해 보면 중국 FSD 구매 가격인 64,000위안을 기준으로 계산 시 약 77억 위안(한화 약 1.5조원)의 시장 규모가 예상된다. 이는 테슬라에게 상당한 추가 수익원이 될 수 있다. 또 한 중국에서 새로운 테슬라 모델 Y "주니퍼"의 인도가 2월 26일부터 시작되었다는 점도 고 무적이다. 시기상 신차를 사는 고객들은 중국판 FSD도 함께 구매했을 가능성이 높기 때문 이다.

더 넓게 보면, 2020년 이후 중국에서 판매된 테슬라의 누적 차량은 약 219만 대인데, 이중 HW 3.0 탑재 차량은 약 160만 대로 추정된다. 이들은 잠재적인 고객층이 될 수 있다. 하 지만 HW 3 차량에 대한 FSD 적용은 최적화 및 배포에 시간이 필요할 것으로 보인다.

테슬라 자율주행팀이 HW 3.0에서 FSD를 단기간 내 구현하지 못할 경우, 미국 시장과 마 찬가지로 무료 하드웨어 교체(HW 3.0을 HW 4.0로 교체)가 이루어질 수 있다. 이때 보수 적으로 10~15%의 도입률로 가정해보면 약 20만 명의 FSD 사용자가 선불로 8,800달러 를 지불하는 규모가 될 것이다. 중국 시장에서의 FSD 사업이 가진 경제적 잠재력을 보여 주는 데 충분한 요소다.

(3) 중국판 FSD의 실제 성능과 사용자 반응

중국판 FSD가 풀리자 실제로 작동하는 모습이 많은 블로거들이 실시간 스트리밍 중계를 진행했다. 미국산 자율주행 기계가 얼마나 중국에서도 운전을 잘하는지 테스트하는 "라방" 을 한 것이다. 특히 주목할 점은, 북미 지역 테스트 영상은 많이 접했지만 중국 현지 테스 트 영상은 모두 라이브 방송이라 더욱 신뢰할 수 있다는 점이다.

중국의 테슬라 차주들은 주로 임의의 목적지를 설정해 FSD를 구동하는 모습을 보였는데, 그들의 전반적인 소감은 예상보다 뛰어난 성능에 놀랍다는 점이었다. 기본 주행 능력이 매 우 훌륭하며, 애초에 FSD는 운행 설계 영역(ODD) 제한도 없다. 즉, 기계가 모르는 길, 안 가본 길도 그냥 갈 수 있다는 말이다.

그림 23. 중국판 FSD 사용자들이 실제로 업로드한 주행 영상 (1) 오토바이가 쉴 새 없이 끼어들고 반대편 차가 불법 유턴하는 등 매우 복잡한 환경에서도 잘 달림





자료: 샤오홍슈(@月亮爱生活), X(@XieJackie), 미래에셋증권 리서치센터

그림 24. 중국판 FSD 사용자들이 실제로 업로드한 주행 영상 (2) 매우 좁고 혼잡한 시골길을 달림, 지도에 전혀 제대로 나와있는 않은 길도 알아서 달림





자료: X(@niccruzpatane), Douyin(@傑克ni), 미래에셋증권 리서치센터

그림 25. 중국판 FSD 사용자들이 실제로 업로드한 주행 영상 (3) 지하 주차장에서도 주행 가능하고 출퇴근 시간대의 폭우 환경에서도 주행을 처리



자료: X(@Veowow), X(@boolusilan), 미래에셋증권 리서치센터

그림 26. 중국판 FSD 사용자들이 실제로 업로드한 주행 영상 (4)

2차선 도로에서 늦게 달리는 트럭을 추월 주행하고, 갑작스럽게 콘이 있어도 부드럽게 피해서 주행



자료: Douyin(@第一智驾), X(@boolusilan), 미래에셋증권 리서치센터

테스터들은 도시, 좁은 길, 시장, 고가도로, 지하 주차장, 심지어 산악 도로 등 다양한 환경 에서 주행 테스트를 진행했다. 그리고 악천후 상황에서의 테스트를 진행한 오너들도 곧바로 나왔다. 일부 매체에서 예상했던 '주행 불가'나 잦은 사고 발생과 같은 실패 사례는 거의 보 고되지 않았다. 주행은 놀라울 정도로 인간과 유사했으며, 전반적인 주행 경로 계획 능력은 최고 수준으로 평가되었다. 이는 테슬라 FSD가 중국의 복잡한 교통 환경에서도 효과적으 로 작동할 수 있음을 보여주는 증거다.

특히 테스트를 진행한 팀 중에 전문가들로 구성된 팀도 있었다. 이들은 화웨이, 리오토, 샤 오펑, BYD 등 중국의 여러 경쟁 자율주행 시스템을 직접 경험한 전문가들로 구성된 팀이었 다. 그리고 이들의 남긴 첫 인상은, "다른 중국 자율주행 시스템과 비교했을 때 테슬라 FSD가 훨씬 사람처럼 운전하고, 부드럽고, 자신감 넘친다"는 평가였다. 특히 주목할 만한 점은 경로를 잘못 들거나 회전 구간을 놓치는 등의 실수가 발생해도, 스스로 "수정"하는 능 력을 보여주었다는 것이다. 화웨이를 포함한 중국의 다른 자율주행 시스템들은 이런 상황에 서 대부분 주저하는 경우가 많다고 덧붙였다.

사실 약 5년 전부터 end-to-end 방식을 시도해온 중국의 자율주행 업체인 Mometa는 이러한 방식을 구축하기 어려운 이유가 "의사결정 계획을 세우는 것" 때문이라고 언급한 바 있다. 테슬라 FSD는 이 부분에서 뛰어난 성능을 보여주며 기술적 우위를 입증한 셈이 다. 결국 해당 테스터들은 중국판 FSD가 지금까지 경험한 시스템 중 가장 사람과 비슷하 게 운전한다고 평가했으며, 10점 만점에 9점을 부여하며 영상을 마쳤다.

다만, 그들의 전체 주행 중에도 총 4번의 운전자 개입(takeover)이 있었다. 그 중 3번은 FSD가 중국의 버스 전용 차선이나 자전거 도로를 일반 주행 차선으로 착각한 경우였다. 나머지 한 번은 밴 차량이 갑자기 끼어들어서 측면 충돌을 피하려고 운전자가 직접 운전대 를 잡은 상황이었다.

> 또한 다른 테슬라 오너들은 테슬라의 이번 중국판 FSD 버전은 "빨간불 무시, 좌회전 차선 에서 직진, 차선 잘못 진입"과 같은 문제도 여전히 발생한다고 보고하면서 다양한 비판을 쏟아내기도 했다. 하지만, 빨간불 무시와 같은 문제는 AI 모델 개발에 있어 "검증 가능한 보상"이 있는 고치기 비교적 수월한 것이다. 따라서 중국판 FSD 또한 매번 업데이트를 거 듭하면서 점점 중국의 특정 주행 규칙을 마스터함에 따라 점차 안정적인 운전 기계로 변모 할 것이다.

3. "모든 것이 컴퓨터다": FSD의 중국 진출이 보여준 비교우위

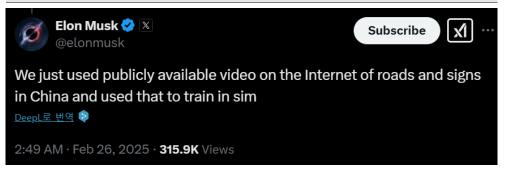
(1) 중국판 FSD의 기술적 특성과 테슬라의 방법론

FSD가 중국 현지 차량 데이터가 아니라 공개된 영상만으로 학습했다는 점을 감안하면 매 우 놀라운 성과다. 중국에 한 번도 와본 적 없는 미국인이 현지 교통 규칙이나 혼잡한 도로 상황에서 중국 운전자들의 운전 습관을 전혀 모르는 상태에서 이 정도로 운전하는 것은 상 상하기 힘든 일인데, FSD가 그것을 해낸 셈이다.

다만, 현재로서는 중국판 FSD가 방대한 양의 현지 사정의 도로 데이터로 훈련되지 않았기 때문에, 마치 '미국인 운전자가 중국에서 일주일 정도 적응한 수준'으로 작동하는 게 아닐 까 한다.

그렇다면 테슬라 차이나 팀은 FSD를 어떻게 훈련하고, 데이터는 어떻게 구했을까? 이 훈 련에는 테슬라 차량들이 중국 주요 도시들의 특정 도로를 주행하면서 수집한 데이터와, 유 튜브 등에서 찾을 수 있는 공개된 중국 도로 및 표지판 영상 데이터를 활용했다.

그림 27. 일론 머스크 "우리는 인터넷에서 공개적으로 사용 가능한 중국 도로 및 표지판 비디오를 사용하여 시뮬레이션 환경에서 훈련했습니다."



자료: Elon Musk, 미래에셋증권 리서치센터

CEO 일론 머스크는 "우리는 인터넷에서 공개적으로 사용 가능한 중국 도로 및 표지판 비 디오를 사용하여 시뮬레이션 환경에서 훈련했습니다"라고 직접 밝히기도 했다. 기본적으로 엔비디아의 "합성 데이터 생성 기반의 자율주행 구축" 전략과 대동소이하다. 머스크가 공개 한 방식에 따르면, 인터넷에서 공개된 주행 데이터를 캡처하여 시뮬레이션 환경에서 훈련한 다음, end-to-end 모델을 훈련하는 방식을 사용한 것으로 풀이할 수 있다.

^{*} 중국과 미국을 오가면서 FSD 모델을 개선할 수 있는 테슬라의 방식을 추측해봄

> 1. 초기 모델 배포 (미국 → 중국): 미국에서 개발된 FSD V13 모델(가중치 포함)을 중국 서버로 보냄. 이 과정에 서 중국의 데이터 규제(원시 데이터 국외 반출 금지)를 준수하기 위해 모델 가중치는 암호화되어 전달됐을 가능성

> 2. 중국 로컬 데이터로 미세 조정: 중국 서버에서 상하이, 베이징, 선전 등 현지 주행 데이터(카메라 영상, 센서 데이터 등)를 사용해 모델을 미세 조정. 이 과정에서 모델의 가중치가 업데이트되고, 그래디언트가 계산

> 3. 암호화 및 본사 전달 (중국 → 미국): 中에서 계산된 그래디언트를 美 본사로 보낼 때, 데이터 보안을 위해 동 형 암호화. 암호화된 상태로 전달되더라도 미국에서 바로 연산(예: 글로벌 모델에 통합)에 사용할 수 있음.

> 즉, 중국판 FSD는 미국 FSD 모델의 가중치(weights)를 가져와 중국의 제한적인 로컬 데 이터 및 시뮬레이션 데이터로 추가 조정한 모델로 사료된다. 미국의 FSD는 이미 일반적인 도로에서 인간 운전자처럼 주행할 수 있는 '주행 지능'을 갖추고 있다.

> 따라서 중국 현지 데이터로 약간의 조정(fine-tuning)만 거쳐도 중국 도로에서 운전이 가 능해졌다. 미국 FSD 모델의 가중치를 초기값으로 사용하고, 중국 관련 데이터로 중국 도로 환경(표지판, 차선, 교통 흐름 등)에 관련된 부분을 중점적으로 추가 학습을 진행하는 것이 라 할 수 있다. 쉽게 말해 처음부터 훈련하는 대신, 중국의 특정 도로 상황에 맞게 알고리 즘을 최적화하는 훈련을 한 것이다.

> 실제로 테슬라와 접촉한 여러 사람들은 2025년 2월 초. 테슬라가 미국 자율주행 엔지니어 일부를 중국으로 이동시켜 FSD를 현지화하고 알고리즘을 최적화하기로 결정했다고 언급했 었다. 한편, 현재 중국에서의 추가 훈련을 위한 연산 능력은 중국 내 GPU 클러스터(1,024 개의 NVIDIA H100)를 임대하여 사용했을 것으로 추정된다. 이는 중국 법규에서 허용되는 방식이기 때문이다.

> 물론 이런 방식의 효과는 오리지널 FSD에 비해서 제한적일 수 있다. 미국 내 사정과는 다 르게 현지의 데이터가 많이 부족하기 때문이다. 일론 머스크도 FSD 중국 배포의 가장 큰 과제가 데이터 부족임을 확인한 바 있다.

> 또한 시뮬레이션 데이터와 실제 데이터의 품질 및 양적 차이가 너무 크기 때문에, 상당 기 간 동안은 미국 FSD처럼 빠른 업데이트 주기 및 개발 진전 속도를 기대하기 어려울 것이 다. 그리고 데이터를 수집할 수는 있지만 북미 FSD 팀에 바로 전송할 수 없다는 점도 빠른 업데이트 반복 주기를 방해하는 요인으로 작용할 것이다.

> 그럼에도 이제 중국판 FSD가 실 사용자의 자동차에 풀렸기 때문에, 실제 주행 데이터를 공급받아 훈련에 사용하면서 이를 차차 극복해 나갈 것으로 예상한다.

(2) End-to-end 방식이 좋은데 왜 빨리 못 따라 하나?

자율주행 기술 개발에 있어 두 가지 주요 접근법이 존재한다. End-to-end 방식과 여러 모 듈 방식이다. End-to-end 방식은 입력(센서 데이터)에서 출력(차량 제어 명령)까지 하나의 신경망으로 처리하는 방식이다. 예를 들어, 자율주행에서 카메라 데이터를 입력받아 바로 차량의 조향, 가속, 브레이크 명령을 생성한다.

반면, 여러 모듈 방식은 자율주행 과정을 여러 단계로 나누어 각각에 대해 명시적인 규칙과 알고리즘을 적용한다. 예를 들어, 차선을 감지하는 모듈, 경로를 계획하는 모듈, 차량을 제 어하는 모듈 등이 별도로 작동한다.

여러 모듈을 함께 사용하는 방식의 장점(해석 가능성, 유연성)도 있다. 그러나 End-toend 방식의 장점이 훨씬 명확하다. 전체 과정을 하나의 신경망으로 처리하기 때문에 중간 단계에서 발생하는 오류가 누적되지 않으며, 시스템 전체를 한 번에 최적화할 수 있다. 또 한 대량의 데이터를 통해 복잡한 패턴을 학습할 수 있어, 특정 상황에서의 운전 패턴을 데 이터로 학습하여 더 자연스러운 주행이 가능하다.

하지만, 이것을 쉽게 모방하기란 쉽지 않다. End-to-end 방식은 모든 입력-출력 관계를 학습해야 하므로, 매우 많은 데이터가 필요하다. 단순히 인간의 주행 데이터를 모방하는 것 만으로는 신경망을 충분히 훈련시키기 어렵다. 이를 극복하려면 데이터가 정말로 방대해야 한다. 데이터가 부족하면 오히려 일반화 성능은 극히 떨어질 수 있다는 말이다.

예를 들어, 빨간색 신호등을 보고 차가 멈춰야 한다는 행동을 학습하려면 "신호등의 색상 (빨간색, 초록색), 신호등의 위치, 주변 환경(도로, 차량, 보행자 등)"과 같은 사항들을 종합 적으로 이해해야 한다. 인간은 이러한 지식을 운전 면허증을 따기 전부터 자연스럽게 알고 있지만, 기계는 그렇지 않다. "end-to-end로 하겠다"고 아무런 사전 지식 없이 맨바닥부 터 학습하려면 정말 많은 데이터가 필요하다는 점을 알아야 한다.

자연스럽게도, 복잡한 작업(예: 자율주행)을 하나의 신경망으로 처리하려면 학습 과정이 매 우 복잡하고 시간이 오래 걸린다. 모든 자율주행 기술 기업들이 end-to-end로 빠르게 전 환하지 못하는 이유가 바로 이런 데에 있다. 이러한 현실적인 어려움 때문에, 안드레 카파 시(전 FSD 총괄)조차 중간 표현(특징, 검출기 등)을 먼저 활용하여 사전 훈련을 수행하고, 이를 바탕으로 점진적으로 end-to-end로 나아가는 것이 효과적이라고 말한 적이 있다.

오히려 여러 신경망 모듈을 보유한 기업들은 특정 작업에 특화된 지식(신호등 인식 모듈, 객체 감지 모듈 등)을 이미 학습한 상태이기 때문에, end-to-end 시스템을 처음부터 구축 하는 것보다 유리한 출발점을 가질 수 있다.

> 즉, 갑자기 신생 기업이 나타나 마법처럼 만들 수 있는 것이 end-to-end가 아니라는 말이 다. 이와 관련해 중국 샤오펑 CEO인 허 샤오펑은 작년 WeChat 게시글에서 아래와 같이 말한 적이 있음을 유념하자. 이것이 중국의 자율주행 기술의 현황에 관한 컨센서스일 가능 성이 높기 때문이다.

> "AI/자율주행/Level 4 업계 친구들과 로보택시에 대해 이야기를 나눴습니다. 대략 2년 후(2026년 하반기)에 전환점에 들어설 것 같습니다. 오늘날 많은 L4 회사의 기술 노선은 여전히 라이다+고정밀 지도+소형 모델의 조합입니다. 이 업계의 모두가 End-to-end로 전환해야 할지 말지를 고통스럽게 고민하고 있습니다. 하지만 제 제안은 '고민하지 말고 빨리 바꾸자. 나중에 그게 대세가 될 것'이라는 점입니다"

(3) 미국 기업 테슬라의 SW 격차 우위와 함의

end-to-end를 이제 본격적으로 접목하고 있는 중국 업체들과 테슬라의 기술적 격차는 "몇 년"으로 측정할 수 없는 복잡한 문제인 것 같다. 소프트웨어 알고리즘의 진입 장벽은 매우 높다. 화웨이, 샤오펑, NIO와 같은 중국 기업들에도 매우 우수한 AI 인재들이 있지만, 테슬라의 방법론을 제대로 이해하지는 못하는 상황이다.

이와 관련해 테슬라는 더 이상 AI DAY를 개최하지 않고, 그들의 아키텍처를 공개하지 않는 전략을 펼치고 있다. 물론 모든 업체들이 테슬라가 AI DAY 2021에서 공개한 여러 개념들, VLA(Vision-Language-Action) 같은 end-to-end 방식에 OCC(Occupancy Network) 와 같은 개념을 활용해야 한다는 것은 알고 있지만, 인식부터 제어까지 전 과정을 "매끄럽 게" 실행하는 것은 여전히 큰 과제다.

일부 오픈소스식 end-to-end 아키텍처와 코드가 존재하지만, 대부분 연구용 데모일 뿐 실제 도로 환경에서 사용할 수 있는 실용성을 갖추지 못했다는 평가도 많다. 다시 말해, FSD는 쉽게 리버스 엔지니어링이나 모방하기 어려운 복잡성을 가지고 있다.

그림 28. 테슬라가 전 직원인 Guangzhi Cao를 상대로 제기한 소송

Cao는 테슬라의 자율주행 기술 관련 팀에서 일하다가 중국 전기차 기업인 샤오펑으로 이직 자율주행 기술 관련 소스 코드와 기밀 정보를 무단으로 가져가 샤오펑에 제공했다고 주장하며 소송



자료: 캘리포니아 북부 연방지방법원, 미래에셋증권 리서치센터

> 그렇기 때문에, 이러한 소프트웨어를 탈취하려는 시도가 없었던 것은 아니다. 2020년 9월, 테슬라의 전 직원인 Guangzhi Cao가 테슬라 자율주행 기술 관련 팀에서 일하다가 중국 샤 오펑으로 이직하면서 자율주행 기술 관련 소스 코드와 기밀 정보를 무단으로 가져갔다는 소송이 제기된 바 있다. 이 사건 이후로, 중국 내 어떤 회사도 테슬라의 인재를 감히 스카 우트하려는 시도조차 하지 못하게 되었으며, 실리콘 밸리에서도 비공식적인 접촉이 어려워 졌다고 한다.

> 그리고 이번 롤아웃을 통해, 결국 테슬라의 end-to-end 및 카메라 비전 접근 방식은 대규 모 현지 차량이나 전담 훈련 없이도 FSD가 중국에서 높은 수준으로 작동한다는 점에서 큰 검증을 받는 데 성공했다. 반면, 중국의 대부분 end-to-end 팀은 특정 도시에 과적합 (Overfitting)된 결과를 보이는 경향이 있으며, 전국적인 end-to-end 시스템 출시는 아직 요원한 상황이다. 거기다가 지역 간 적응은 더욱 어려운 과제로 남아 있다.

> 한편, 샤오펑의 자율주행 제품 총괄인 류이린(Liu Yilin)은 테슬라 FSD의 중국 진출에 대해 전문가로서 상세한 평가를 제공했다. "FSD가 중국 특유의 교통 환경에 완전히 적응하지 못 했지만 기본적인 기술력은 매우 뛰어나다"는 관점을 내비쳤다.

> 류이린은 테슬라 FSD가 버스 전용 차선, 대기 구역, 복잡한 적색 신호등 등 중국 특유의 교통 상황에서 불편함을 보이는 점을 지적하긴 했다. 그러나 현지 교통 인프라와 교통 규칙 에 적응하지 못하는 부분은 비교적 쉽게 풀리는 문제다. "검증 가능한 보상"의 문제이기 때 문에 추가적인 현지 데이터 기반의 강화학습을 통해 극복이 될 수 있다. (예를 들어, 중국 환경을 시뮬레이션하는 3D 환경에서 "실선 침범 시 감점, 신호 위반 시 감점 등의 규칙 적 용"등의 방식으로 학습)

> 동시에 류이린은 테슬라 FSD의 우수한 기본 능력에 평가했던 부분을 아래와 같이 정리할 수 있다.

- 정적 감지: 정지된 물체의 감지가 매우 정확하고 안정적이며, 다양한 날씨와 교통 조건에서도 안정적으로 작동
- 동적 감지: 매우 뛰어나며, 수 시간 동안 주행해도 속도 측정이나 각도 오류가 발생하지 않음
- 대응 능력: 차량 제어가 예상보다 훨씬 뛰어나며, 사람과 유사하고, 다른 차량과의 상호작용이 자연스러움

종합적으로, 류이린은 중국 시장에서의 실질적인 경험 측면에서는 테슬라 FSD가 중국 자 동차 회사들보다 뒤처지지만, 기본적인 기술력은 "무시무시하게(恐怖" 뛰어나다고 평가했 다. 현지화의 부족에도 불구하고, 기본 기술력은 중국 국내 자동차 회사들을 "멀리 앞서고 있다(远甩)"는 그의 평가는 테슬라의 기술적 우위를 분명히 보여준다. 이는 테슬라가 추가 적인 소프트웨어 개발 없이 당국의 승인만 받으면 어디에서든 FSD를 사용할 수 있다는 것 을 예고하기도 한다.

> 다시 말해, 테슬라 FSD의 중국 진출은 end-to-end 기술의 글로벌한 적용 가능성과 방법 론적 우위를 보여주는 중요한 사례다. 현지 데이터의 제한적 활용에도 불구하고 말이다.

> 또한 이러한 방식은 다양한 환경에 적응할 수 있는 범용 기술로, 로봇 산업 전반의 발전에 도 중요한 시사점을 제공한다. 로보틱스 발전에 사활을 걸고 있는 중국으로서는 절대 갈고 닦아야 하는 부분이라고 할 수 있다.

> 따라서, 중국판 FSD의 배포를 허용하게 한 것은 중국 규제 당국이 end-to-end 기술의 격차를 인정하고 빠르게 테슬라 방식을 수용하게끔 하는 조치다. 미국이 가지는 가장 소중 한 비교우위는 역시 소프트웨어다. 이 세상의 "모든 것은 컴퓨터"고 그 컴퓨터의 구동을 정 의하는 것이 소프트웨어가 될 가능성이 커졌다.

그런 면에서 볼 때 우리는 이미 "Software-Defined World"를 살고 있는 것 아닐까?

IV. Paper of the Week

Al Scribe 애플리케이션의 효용 조사

의료진이 외래 환자를 볼 때에는 문진을 통해 환자의 증상을 듣고 적합한 진단명 및 처치 의 종류를 결정해야 한다. 동시에 이러한 판단이 적절했음을 증명하기 위해 의료진은 문진 내용을 담은 진료 기록을 작성해야 한다. 여기에는 환자의 인적사항, 주된 증상, 진단 결과 또는 진단명, 진료 경과, 치료 내용과 같은 항목이 빠짐없이 들어가야 한다.

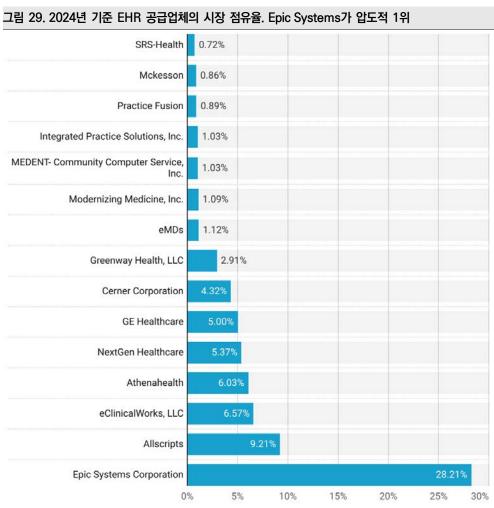
이러한 진료 기록은 환자를 보면서 동시에 작성해야 하기 때문에 의사에게는 정신적 피로 감을 줄 뿐 아니라, 신경이 분산되기 때문에 환자에게 온전히 집중하기도 어렵다. 그런데 심지어 진단학에서는 문진 과정의 첫 단계인 Chief Complaint 파악, 즉 환자의 주요 증상 을 파악하는 단계에서 가능한 환자가 이야기한 단어를 그대로 적으라고 가르치고 있다. 환 자가 "양치할 때 잇몸에서 피가 나요."라고 한다면 진료 기록에도 그대로 적어야 한다는 뜻 이다. 이는 나중에 다시 보더라도 환자가 이야기하는 뉘앙스를 온전히 파악할 수 있도록 하 기 위함이지만, 의사 입장에서는 업무 부담이 늘어나는 것도 사실이다.

한편, AI를 이용한 여러 분야에서의 업무 효율성 상승은 이미 다양한 연구에서 입증한 바 있다. 지난 2월 AI Weekly에서 소개한 앤트로픽의 Economic Index도 약 4%의 직업이 작업의 75% 이상에서, 약 36%의 직업이 작업의 25% 이상에서 AI를 사용하고 있음을 발 견했다. 물론 높은 사용률을 보이는 직종은 작업에서 코드나 텍스트가 높은 비중을 차지하 는 경우였다. 그런데 이는 의사들이 임상 문서를 작성하는 과정 역시 AI를 통해 효율화 할 여력이 있다는 의미이기도 하다. 그리고 이러한 아이디어를 기반으로 개발된 것이 AI 기반 임상 문서 작성(scribing) 기술이다.

Al Scribe 소프트웨어는 Abridge, Nuance, Augmedix 등 여러 회사에서 출시되었으나, 이를 실제로 사용했을 때 얼마만큼의 효용이 있는지에 대한 연구는 진행되지 않았다. 이는 Al Scribe 소프트웨어의 종류뿐 아니라, 각 병원마다 사용하는 EHR 소프트웨어의 종류, Al Scribe 소프트웨어별로 지원하는 EHR의 종류, EHR 시스템 내 통합 여부 등이 모두 달라 비교가 어렵기 때문이다. 그런데 지난 2월, 비록 한 메디컬 센터 내에서만 진행된 것이기는 하나, Al Scribe을 사용하지 않던 의료진이 Al Scribe을 사용한 후 발생한 업무 환경 변화 를 조사한 결과가 공개됐다.

해당 조사에서는 Nuance의 DAX Copilot을 사용했다. Nuance는 2022년에 마이크로소프 트에 인수된 진료 문서화 소프트웨어 기업으로, 대표 제품인 DAX Copilot은 지난 2024년 1월, 미국 내 EHR 중 가장 높은 점유율을 보유한 Epic Systems에 통합되어 있다. 따라서 연구에 참여한 의사들이 Epic의 EHR 모바일 애플리케이션을 통해 문진 과정을 녹음하고, 문진이 끝난 후 이를 저장하면 DAX Copilot이 요약한 진료 기록 초안이 Epic의 EHR 시스 템에 자동으로 제공된다.

> *EHR: Electronic Health Records의 약자로, 외부로 공유가 가능한 환자의 건강 기록 시 스템. 따라서 환자가 병원을 옮기더라도 지난 의료 기록을 공유해 치료 과정에서의 시행착 오를 줄일 수 있음. 환자의 건강 상태를 기록한다는 점에서는 EMR(Electronic Medical Records)과 유사하나, EMR은 기록이 작성된 망 바깥으로 벗어날 수 없음.



자료: Market.us Media, 미래에셋증권 리서치센터

연구진들은 참여자들의 다양성을 확보하는 데에도 초점을 맞췄다. 참여자들은 7개의 서로 다른 전문 분야, 22개의 임상 현장, 그리고 서로 다른 Epic EHR 시스템 사용 숙련도를 가 지고 있었다. 이렇게 선정된 의료진들은 평균 11.1년의 진료 경력을 가지고 있었으며, 주당 평균 36.5명의 환자를 진료했다. 그리고 이들이 본 환자들은 평균 10.0개의 활성 문제 목 록을 가지고 있었다. 선정된 46명의 임상의들은 2주간 DAX Copilot 사용법에 대한 교육을 받은 뒤 5주 동안 실제 외래 진료 환경에서 DAX Copilot을 사용했다.

또한 연구진은 연구 기간 및 연구 시작 전 9개월동안 참여자들의 EHR 사용 패턴 데이터를 수집했다. 여기에는 진료 당 기록 작성 시간, 당일 진료 마감률, 근무일 당 근무 시간 외 업 무 시간(일명 "파자마 시간")이 포함되어 있었는데, 이 데이터는 참여자의 EHR 효율성 지 표가 됐다. 마지막으로 5주간의 실사용 후 개방형 설문을 통해 정성적 데이터도 확보했다.

> 데이터가 두 종류였기 때문에 연구 결과 역시 정량적 지표와 정성적 분석으로 나누어 분석 했다. 정량적 지표 분석 결과, 주변 음성 인식 도구 사용은 진료 당 기록 작성 시간을 20.4%, 당일 진료 마감률을 9.3%, 주간 근무 시간 외 업무 시간을 30.0% 개선했다. 또 한, 임상의가 직접 입력한 문서 문자의 비율은 29.6% 감소했지만, 전체 기록 길이는 20.6% 증가했다. 이렇듯 정량적 지표 전반적으로 AI Scribe을 이용한 개선이 나타났으며, 이러한 결과는 임상의의 전문 분야나 환자의 복잡성, 작업량 및 임상의의 경력 등의 요인을 보정했 을 때에도 일관되게 나타났다.

표 8. Al Scribe 사용 전후의 정량적 비교

항목	Al Scribe 사용 전후 비교 전	후	전후 변화율(%)
진료 당 기록 작성 시간(분)	10.3	8.2	-15.2
당일 진료 마감률(%)	66.2	72.4	9.1
주간 근무 시간 외 업무 시간(분)	50.6	35.4	-16.9
수동으로 입력한 문자의 비율(%)	11.2	7.9	-20.7
전체 노트 길이(글자 수)	202,637.5	244,427.1	15.5

자료: JAMA Network, 미래에셋증권 리서치센터

정성적 분석은 설문 조사에 포함된 개방형 질문에 대한 답변을 바탕으로 진행했다. 연구진 이 ChatGPT 4.0를 활용해 답변을 분석한 결과. 임상의들은 AI Scribe를 사용함으로써 업 무 효율성이 향상되고 기록 작성에 소요되는 시간이 단축되었다고 보고했다. 이는 정량적 지표 결과와 일치하는 내용이다. 또한, 환자와의 대화에 더 집중할 수 있게 되었고, 문서 작성으로 인한 정신적 부담이 줄었다고 응답했다.

그러나 생성된 기록의 품질에 대해서는 엇갈린 반응을 보였다. 일부 임상의는 기록이 정확 하고 상세하다고 평가한 반면, 다른 임상의는 오류가 많고 상당한 수정 및 교정이 필요하다 고 지적했다. 특히, AI가 생성한 기록이 지나치게 장황하거나 부자연스럽고, 때로는 의학 용 어 대신 일반적인 용어를 사용하여 부정확한 내용을 담고 있다는 의견이 있었다. 이는 임상 의들이 생성된 기록을 그대로 사용하기 어렵게 만들고, 추가적인 편집 시간을 소요하게 하 는 요인이었다.

이 연구를 통해 AI Scribe가 임상의의 문서 작성 부담을 줄이고 업무 효율성을 향상시킬 수 있는 잠재력을 가지고 있음을 확인했다. 특히, 진료 중 실시간으로 기록 초안을 생성함 으로써 임상의가 환자와의 대화에 더 집중하고, 진료 후 기록 작성에 소요되는 시간을 줄일 수 있도록 돕는다는 점이 긍정적이다. 그러나 현재 기술 수준에서는 생성된 기록의 품질 문 제가 여전히 존재해. 임상의의 검토 및 수정이 필요하다는 점이 효율성을 떨어뜨리는 요인 이었다. 이는 환자의 증상은 있는 그대로 작성하면서도, 나머지 부분은 의학 용어를 사용해 야 한다는 점에서 상황에 맞는 어휘 선택이 중요하며, 이 부분에서 모델의 추가적인 성능 개선 여지가 있음을 의미한다.

> 또한 이 연구는 의사의 업무 부담 완화에 대해서만 집중했지만, 이는 병원 차원에서 AI Scribe을 도입하기 위한 충분한 인센티브가 되지 못할 수 있다. 애플리케이션별로 가격은 상이하지만, DAX Copilot의 경우 매월 약 500~800달러의 비용이 발생하는 것으로 알려져 있다. 더 저렴한 애플리케이션들도 존재하지만, 그러한 서비스 대부분은 EHR 시스템에 통 합되어 있지 않다는 점에서 어느 정도의 비용을 들여야 하는 것은 분명하다. 따라서 더 많 은 병원에서 이를 도입하려면 AI Scribe와 환자의 서비스 만족도 개선 및 진단 신뢰도 증 가 등 수익성 지표의 상관관계가 연구를 통해 드러나야 할 것이다.

표 9. 개방형 질문에 대한 답변의 주요 내용 - 정량적 분석과 달리 상반된 의견이 공존

_ UUU E	"	רוא אווו
주제	톤	답변 예시
	+	"원래도 효율적으로 일하긴 했지만, 하루에 15분은 아끼게 해준다."
DAX Copilot으로	+	"차트 작성 시간을 줄이는 데 조금 도움이 되는 것 같다…매주 서류 작성에 있어 총 2시간 정도를
인한 효율성 및		아끼게 해주는 것 같다."
시간 절약		"노트 테이킹 프로그램으로써는 굉장히 가치 있다고 생각하며, 계속해서 사용할 것. 그러나
12 2 1	,	메모를 바로 완성하고 청구할 수 있을 정도로는 생성하지 못한다."
	-	"시간 절약이 내가 기대했던 것에는 크게 못 미침."
	+	"나의 문서화 업무 부담을 드라마틱하게 줄여주었고 환자와의 대화 중간에 컴퓨터 화면을 보느라
환자와의		신경이 분산될 일이 없어졌다."
상호작용	+	"AI 문서화는 환자와의 상호작용과 효율성을 크게 향상시킨다."
0- 18	_	"진료 중 검사 결과를 설명하는 것이 어색할 때가 있어서 (DAX Copliot의) 검사 결과 보고가
		도움이 되지 않았다."
	+	"DAX를 사용할 때는 노트 작성이 힘들게 느껴지지 않아서 (전부 혼자 할 때보다) 그 자리에서
		바로 완성하고 싶은 마음이 더 커지는 것 같다."
	+	"DAX는 제 진료와 업무량을 따라잡는 능력에 긍정적인 영향을 미쳤다."
	+/-	"DAX에서 생성된 노트를 검토하고 조작해야 하긴 하지만 그래도 긴 일과 끝에 노트를
문서화 부담		처음부터 생성하는 대신 수정만 하면 된다는 점이 만족스럽다."
	+/-	"완벽한 건 아니지만 좋은 시작이다. 내가 집에 가져가야 하는 업무의 양을 크게 줄여줬다."
	+/-	"이것으로 아낀 시간이 그렇게 많지는 않지만, [DAX와 내 수정을 더한] 노트가 더 낫다… 또한
		제가 기억에 의존하지 않아도 되어서 심리적 부담감을 크게 줄여준다."
	_	"생성된 텍스트를 확인하고 수정하는 데 걸리는 시간은 DAX 없이 차트를 작성할 때와
		비슷하거나 더 길고, 차트 작성에 오류가 생길 위험이 더 크다."
	+	"환자의 히스토리 중 어느 것도 입력할 필요가 없었다. 모든 중요 요소를 포착할 수 있다는
		사실에 금방 익숙해졌다."
	+	"DAX에서 가장 많이 사용하는 기능은 주관적인 부분. 80~85% 정도 정확했다."
	-	"DAX는 대화를 의학 용어가 아닌 일반 용어로 바꿔서 표현하는 경우가 많다. 이는 진료 기록으로 사용할 수 없어 직접 편집해야 한다."
생성된 노트의		기속으로 자동될 수 없이 역합 단합에야 한다. "DAX가 HPI를 여전히 장황하고 부자연스럽게 표현해 DAX가 작성한 내용을 편집하는 데 상당한
800 포트 리 품질	_	시간을 할애해야 한다."
62	_	"언어가 어색하고 너무 격식적/전문적이다. HPI가 부정확한 경우가 많아서 수정해야 한다."
		"일관성이 없고, 긴 대화가 짧은 문장 1~2개로 요약된다. 의학적-법적 목적의 문서화에는
	_	충분하지 않다고 생각한다."
		"안타깝게도 출력물을 많이 수정해야 한다(HPI 재구성, 중복되고 불필요한 정보 삭제, 의학에서
	-	사용되지 않는 용어 변경, 결과를 올바른 섹션에 배치)."
	+	"이 기술이 최적화되면 일차 의료 서비스 제공자에게 수십 년 만에 가장 큰 발전일 것이다."
	+	"향후 업데이트가 이보다 더 도움이 될 것이라고 확신한다."
	+	"[AI 문서화] 기술 발전에 따라 결과물이 개선되고 교정 부담이 줄어들 것으로 기대한다."
DAX Copilot의		"적절한 기대치를 가지고 사용하면 오늘날에도 유용하다. 아직은 '키보드-프리'나 엄청나게
현재 및 미래	1/	시간을 아껴주는 수준은 아니다."
사용성		"AI가 이 부분에서의 목표를 달성할 수 있고, 이 목적으로 배포할 준비가 된 소프트웨어들도
	_	있다고 생각하지만, DAX는 그에 속하지 않는다. 시스템 전반에서 사용하도록 추천하지 않는다."
		"산부인과에서는 그다지 유용하지 않을 것 같다 – 템플릿화되어 있는 이러한 방문들에서는 증상
	_	리스트 작성이 가장 부담인데, 이 AI는 해결하지 못한다."
자료: 미래에셋증권 i	리서치센	

자료: 미래에셋증권 리서치센터 주: HPI = 현재 질환의 히스토리

V. Al Peer Table

표 10. Peer Table (1)

종목명	티커	(현지통화)	(조원)	1W	1M	3M	6M	1Y	YTD	의 발류체인 관련 한줄평 예정일
최종 소비자(온디바이스, 앱)										
애플	AAPL US	213.49	4,659	-6.1	-12.7	-15.7	-1.3	24.3	-14.7	05-02 엔드유저 데이터 보유 및 서비스 배포능력
테슬라	TSLA US	249.98	1,168	12.5	-29.7	-47.9	9.7	52.8	-38.1	04-23 FSD v13 공공 출시 및 1Q25 중국 출시 가능성
삼성전자	005930 KS	54,700	324	1.9	-2.3	1.6	-14.0	-22.7	2.8	04-04 온디바이스 AI, HBM, 파운드리 영역에서의 가능성
퀄컴	QCOM US	156.58	252	1.0	-8.6	-0.3	-6.2	-4.5	2.5	05-01 생성 AI 처리 강화용 온디바이스 CPU 설계(X Elite)
스냅	SNAP US	8.64	21	-4.8	-20.4	-22.7	-8.3	-22.8	-19.8	04-25 이미지 생성 모델을 탑재한 증강현실 서비스
크래프톤	259960 KS	356,000	17	1.6	10.4	4.1	5.3	54.4	13.9	05-08 PUBG, inZOI 등 자사 게임에 AI CPC 기능 탑재
비즈니스 효율화										
어도비	ADBE US	394.74	249	-9.3	-14.2	-13.3	-23.4	-19.8	-11.2	06-13 세계 최고의 미디어편집 툴. 서비스 배포능력(Firefly)
세일스포스	CRM US	279.40	390	2.4	-14.4	-20.3	9.8	-4.7	-16.4	05-29 세계 최고의 CRM 업체. AI로 사용성 강화(Einstein)
서비스나우	NOW US	844.33	253	7.7	-14.4	-24.3	-4.7	13.5	-20.4	04-24 워크플로우 자동화. AI로 사용성 강화(Now Assist)
크라우드스트 라이크	CRWD US	353.74	127	14.5	-21.7	-6.0	31.6	12.1	3.4	06-04 기업 고객 대상 엔드포인트 보안(Charlotte AI)
IBM	IBM US	248.35	335	-3.3	-4.9	9.2	17.7	34.4	13.7	04-23 기업 고객 대상 AI 모델 개발, 배포 플랫폼(watsonx)
액센츄어	ACN US	318.82	290	-4.9	-17.8	-10.4	-4.6	-13.5	-9.0	03-20 각 산업별 맞춤형 AI 에이전트를 만들기 위한 첨병
SAP	SAP GY	242.60	471	0.2	-13.4	1.5	20.9	42.9	2.7	04-22 기업용 SW 솔루션에 AI를 내장한 에이전트(Joule)
인포시스	INFO IN	1,579.85	110	-7.1	-14.2	-20.1	-18.2	-0.3	-16.0	04-17 각 산업별 맞춤형 AI 에이전트를 만들기 위한 첨병
타타 컨설턴시	TCS IN	3,511.05	212	-2.2	-10.1	-17.4	-20.4	-14.2	-12.7	04-11 각 산업별 맞춤형 AI 에이전트를 만들기 위한 첨병
일루미나	ILMN US	85.60	20	-0.4	-14.6	-40.2	-34.3	-31.0	-35.9	05-02 AI 기반 첨단 유전체 분석
템퍼스 Al	TEM US	47.17	12	4.4	-47.3	21.1	-4.2	-	39.7	05-15 AI 기반 정밀 진단 업체
루닛	328130 KS	55,000	2	3.8	-23.8	-34.4	33.3	-14.7	-13.8	03-21 AI 기반 암 검출/진단 보조 솔루션(인사이트)
Al Ops										
몽고DB	MDB US	185.37	22	2.8	-36.0	-30.1	-35.4	-47.8	-20.4	05-30 AI 모델 구축시 핵심 인프라인 '벡터 DB' 기능(Atlas)
클라우드플레어	NET US	116.15	58	-0.4	-32.1	-1.2	46.8	26.8	7.9	05-02 실시간 AI 추론에 CDN 수요증가 가능성(Workers AI)
팔란티어	PLTR US	86.24	294	12.9	-27.6	15.9	136.6	267.1	14.0	05-06 비즈니스 현장의 의사결정을 돕는 AI 플랫폼(AIP)
포스 패러다임	6682 HK	47.75	4	-6.3	-17.6	-2.8	43.0	-34.4	-6.3	03-20 중국의 팔란티어(Sage)

자료: Bloomberg, 미래에셋증권 리서치센터 / 주: 실적발표 일정은 변경될 수 있음.

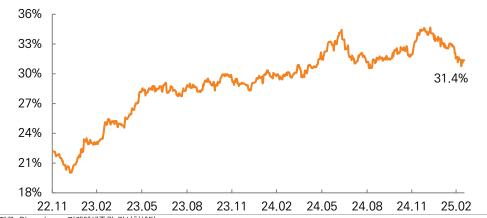
丑 11. Peer Table (2)

종목명	티커	주가	시가총액	4147		기간별 수	,	41/	\(TD	실적발표 AI 밸류체인 관련 한줄평
		(현지통화)	(조원)	1W	1M	ЗМ	6M	1Y	YTD	예정일
파운데이션 모델										
	GOOGL US	165.49	2,949	-0.2	-10.5	-15.2	4.1	17.8	-12.5	04-25 OpenAl와의 AGI 경쟁이 가능한 업체(Gemini)
메타 플랫폼스	META US	607.60	2,237	1.7	-17.4	-1.8	13.5	26.0	3.9	04-24 오픈소스 AI 개발의 선구자(Llama 등)
알리바바	9988 HK	135.80	482	1.0	11.1	62.8	63.9	93.0	64.8	05-14 중국의 CSP이자 중국의 대표적인 언어모델(Qwen)
리에브	9888 HK	91.65	48	-0.4	2.0	6.3	10.4	-9.5	10.8	05-16 NLP 및 자율주행 모델의 전통강호(Ernie Bot, Apollo)
센스타임	20 HK	1.71	12	-1.2	-6.6	15.5	55.5	103.6	14.8	03-26 중국의 멀티모달 AI의 다크호스(SenseNova)
네이버	035420 KS	208,000	33	-3.3	-5.4	-0.2	30.6	12.3	5.1	05-07 한국형 LLM 개발 및 AI 서비스(HyperCLOVA)
클라우드										
마이크로소프트	MSFT US	388.56	4,197	2.2	-4.7	-14.3	-10.3	-6.0	-7.6	04-25 OpenAl 모델 라이선스 독점권 보유한 세계 2위 CSP
아마존 닷컴	AMZN US	197.95	3,048	1.8	-13.4	-14.4	5.9	13.5	-9.8	04-30 최고의 CSP로서 Anthropic에 수 조원 투자
오라클	ORCL US	149.27	608	0.3	-14.3	-11.8	-10.4	20.2	-10.2	06-11 AI 데이터센터 capex 경쟁에 진입한 전통적 강자
소프트뱅크	9984 JP	7,836.00	113	-2.9	-20.1	-20.3	-3.8	-8.8	-14.7	05-13 '비전펀드'는 AI에 집중. 일본 최고의 AI 슈퍼컴퓨터.
하드웨어	인프라									
엔비디아	NVDA US	121.67	4,313	13.7	-12.4	-6.7	5.3	38.6	-9.4	05-28 AI 모델 훈련 및 추론에 필수인 GPU계의 현존 최강자
브로드컴	AVGO US	195.54	1,336	6.0	-16.1	-18.4	21.0	60.4	-15.7	06-12 이더넷 기반 네트워킹 반도체의 최강자
AMD	AMD US	100.97	238	4.5	-10.7	-19.2	-33.1	-47.2	-16.4	04-30 서버용 CPU의 최강자. AI 가속기 분야 패스트팔로어
인텔	INTC US	24.05	151	20.7	1.9	17.7	12.0	-43.0	20.0	04-25 파운드리로서 소버린 AI 미 지정학적 가치 부상
마이크론	MU US	100.79	163	15.7	1.3	-7.1	13.9	8.6	19.8	03-20 SK하이닉스 추격 중인 "미국"의 메모리 반도체
SK하이닉스	000660 KS	204,500	149	8.8	-2.9	11.9	26.6	28.3	18.4	04-25 HBM 부문 전세계 최강자
시놉시스	SNPS US	446.23	100	3.8	-14.6	-12.8	-11.2	-18.9	-8.1	05-22 반도체 EDA 부문 리딩 업체, AI 에이전트 적극 활용
케이던스 디자인 시스템	CDNS US	247.43	99	5.8	-16.2	-20.5	-10.0	-17.1	-17.6	04-22 반도체 EDA 부문 리딩 업체
암페놀	APH US	63.00	111	2.0	-8.5	-15.2	-1.3	16.2	-9.3	04-23 엔비디아향 AI용 서버 내 고속 I/O 및 커넥터 제조
버티브	VRT US	87.45	48	12.1	-19.1	-27.5	0.2	18.1	-23.0	04-24 액체 냉각 방식에 있어 글로벌 선두업체
Arm	ARM US	117.94	180	1.3	-26.1	-17.4	-13.8	-7.1	-4.4	05-08 AI 가속기 설계 위한 다수의 IP를 소유한 팹리스
마벨 테크놀로지스	MRVL US	68.74	86	4.7	-35.5	-38.7	-5.4	3.1	-37.7	05-30 데이터 인프라용 네트워킹 및 스토리지 ASIC 설계
중과서광	603019 CH	71.42	21	0.5	-6.2	-0.0	101.4	36.7	-1.2	04-25 중국과학원과 긴밀한 관계를 유지 중인 HPC 제조업체
SMIC	981 HK	50.50	94	-3.0	6.3	100.0	219.2	199.5	58.8	05-09 7나노 공정을 달성한 중국 유일의 미세공정 파운드리
TSMC	2330 TT	959.00	1,095	-3.9	-11.6	-10.8	1.6	29.5	-10.8	04-18 명실상부 반도체 파운드리 최강자
콴타	2382 TT	256.50	44	-1.7	-1.5	-6.7	1.0	2.9	-10.6	05-15 GPU 기반 보드 및 서버 시스템 조립 담당, 대만 업체
위스트론	3231 TT	104.50	13	-1.4	-7.5	-3.7	1.5	-10.4	0.5	05-12 GPU 기반 보드 및 서버 시스템 조립 담당, 대만 업체
Alchip	3661 TT	2,795.00	10	-2.1	-22.1	-15.8	16.5	-21.8	-14.8	05-05 빅테크들의 AI 가속기 설계를 돕는 대만의 팹리스

자료: Bloomberg, 미래에셋증권 리서치센터 주: 실적발표 일정은 변경될 수 있음.

VI. Appendix: Charts

그림 30. ChatGPT 출시일 이후, S&P 500에서 Magnificent Seven이 차지하는 비중(시가총액 기준)



자료: Bloomberg, 미래에셋증권 리서치센터

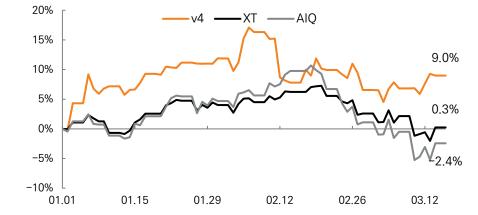
그림 31. ChatGPT 출시일 이후, AI 밸류체인에 속한 유니버스 종목들의 평균 수익률(동일가중 방식)



자료: Bloomberg, 미래에셋증권 리서치센터

그림 32. 대표 AI ETF vs AI 유니버스(v.4.1) 수익률 추이(24.12.31 기준)

- Global X Artificial Intelligence & Technology ETF(AIQ), iShares Exponential Technologies ETF(XT)



자료: Bloomberg, 미래에셋증권 리서치센터

표 12. LLM 벤치마크 챗봇아레나의 "Hard Prompts" 기준 성능 상위 모델 10개(동일 모델 생략)

순위 (전 주 변동)	모델명	아레나 점수	기관명	라이선스	지식 컷오프
1(▲1)	Grok-3-Preview-02-24	1409	xAl	독점 폐쇄소스	2025년 2월
2(V 1)	GPT-4.5-Preview	1401	OpenAl	독점 폐쇄소스	2023년 10월
3(-)	Gemini-2.0-pro-exp-02-05	1385	구글	독점 폐쇄소스	2024년 6월
4(-)	Gemini-2.0-flash-thinking-exp-0121	1381	구글	독점 폐쇄소스	2024년 6월
5(-)	o1-2024-12-17	1370	OpenAl	독점 폐쇄소스	2023년 10월
6(-)	o3-mini-high	1362	OpenAl	독점 폐쇄소스	2023년 10월
6(▲1)	DeepSeek-R1	1362	DeepSeek	상업적 사용허가	2023년 12월
8(-)	Gemini-2.0-flash-001	1358	구글	독점 폐쇄소스	2024년 8월
9(-)	o3-mini	1355	OpenAl	독점 폐쇄소스	2023년 10월
10(New)	ChatGPT-4o-latest(2025-01-29)	1353	OpenAl	독점 폐쇄소스	2023년 10월

자료: Imarena.ai, 각 사, 미래에셋증권 리서치센터

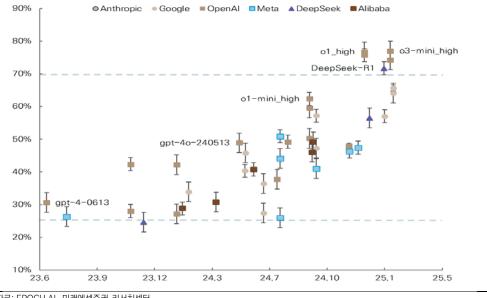
표 13. LLM 벤치마크 LiveBench의 성능 상위 모델 10개(동일 모델 생략)

순위 (전 주 변동)	모델명	전체 평균 점수	기관명	라이선스	지식 컷오프
1(-)	claude-3-7-sonnet-thinking	76.10	Anthropic	독점 폐쇄소스	2024년 10월
2(-)	o3-mini-2025-01-31-high	75.88	OpenAl	독점 폐쇄소스	2023년 10월
3(-)	o1-2024-12-17-high	75.67	OpenAl	독점 폐쇄소스	2023년 10월
4(-)	grok-3-thinking	-	xAl	독점 폐쇄소스	_
5(-)	deepseek-r1	71.57	Deepseek	상업적 사용허가	2023년 12월
6(-)	o3-mini-2025-01-31-medium	70.01	OpenAl	독점 폐쇄소스	2023년 10월
7(New)	gpt-4.5-preview	68.95	OpenAl	독점 폐쇄소스	2023년 10월
8(▼ 1)	Gemini-2.0-flash-thinking-exp-0121	66.92	구글	독점 폐쇄소스	2024년 6월
9(▼1)	claude-3-7-sonnet	65.56	Anthropic	독점 폐쇄소스	2024년 10월
10(▼1)	Gemini-2.0-pro-exp-02-05	65.13	구글	독점 폐쇄소스	2024년 6월

자료: LiveBench, Huggingface, 미래에셋증권 리서치센터

표 14. 각 AI 모델들의 GPQA Diamond 기준 성능 비교

GPQA Diamond는 박사 수준의 과학 분야 질문으로, 인간 전문가들도 약 65%의 정확도를 보임



자료: EPOCH AI, 미래에셋증권 리서치센터

주: Hard Prompts는 기존의 일반적인 요청보다 훨씬 더 까다롭고 복잡한 문제를 제시하여 LLM의 한계를 테스트하는 벤치마크

Compliance Notice

- 당사는 자료 작성일 현재 조사분석 대상법인과 관련하여 특별한 이해관계가 없음을 확인합니다.
- 당사는 본 자료를 제3자에게 사전 제공한 사실이 없습니다.
- 본 자료를 작성한 애널리스트 김은지은(는) 자료작성일 현재 테슬라 17주 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.