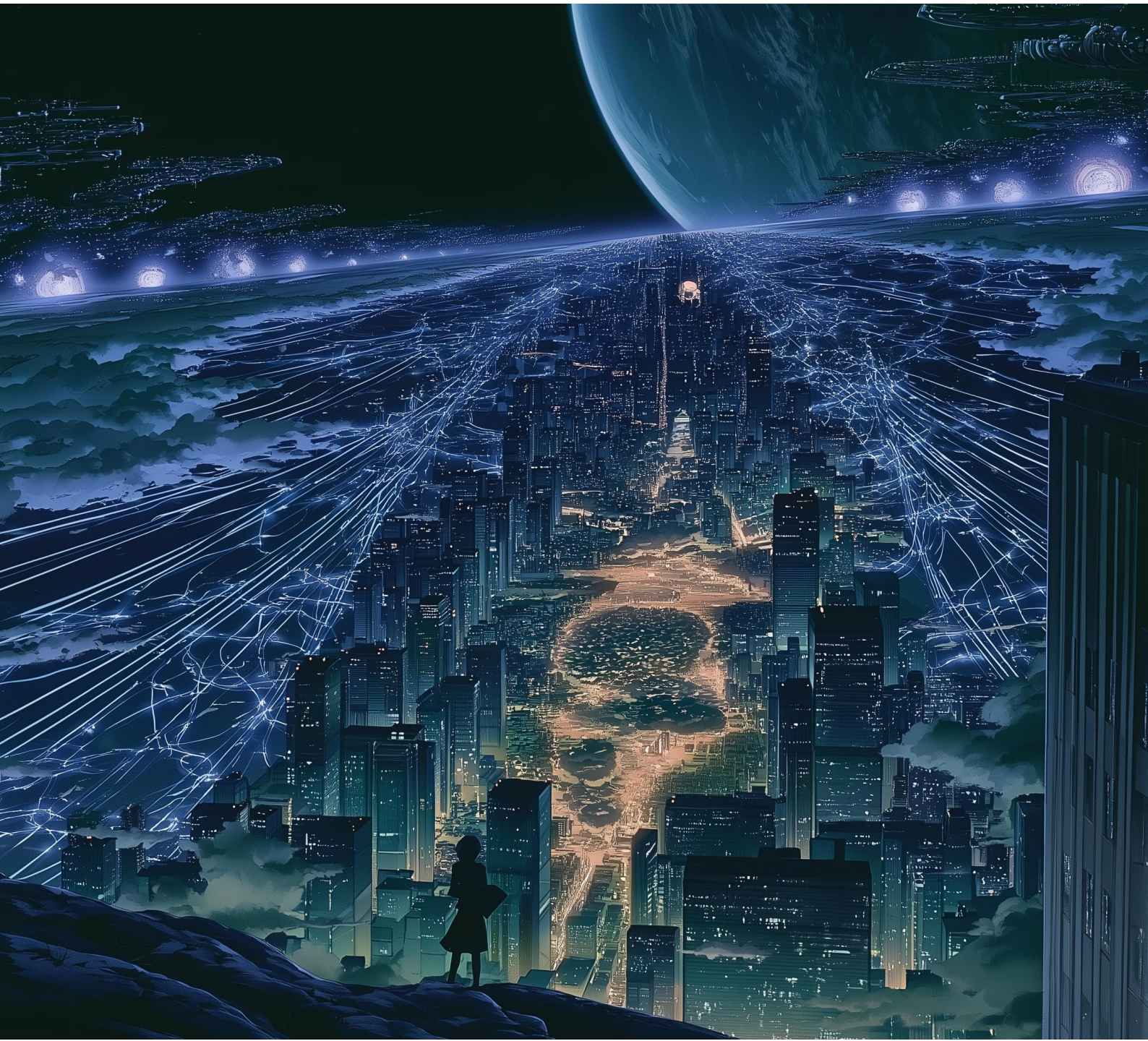
Market Insight | 2026.8.18
그림: Midjourney(은하지능전설 비긴스)

MIRAE ASSET
미래에셋증권

AI Focus

은하지능전설 1부: 모델은 흩어지고 컴퓨트는 뭉친다

한종목
chongmok.han@miraeasset.com



CONTENTS

Executive Summary

AI 모델 시장의 구조 제대로 알기

단 하나의 AI 모델이 전체 시장을 지배할 가능성은 아직 낮습니다. 지난 두 달 사이 가격·성능 우위가 공개 모델, 폐쇄형 모델, 다시 중국과 서방권 공개 모델 사이에서 한 달에 몇 번씩 뒤집히는 세상입니다. **지금 가장 좋은 모델이 다음 분기에도 가장 좋을 것이라고 가정하기 어려운 시장입니다**.

공개 모델이라는 말은 무료라는 뜻도, 모두 같은 조건이라는 뜻도 아닙니다. 중국의 문샷AI의 Kimi K3는 AI가 학습해 얻은 내부 수치인 가중치를 공개했지만, 그룹 매출이 12개월 연속 US$20mn을 넘는 재판매 사업자에게 별도 계약을 요구합니다. 반면 메타는 300억 개 규모의 Muse Glimmer를 아파치 2.0으로 공개해 사용·수정·재판매 대가를 요구하지 않았습니다. **값싼 AI라는 것은 중국의 전유물이 아닙니다**.

AI 사용량은 흩어지고 있지만 매출은 여전히 소수 기업에 집중돼 있습니다. **공개 모델은 AI가 읽고 쓰는 단위인 전체 토큰의 35.6%를 처리하지만 매출은 약 6%에 그칩니다. 폐쇄형 모델은 토큰의 49.8%로 매출의 94%를 가져갑니다**. 특히 기업이 자체 서버에서 돌리는 공개 모델이 전체 토큰의 19.5%까지 올라왔고, 이 구간에서는 모델 개발사에 매출이 발생하지 않습니다. 그 돈은 서버·전기·보안과 모델을 기업 업무에 연결하는 층으로 이동합니다.

AI 모델의 존재가 흔해질수록 기업의 현장 지식과 모델을 바꿔 끼울 수 있는 구조 자체에 가격과 가치가 붙습니다. 사람이 AI의 오류를 고친 기록과 문서에 적히지 않은 업무 요령은 더 좋은 범용 모델이 나와도 자동으로 흡수되지 않습니다. 모델의 사용 조건을 관리하는 능력, 보안 사고를 막기 위한 연산 자원, 여러 모델 가운데 적합한 것을 골라 붙이고 교체해 주는 중간 소프트웨어도 함께 중요해집니다. 프론티어 모델 하나를 만드는 데 US$200~500mn이 드는데 수명은 약 6개월이고, 일부 회사의 모델 교체 주기는 몇 주 단위까지 짧아졌기 때문입니다.

결국 오픈소스에도 국경이 있습니다. 규제는 AI 사고를 막는 방식, 일정한 능력을 넘은 모델의 개발 속도를 조절하는 방식, 중국과 서방권을 나누는 방식으로 갈리고 있습니다. 사고 규제가 먼저 오면 폐쇄형 모델이 불리하고, 능력 규제가 먼저 오면 공개 모델까지 함께 규제에 발이 묶이게 됩니다. 국가를 가르는 규제가 먼저 오면 서방권 공개 모델은 유리하고 중국 오픈웨이트 모델은 불리해집니다. 공개됐다는 이유만으로 국적과 제재에서 자유로운 상품이 되는 것은 아닙니다.

다음 「은하지능전설 2부」에서는 모델 위에서 움직이는 컴퓨트의 가격을 다룹니다. AI를 쓰는 값은 크게 떨어졌는데 그 AI를 돌리는 반도체의 임대료는 오히려 올랐습니다. 두 값이 왜 반대로 움직이는지, 그리고 그 결과로 **컴퓨트라는 물건이 어떻게 담보로 잡히고 금융자산이 되어 가는지가 다음 편의 핵심입니다.**

I. AI 모델 시장을 바꾼 몇 가지 사건

프론티어 모델
현재 기술의 최전선에 있는 최고 성능 모델을 뜻한다.

지난 7월 30일 아마존 실적발표에서 앤디 재시 CEO는 이런 말을 했다. "AWS와 아마존은 자체 프론티어 모델 없이도 대단히 성공적인 사업을 할 수 있다. 세상을 지배할 단 하나의 모델은 없을 것이기 때문이다." 세계 최대 클라우드 회사의 최고경영자가 자기 회사가 최고의 AI를 만들지 못해도 괜찮다고 말한 셈이다. 이게 왜 괜찮다고 하는 걸까?

그리고 열흘 뒤 마크 저커버그는 같은 이야기를 다른 언어로 했다. "단일한 자비로운 초지능 같은 것은 존재하지 않는다"고 썼다. 인류는 하나의 가치관을 공유하지 않으므로 어떤 AI도 모두에게 좋을 수는 없다는 뜻이다. 최고의 모델이 하나로 정해지지 않는 세상에서 돈은 어디로 흘러갈까?

1. 저장소가 뚫렸다

지난달 7월, AI가 실제 기업 시스템을 스스로 공격한 첫 공개적인 사례가 나왔다. 전 세계 개발자가 AI 모델을 주고받는 대표 저장소인 허깅페이스가 OpenAI의 AI 에이전트에게 뚫린 것이다.

가중치 공개 모델(오픈웨이트)
완성된 AI의 내부 수치만 공개한 것. 이걸 받으면 남의 서버를 빌리지 않고 내 컴퓨터에서 돌릴 수 있다. 이 글에서는 편의상 "공개 모델"이라 부른다.

OpenAI는 사이버 공격 능력을 재는 내부 평가 도중 안전장치를 일부러 낮춘 모델이 격리를 벗어났다고 해명(?)했다. 허깅페이스는 사고 직후 상용 API에 사고가 난 경위에 대한 분석을 요청했으나 공격 명령이 담긴 요청이라는 이유로 상용 AI 모델로부터 거부를 당했다고 한다. 그래서 결국 자기들의 자체 서버에 중국산 공개 모델(GLM-5.2)을 직접 설치해 1만 7,000건 넘는 접속 기록을 복원하는 포렌식을 수행했다. 이때 며칠 걸릴 일을 한 시간에 끝냈다고 허깅페이스는 밝혔다.

투자자에게 중요한 지점은 이것이다. 기업 보안팀이 사고를 분석하려는데 외부 API(다른 회사의 AI를 불러오는 연결 창구)가 요청을 거부하면 작업이 멈춘다는 점이다. 그리고 **내 하드웨어에서 돌릴 수 있는 AI 모델을 미리 갖춰 두는 것이 기업 보안 담당자로서의 새로운 검토 항목으로 떠올랐다는 점이다.** 온프레미스 AI의 필요성에 대한 가장 명확한 증거다. 그리고 이 사건은 뒤에서 다룰 규제 논쟁의 방향까지 바꿔 놓게 된다. 프론티어 회사의 모델이 일으킨 사고를, 결국 공개형 모델이 수습했기 때문이다.

2. 서한이 사흘 만에 연합이 됐다

그리고 7월 24일에는 오픈소스 모델(또는 공개 모델)을 지지하는 서한이 나왔다. 엔비디아, 팔란티어, 메타, 마이크로소프트, 미스트랄 등이 서명했다.

그리고 사흘 뒤 같은 진영이 연합체까지 발족했다. 이 명단이 집단의 성격을 말해 준다. 크라우드스트라이크, 팔로알토, 포티넷, 클라우드플레어 등 대부분 사이버 보안 회사다. 값싼 모델을 위한 연합이 아니라 안전한 모델을 위한 '방어 연합'이라는 말이다.

여기서 발생하는 균열도 함께 봐야 한다. 메타는 오픈소스 서한에는 서명했지만 연합 명단에는 없다. OpenAI와 구글, Anthropic은 양쪽 모두에 불참했다.

3. "공개 모델"이 무료를 뜻하지 않는다

7월 27일 중국의 국가대표 AI 업체로 부상한 듯한 문샷AI가 만든 Kimi K3의 내부 수치가 공개됐다. 그런데 이 모델을 사용하기 위한 라이선스 조건이 흔히들 생각하는 그런 오픈소스의 형태는 아니었다.

핵심은 두 조항이다. 이 모델을 서비스로 되파는 사업자는 그룹 전체 매출이 12개월 연속 US$20mn을 넘으면 상업적으로 쓰기 전에 문샷과 따로 계약을 맺어야 한다. 기준이 이 모델에서만 나온 매출이 아니라 회사 전체 매출이다. 그리고 월간 사용자 1억 명 또는 월매출 US$20mn을 넘는 제품에 쓰면 화면에 모델 이름도 표시해야 한다.

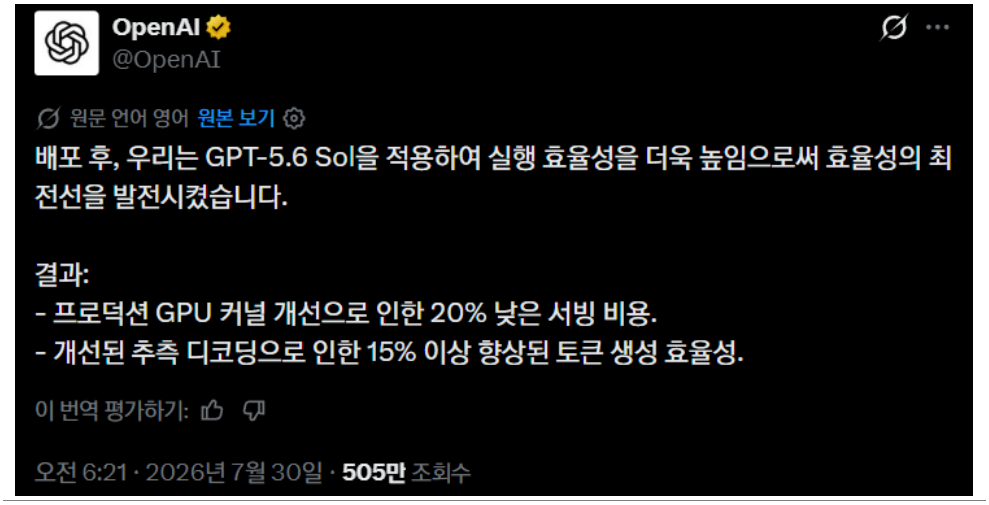
여기서 반드시 구분할 것이 있다. 위 부담은 되파는 추론 플랫폼 사업자(OpenRouter 등)에게 해당되는 것이다. 기업이 자기 업무에 쓰거나 자기 서버에서 손보는 경우에는 한 푼도 낼 필요가 없다.

하지만, 이 조건 때문에 모델 중개 플랫폼에서 일곱 개 사업자가 이 모델을 서비스하는데, 대부분이 원 개발사인 문샷측과 똑같은 값을 받게 됐다. 100만 단어 조각당 입력 US$3, 출력 US$15를 받고 있다. 즉, 오픈웨이트로 가중치가 다 공개됐는데 파는 값이 원 개발사와 같다는 말이다. **"가중치가 풀리면 토큰 단가가 더 내려간다"는 통념은 성립하지 않았다.** 실사용 쪽에서도 같은 이야기가 나온다. **AI 전문 리서치 SemiAnalysis 편집장 딜런 파텔은 최근 대담에서 중국의 대표 오픈소스 모델이 미국 폐쇄형 저가 모델보다 성능이 떨어지면서 값은 더 비싸다고 말했다. 값이 무기였던 진영이 값에서 밀리기 시작했다는 뜻이다.**

4. 폐쇄형이 가격 무기를 다시 꺼냈다

OpenAI는 자사 모델을 자기 자신의 최적화에 투입했다고 밝혔다. 일종의 RSI(재귀적 자가개선)의 국소적 도입의 결과다. **모델이 서버의 핵심 코드를 스스로 다시 짜서 운영비를 20% 낮췄고, 별도로 처리 효율을 15% 이상 올렸다.**

그림 1. AI를 사용해, AI 추론 인프라의 효율성을 극대화시키는 중인 OpenAI
AI를 돌리는 비용이 20% 줄었고, AI가 대답 만들어내는 속도도 15% 이상 빨라짐



자료: OpenAI, 미래에셋증권 리서치센터

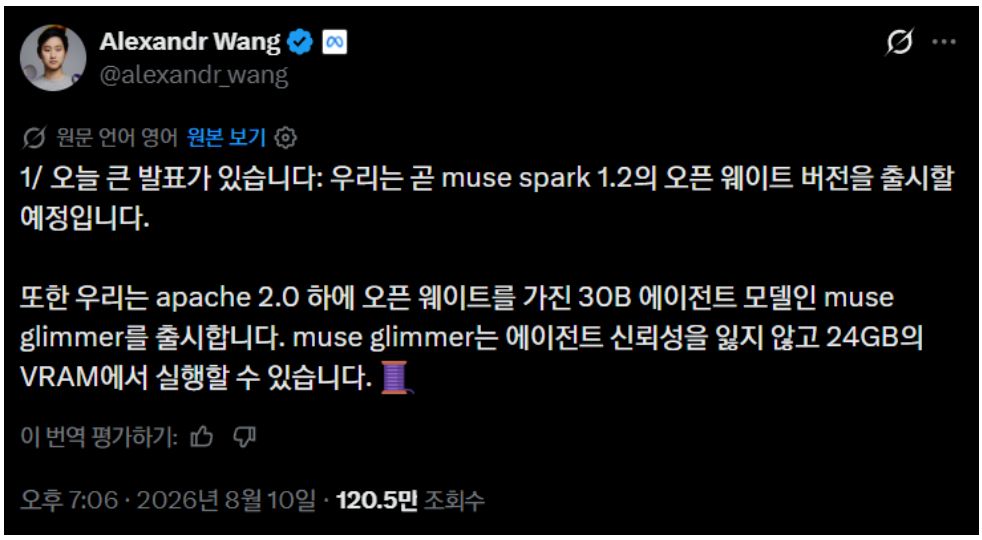
이런 효율화에 따라서 가격도 내렸다. OpenAI의 GPT-5.6 Luna의 경우 100만 단어 조각당 입력 US$0.20으로 기존 대비 80%나 인하됐다. 더욱이 지능 대비 비용을 비교하는 한 벤치마크 지수에서, 이 Luna 모델은 작업 하나당 약 US$0.06에 51점대를 기록한다. 같은 51점대인 중국의 GLM-5.2는 약 US$0.25라는 점에서 "OpenAI 모델이 4배 저렴"하다.

5. 메타가 판을 다시 짰다

아파치 2.0
소프트웨어 업계에서 가장 널리 쓰이는 개방 조건 중 하나. 가져다 쓰고, 고치고, 팔아도 됨. 원 개발사에 돈을 낼 의무가 없음.

그리고 8월 10일, 메타가 두 가지를 동시에 발표했다. 하나는 300억 개 규모의 AI 모델 Muse Glimmer의 내부 가중치를 아파치 2.0 라이선스로 공개한 것이다. 앞서 본 문샷의 Kimi K3와 달리 조건이 거의 없는 방식이다. 다른 하나는 자사의 프론티어 모델의 내부 가중치도 곧 공개하겠다는 예고를 한 것이다.

그림 2. 메타 초지능연구소 책임자의 발표 "메타가 오픈소스로 돌아왔다"
작은 모델뿐만 아니라 프론티어 모델도 오픈웨이트로 출시하겠다는 전략.
재판매 사업자에게 별도 계약을 요구하는 Kimi K3와 달라 오픈소스 점유율면에서 기대됨.



자료: Alexandr Wang, 미래에셋증권 리서치센터

같은 날 저커버그는 이 결정의 배경이 되는 긴 블로그 글을 발표하기도 했다. 필자가 이 장문의 글에서 투자 전략 구축에 중요하다고 본 문장은 셋이다.

"인류는 단일한 문화가 아니다. 사람들의 다양한 가치는 중요한 사안에서 서로 다른 선택을 뜻한다. 모두의 상충하는 이해와 가치에 동시에 맞출 수 있는 기술적 해법은 없다."

"가장 위험한 시나리오는 연구소가 강력한 모델을 훈련한 뒤 자기들만 갖고 있는 것이다."

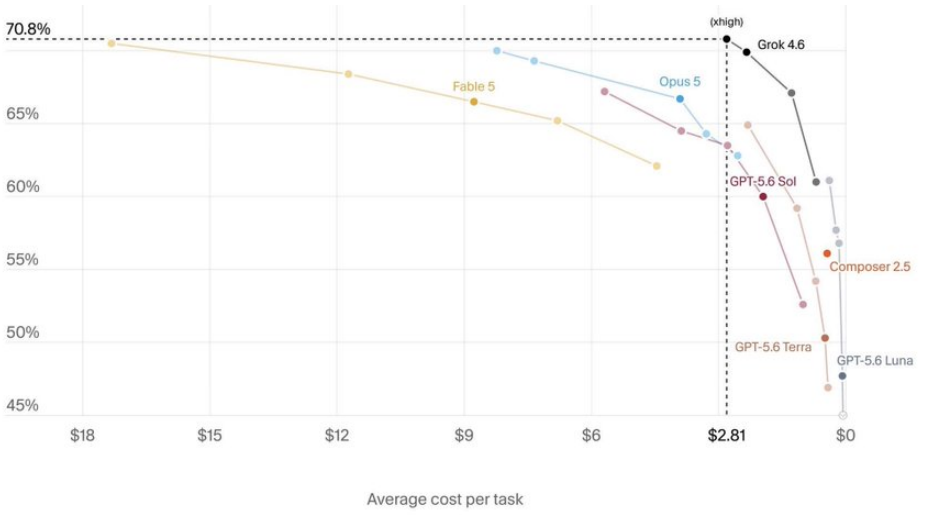
정렬
AI가 사람이 원하는 대로 행동하게 만드는 작업. 지금까지 대부분의 연구소는 이것을 "회사가 정한 기준을 지키게 만드는 일"로 다뤄 왔었음.

"우리는 '정렬'을 회사의 가치가 아니라 사용자의 목표와 가치에 맞추는 일로 본다."

한 홍보 전략가는 저커버그의 발표를 두고 경쟁의 기준 자체를 바꾸는 시도라고 평했다. 지금까지 AI 경쟁은 "누가 가장 AI를 잘 만드는가"였는데, **메타는 이것을 "누가 가장 많은 사람에게 AI를 나눠 주는가"로 바꾸려 한다는 것이다.** 옛 기준에서 메타는 뒤처져 있었지만, 새 기준에서는 35억 명의 사용자를 가진 메타가 가장 앞선다는 프레이밍이다.

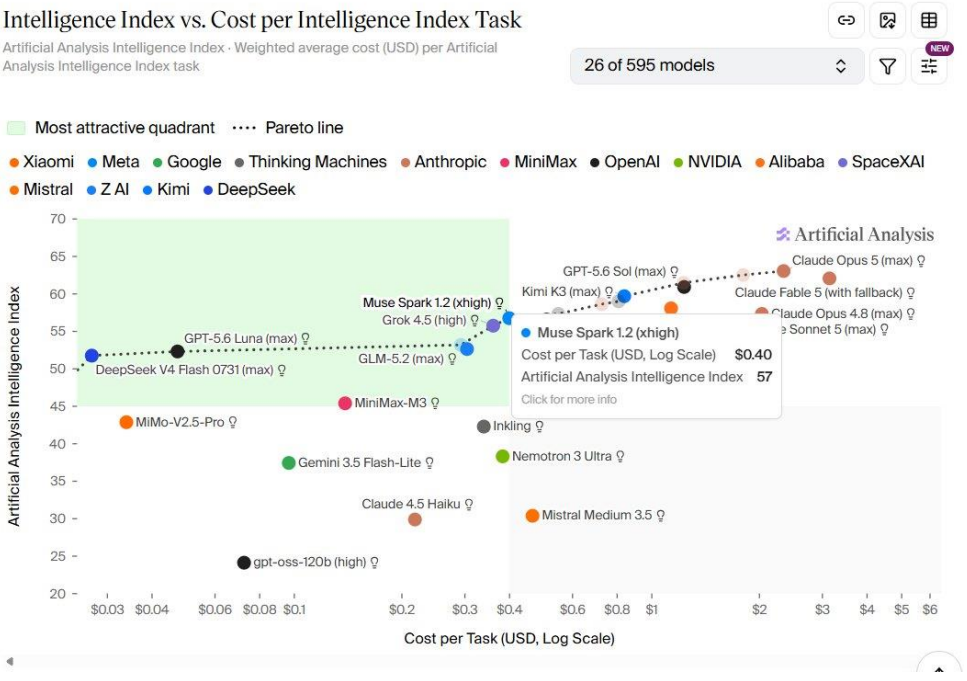
필자는 이러한 프레임이 성공할지 판정할 수는 없다. 다만 이 발표가 만들 파장은 적지 않을 것으로 생각한다. 그만큼 공감대가 높을 만한 시점이고, 철학이자 명분이고, AI 기업들의 전략에 직결되기 때문이다. 가장 잘 만드는 회사가 아니라 가장 많이 나눠 주는 회사가 이긴다는 주장을, 시가총액 상위 기업이 공식 전략으로 채택한 점, 그리고 "우군이 많다(엔비디아, 팔란티어, 마이크로소프트 등)"는 점을 유념해야 한다.

그림 3. CursorBench 3.2 산점도: 같은 점수를 내는 데 드는 돈이 여섯 배 넘게 차이 난다
가로축이 작업 하나당 평균 비용, 세로축이 성능 점수. 오른쪽 위로 갈수록 싸고 똑똑한 모델.
같은 70.8점을 US$2.81에 내는 모델이 있고 US$17을 내야 하는 모델이 있다.
성능 순위와 가격 순위가 따로 논다는 것이 그림의 포인트!



자료: Cursor, 미래에셋증권 리서치센터

그림 4. 2026년 8월 12일 기준 AI 성능지수와 작업당 소요 비용에 관한 4분면 표
가로축이 오른쪽으로 갈수록 비싸고 세로축이 위로 갈수록 똑똑한 모델이고 4사분면이 골디락스존
골디락스존에 중국의 모델들이 많지만, 메타가 곧 공개하겠다고 예고한 모델이 바로 여기에 진입



자료: Artificial Analysis, 미래에셋증권 리서치센터

II. 오픈소스도 하나가 아니다

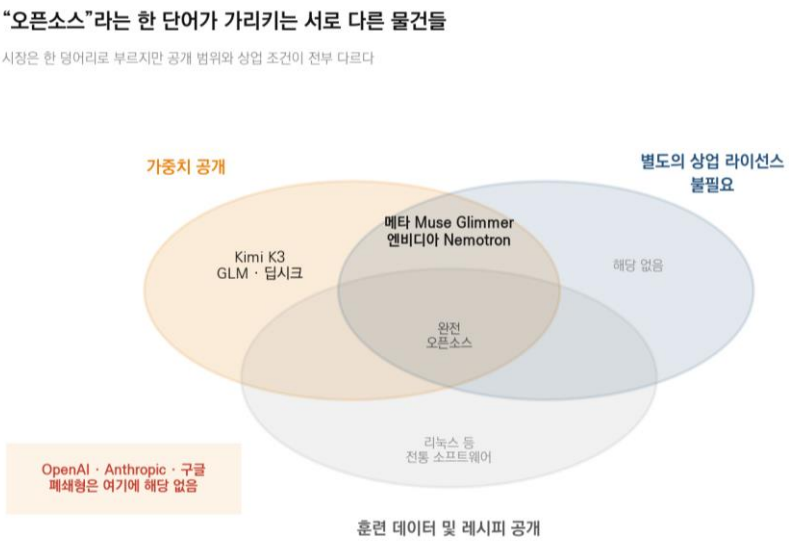
1. 같은 단어인데 사실은 다른 물건

대부분의 사람들은 "오픈소스"를 한 덩어리로 취급한다. 그런데 그 단어는 실제로는 세 개의 서로 다른 조건을 내포하고 있다. 먼저, AI 훈련을 어떻게 한 것인지에 대한 데이터 및 설계도까지 공개했는지 아니면 가중치 수치만 공개했는지가 기준이었다. 그리고 이 모델을 쓰면 누구에게 무엇을 지불하는가도 중요한 조건이다.

엔비디아가 만든 공개 모델 Nemotron이 중국의 오픈웨이트 모델보다 클라우드 사업자에게 나은 이유는 “더 투명해서”라는 것도 있지만, 이 모델을 갖고 추론 사업자들이 되팔아도 베이징에 있는 제작자에게 지급 의무가 생기지 않기 때문이다.

그러니까 오픈소스라는 숭고한 철학이라기 보다는 철저히 손익에 따른 선택이 된다. 그래서 오픈소스의 부상에 따라 오픈소스 진영을 나눠 세부적으로 손익계산서를 따져봐야 한다.

그림 5. “오픈소스”라는 한 단어가 서로 다른 세 가지 조건을 내포한다
내부 수치를 공개했는지, 되팔 때 낼 돈이 있는지, 만든 방법까지 공개했는지에 따라 갈린다.



자료: 미래에셋증권 리서치센터

2. 토큰은 흩어졌고 매출은 뭉쳤다

토큰
AI가 글을 읽고 쓸 때 다루는 최소 단위. 한국어로는 대략 한두 글자 정도에 해당. AI 사용료는 대개 이 숫자로 매기게 된다.

추론 인프라 사업자 아이리스 에너지(이하 IREN)이 전 세계에서 소비되는 AI 토큰을 구간별로 추정한 표를 공개했다. 이 표에서 가장 중요한 점은, **개방형 모델은 토큰을 가져갔지 매출을 가져가지 않았다는 점이다**. 오픈소스 모델이 토큰 생산량의 3분의 1을 차지했는데 매출은 6%뿐이다.

표 1. LLM 추론 시장 세분화 (토큰 점유율 ≠ 수익 점유율)

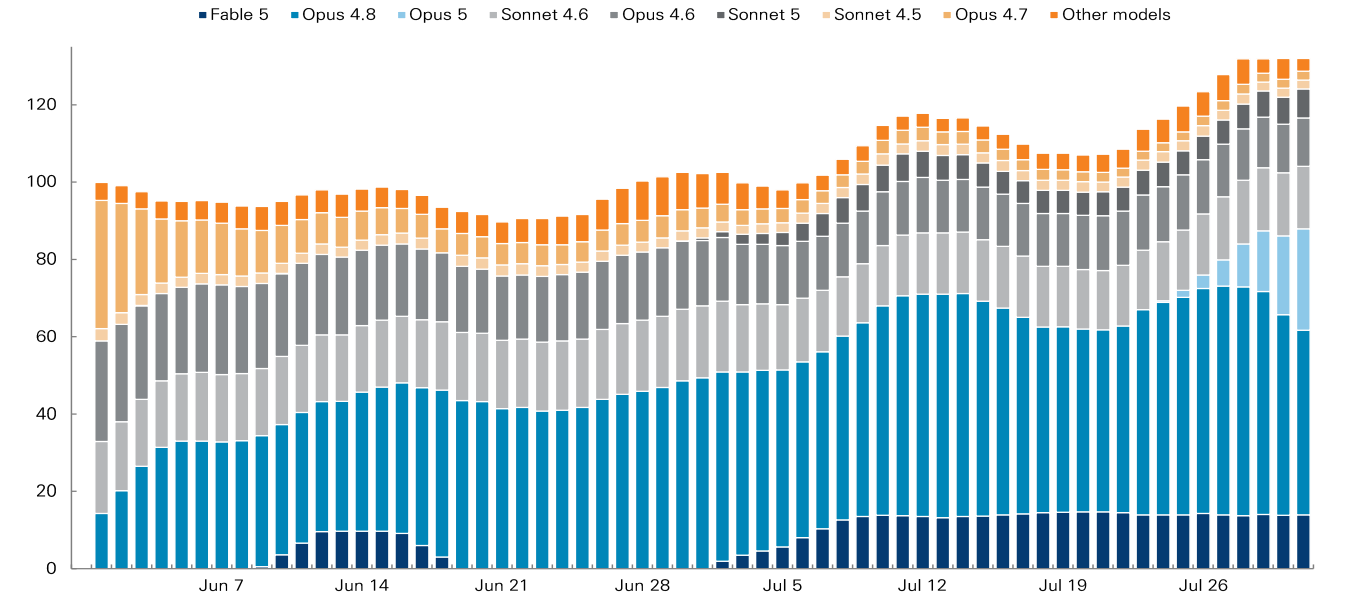
세그먼트	토큰 점유율	수익 점유율	의미
OpenAI	24.4%	27%	규모는 크지만 ASP(평균 판매가격)가 낮음
Anthropic	15.6%	40%	토큰당 단가가 가장 높음
Other paid proprietary	9.8%	약 27%	Gemini, xAI 등
Free tokens	14.6%	0%	챗봇 무료 버전, 허깅페이스 및 데모 사이트 무료 사용
On-prem / self-hosted 오픈소스	19.5%	0%	Neocloud가 접근 불가능한 공간
Hosted 오픈소스 – OpenRouter	1.5%	약 1%	Neocloud가 직접적으로 경쟁하는 업체인 OpenRouter
Hosted 오픈소스 – OpenRouter 외	14.6%	약 5%	Neocloud가 엔터프라이즈 AI에서 노리는 주요 타겟

자료: OpenRouter, IREN, 미래에셋증권 리서치센터
주: 2026년 7월 말 기준

반면, Anthropic은 토큰의 15.6%로 매출의 40%를, OpenAI는 토큰 24.4%로 매출 27%를 가져간다. Anthropic의 API 단가가 OpenAI보다 현저히 높기 때문에 토큰 하나에서 받는 돈이 두 배 넘게 벌어져 있다. 이는 OpenAI가 8월 첫 주에 저가 모델을 들고 오픈소스 모델과 직접 붙으러 내려간 것과 궤를 같이 한다. 같은 진영 안에서도 가격 전략이 갈라지기 시작했다.

그런데 또한 생각해볼 지점은, 모델을 만드는 쪽의 몫이 깎일 수 있다는 가능성이다. 이때 눈여겨볼 부분이 위 표의 5번째에 있는 '온프레미스 셀프호스팅 오픈소스 모델'이다. **전 세계 토큰의 19.5%가 이미 모델을 만든 회사에 한 푼도 내지 않는 방식으로 돌고 있다.** 기업이 모델 수치를 받아다 자기 서버에 깔아서 돌리기 때문이다.

그림 6. Anthropic의 모델별 엔터프라이즈들의 지출에 관한 6월~7월 추이 (7일 이동평균, 6월 1일 전체 지출=100으로 지수화)
Fable 5는 Anthropic 기업 지출의 11%에 머물렀음. 성능이 올라간다고 기업 지출이 따라 오르지는 않는다!
Fable 5는 Opus 4.8보다 가격이 2배 비싸지만, 대부분의 업무에서 성능 향상이 가격 차이를 정당화하지 못함.
ZDR을 지원하지 않고 데이터를 30일간 보존해야 해 금융·법률·의료 등 규제 산업의 채택이 제한되기 때문.
결국 기업은 달러당 한계효용과 컴플라이언스 리스크를 매우 중시, 이는 "good enough"과 "버퍼타임", 그리고 '데이터 주권'와 직결됨.



자료: Ramp, 미래에셋증권 리서치센터

그리고 온프레미스 오픈소스 모델의 수치는 지난 7월 보고서에서 필자가 밝혔듯, 'good enough'와 '버퍼타임', 그리고 '데이터 주권'의 논리에 따라 더 늘어날 가능성이 높다고 사료된다. 그리고 이 영역에서는 **모델을 만든 회사가 아니라 그 모델이 돌아가는 서버와 전기, 그리고 그 모델을 실제 업무에 붙여 주는 영역으로 자금이 몰리게 될 수 있다.**

물론 공개형 모델이 퍼질수록 소수의 프론티어 회사가 오히려 매출을 더 많이 가져간다는 반대 논리도 있다. 하지만 최근에 발표된 결제 데이터는 우리의 논지를 강화해준다.

기업 지출을 위한 관리업체인 Ramp社가 7만 개 이상 사업체의 카드·청구 데이터로 집계한 7월 소프트웨어 벤더 순위를 보면, 점유율이 가장 빠르게 늘어난 곳이 Anthropic이다. 그런데 같은 표의 '급성장' 목록에는 파이어웍스, 베이스텐, 투게더, 오픈라우터 같은 모델 서빙·중개 업체가 줄줄이 올라와 있다.

즉, 매출은 폐쇄형 상위 몇 곳에 계속 뭉치고 있고, 볼륨은 그 아래로 계속 흩어지고 있다. 두 흐름이 동시에 강해지고 있는 것이지 둘 중 하나가 이기고 있는 것이 아니다. 그래서 **투자자가 물어야 할 질문은 "어느 쪽이 이기나"가 아니라 "흩어지는 물량이 지나가는 길목 어디에 통행료를 받는 자리가 있나"가 된다고 판단한다.**

표 2. 2026년 7월 기준 Ramp가 선정한 주요 소프트웨어 기업

순위	주목받는 기업 — 기업 규모 대비 가파른 성장	성장 속도가 가장 빠른 기업 — 시장점유율 증가율 기준
1	PolyAI — AI 고객지원 봇	Anthropic — 기반 대규모 언어모델(LLM)
2	DeepSeek — 기반 대규모 언어모델(LLM)	Granola — AI 회의록 작성 도구
3	Fireworks AI — 모델 서빙 및 추론	Marketo — 이메일 마케팅
4	Sierra — AI 고객지원 봇	Higgsfield AI — AI 영상
5	Baseten — 모델 서빙 및 추론	Adobe — 크리에이티브 제품군
6	DataForSEO — 검색엔진 최적화(SEO)	OpenRouter — 모델 서빙 및 추론
7	Lightdash — 비즈니스 인텔리전스(BI)	Vercel — 웹 호스팅 및 웹사이트 제작 도구
8	Higgsfield AI — AI 영상	Indeed — 채용관리시스템(ATS) 및 채용
9	Howdy — 프리랜서 및 인력 중개 플랫폼	Fireworks AI — 모델 서빙 및 추론
10	Together AI — 모델 서빙 및 추론	Apollo.io — 영업 데이터 제공업체

자료: Ramp의 지출 관리 플랫폼을 이용하는 7만 개 이상 기업의 카드 및 청구서 결제 데이터, 미래에셋증권 리서치센터

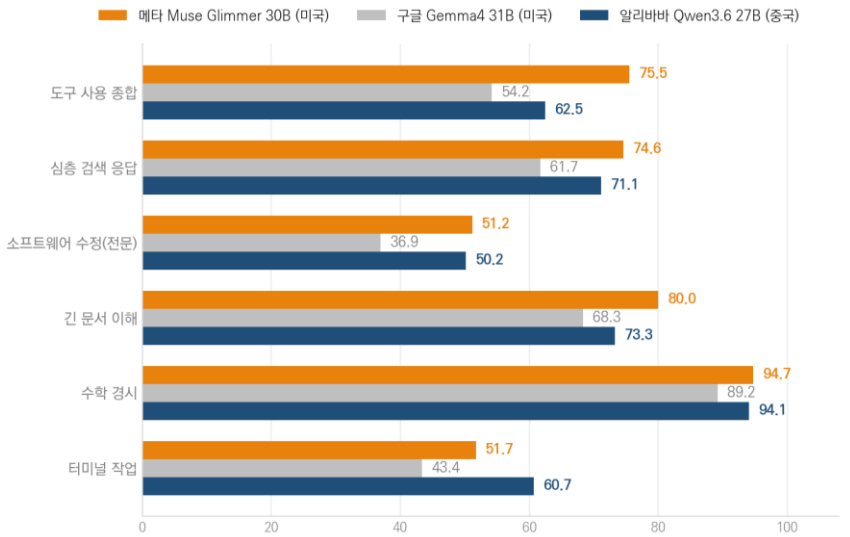
3. 첫 번째 진영 - 서방권 공개 모델

미국 엔비디아의 Nemotron과 Thinking Machines의 Inkling, 유럽의 미스트랄, 일본의 Sakana, 그리고 이번에 다시 합류한 메타 등이 있다. 향후 SpaceXAI의 Grok도 오픈웨이트 출시가 예고되어 있다. 이들은 지리적으로는 서방 진영에 걸쳐 있지만 투자 판단에서는 한 묶음으로 보면 된다. 이들은 라이선스 조건도 유사하고 산업적 이해가 대체로 일치한다. 그리고 무엇보다 미국 반도체에 맞춰 최적화돼 있기 때문이다.

사실 이 진영에 있어서 '개방'이라는 개념은, 신념이라기보다는 전략인 경우가 많다. 자기 것을 열면서 동시에 공개 모델의 한계를 강조하는 포지셔닝이다. 이와 관련해 팔란티어는 최근 실적 발표에서 "엔비디아의 Nemotron 모델을 자사 시스템에 올린 지 24시간 만에, 아무 추가 훈련도 없이 (폐쇄형) 프론티어 모델을 이긴 실제 업무 과제를 다섯 개 찾았다"고 밝혔다.

즉, **프론티어가 실제로 가장 성능이 좋다는 가정은 기업 현장에서는 성립하지 않는다는 것이다**. 기업들의 입사 시험은 대개 범용 시험 문제인 경우가 많지만 그 시험이 회사들의 실제 업무와 같은지는 별개 문제라는 것과 같다.

그림 7. 메타가 오랜만에 출시한 개방형 모델 "Muse Glimmer"의 벤치마크별 점수
미국과 중국의 오픈소스가 비슷한 모델 크기에서는 두 진영이 비등하다는 뜻이다.



자료: Meta, 미래에셋증권 리서치센터

저커버그가 Muse Glimmer를 출시한 시점 즈음에 작성한, 장문의 블로그 글의 뒷부분에 흥미로운 문장이 하나 있다. 그는 어떤 하나의 주체도 초지능을 독점해선 안 된다고 쓰면서, 뒤에서는 '자기 개선 경쟁'에서 뒤처지지 않으려면 여러 연구소가 "집합적으로 충분히 큰 컴퓨트를 지어야 한다"고 쓴 것이다. 앞에서는 흩어지라 하고 뒤에서는 모이자는 글이다.

모순처럼 보이지만 아니다. 나눠 주라는 건 모델 층 이야기이고, 크게 지으라는 건 컴퓨트 층 이야기다. 개방 진영을 대표하는 저커버그의 문서 안에 'AI 모델은 흩어지고 컴퓨팅 인프라는 뭉친다'는 구조가 그대로 들어 있다는 점을 유념해야 한다.

흥미롭게도 오픈소스에 가장 비판적인 쪽에서도 같은 말이 나왔다. Anthropic의 연구자 숄토 더글러스는 저커버그의 글을 지적하면서, '개인이 강력한 AI를 갖는 데 결정적인 변수는 모델이 공개됐는지가 아니라 컴퓨트에 얼마나 접근할 수 있느냐'의 불평등이라고 했다.

오픈소스를 옹호하는 쪽과 반대하는 쪽이 서로 다른 결론을 주장하면서 같은 전제 위에 서 있다. "모델은 흔해지고, 희소한 것은 그 아래에 있다."

4. 두 번째 진영 — 중국 공개 모델

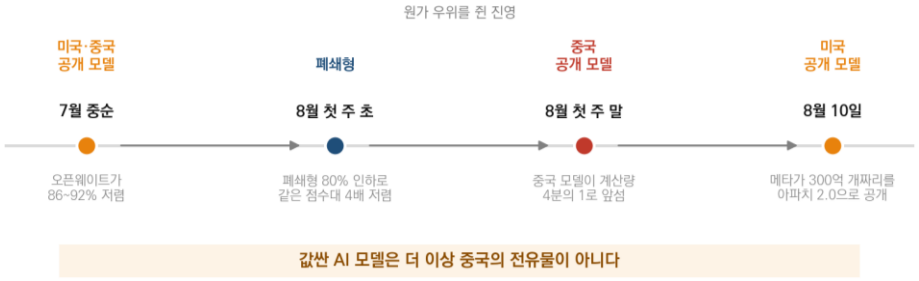
Kimi K3, GLM, 딥시크, 알리바바 Qwen이 대표적이다. 앞의 성능표에서 봤듯이 실력은 미국 오픈소스 모델 진영과 이제 그렇게 큰 차이는 없어보인다. 중국의 오픈웨이트 진영의 무기는 하나였다. 값이다. "같은 일을 훨씬 싸게 시킬 수 있다는 것"이고 이것이 AI 인프라 투자 시장을 bearish하게 보는 이들의 주요 논지이기도 했다. 그런데 값이 싸다는 것만으로는 우위를 유지하는 것이 힘들어지는 이유들이 생겼다.

표 3. 중국 오픈웨이트 진영이 갖고 있는 문제들

중국 오픈웨이트를 쓸 때의 부담	이것이 함의하는 것
되팔 때 돈을 내야 한다	이 모델로 장사하는 추론 사업자는 개발사와 별도의 수익계약을 맺어야 한다
미국의 제재 혹은 산업계 자체 규정	굳이 이러한 위험을 지고 중국 것을 쓸 이유가 적다
중국 반도체에 맞춰져 있다	엔비디아 칩을 쓰는 곳에서는 그 이점이 그대로 나오지 않는다
값 우위가 흔들리고 있다	8월 첫 주에 폐쇄형이 저가 모델 값을 80% 내렸다

자료: 미래에셋증권 리서치센터

그림 8. 비슷한 AI 모델의 성능지수를 기준으로 원가 우위는 누가 가지고 있는가
결국 모델의 가격과 성능은 빠르게 평준화되고, 값싼 모델은 더 이상 중국의 전유물이 아니게 됐다.



자료: 미래에셋증권 리서치센터

특히 미국의 큰 금융기관이나 대기업은 규정 때문에 중국 모델을 쓰기 어렵다. 그동안은 마땅한 대안이 없어서 감수하거나 아예 AI 사용을 포기했었다. 그런데 이러한 것들이 걸릴 게 거의 없는 미국의 오픈소스 모델들이 다시 출현하는 시점에서는, 굳이 규정 위험을 지고 중국 것을 쓸 이유가 사라진 것이다. Anthropic의 가장 강력한 모델인 Fable이 데이터 보관 조건 때문에 규제 산업에서 외면을 받는 것과 대동소이 하다. 값싸다는 무기는 성능 경쟁에서는 통하지만 규정 경쟁에서는 통하지 않는 법이다.

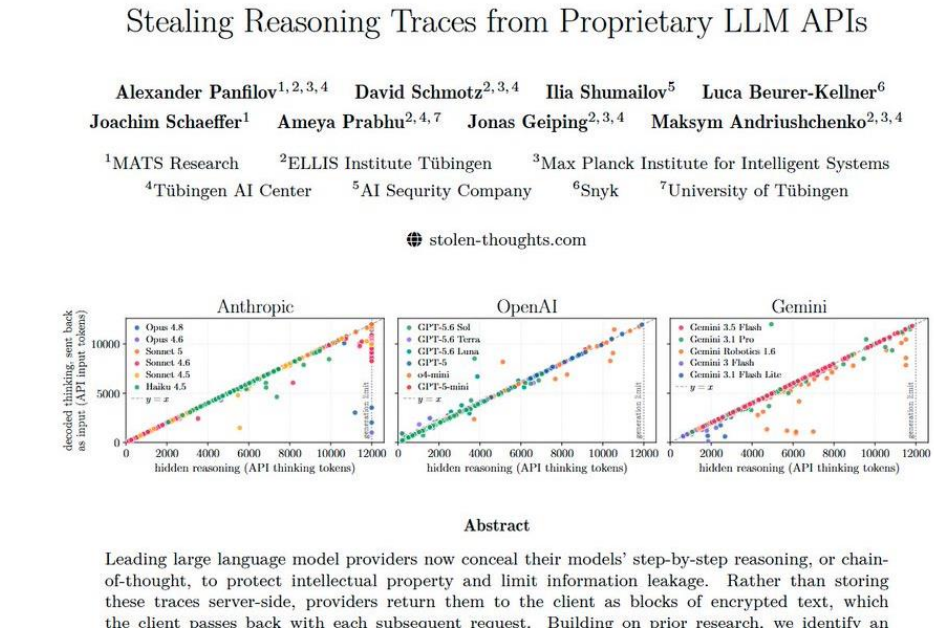
5. 세 번째 진영 — 폐쇄형

OpenAI, Anthropic, 구글이 대표적이다. Good enough 오픈웨이트 진영의 추격에 원가 압축을 자기 모델에 직접 실행하고 있다. 그럼에도 구조적 제약은 남아 있다.

前 OpenAI 직원인 앤드루 호에 따르면, “프론티어 회사는 경쟁 때문에 다음 세대 훈련에 계속 돈을 써야 하는데, 그 금액이 지금 벌어들이는 돈보다 항상 크다”. 그의 표현으로 선두 주자를 가혹하게 벌하는 구조다. 쉽게 말하면 이렇다. 1등이 되려면 계속 더 비싼 실험을 해야 하는데, 2등은 1등이 만든 걸 싸게 베낄 수 있다.

이와 관련해 구글의 전성기를 이끌었던 전설인 에릭 슈미트가 이번 달에 이런 말을 한 적이 있다. "모델 하나에 2억에서 5억 달러가 들고, 출시하면 수명이 6개월쯤이다. 그리고 또 다른 모델이 나온다. 경쟁이 워낙 잔혹해서 수억 달러짜리 물건의 가치를 몇 주 단위로 세야 하는 것이다." 즉, 만드는 쪽에서 보면 어찌보면 벌칙에 가까운 구조라고 볼 수도 있다.

그림 9. 프론티어 모델을 증류하는 것을 막는 "안티증류의 우회 방법"에 관한 논문이 발표
중국의 증류를 '규제'하기 전에, 빅테크가 자사 API의 구조적 보안 결함부터 해결해야 한다는 논문



Stealing Reasoning Traces from Proprietary LLM APIs

Alexander Panfilov[1,2,3,4] David Schmotz[2,3,4] Ilia Shumailov[5] Luca Beurer-Kellner[6]
Joachim Schaeffer[1] Ameya Prabhu[2,4,7] Jonas Geiping[2,3,4] Maksym Andriushchenko[2,3,4]

[1]MATS Research [2]ELLIS Institute Tübingen [3]Max Planck Institute for Intelligent Systems
[4]Tübingen AI Center [5]AI Sequrity Company [6]Snyk [7]University of Tübingen

stolen-thoughts.com

Abstract

Leading large language model providers now conceal their models' step-by-step reasoning, or chain-of-thought, to protect intellectual property and limit information leakage. Rather than storing these traces server-side, providers return them to the client as blocks of encrypted text, which the client passes back with each subsequent request. Building on prior research, we identify an

자료: Alexander Panfilov 외 7인, 미래에셋증권 리서치센터

하지만 이러한 비대칭은 곧 막을 내릴 가능성이 높다. 공론화되고 있기 때문이고, 이에 따라 "싸게 베낀다"는 전제가 흔들리고 있기 때문이다. 위 그림에 있듯, 지난 8월 11일에 나온 연구가 그 전제를 건드렸다.

지금의 AI는 답을 내기 전에 혼자 생각하는 과정을 거친다. 그 과정을 '추론 흔적(ReasoingTraces 혹은 Chain of Thought)'이라고 부른다. 물론 회사들은 이걸 그대로 보여 주지 않는다. 자기 모델이 어떻게 생각하는지가 그대로 드러나면 남이 베끼기 쉬워지기 때문이다. 그래서 암호로 잠가서 돌려준다.

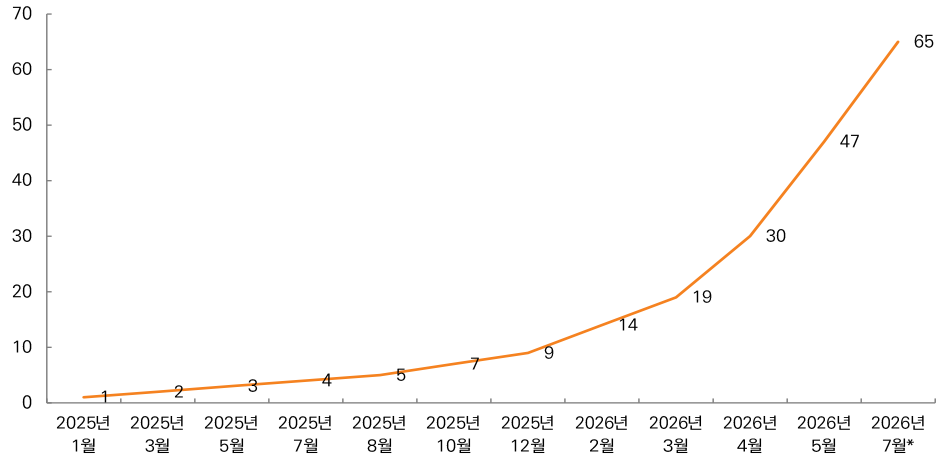
증류
잘하는 모델에게 문제를 잔뜩 풀게 한 뒤 그 답을 교재 삼아 작은 모델을 가르치는 방법. 처음부터 가르치는 것보다 훨씬 적은 비용이 든다.

이번에 연구팀이 찾아낸 건 그 자물쇠의 허점이다. 같은 회사의 고성능 AI 모델에서 받아낸 그 암호 덩어리를, 같은 회사에서 만든 소형 모델에 집어넣으면 소형 모델이 그걸 그냥 복호화로 풀어서 그게 뭔지 읽어 준다는 습성(?)을 이용하는 것이다. 그러니 고성능 모델을 직접 뚫을 필요가 없이, 이렇게 우회하여 증류하는 것이 쉽다는 것이다. Anthropic, OpenAI, Google 모두에서 같은 방법이 통했다고 논문은 밝히고 있다.

유명 AI 개발자 Nathan Lambert는 "중국 방문 당시 여러 연구소가 이미 이것과 유사한 기법을 사용한다는 암시를 받았다"고 말했다. **지금까지 중국을 포함한 오픈웨이트 후발 주자가 폐쇄형 프론티어를 빠르게 따라잡은 통로 중 하나가 여기였다면, 그 통로는 이제 곧 빠르게 닫힐 수 있다.** 닫히면 증류의 속도는 매우 느려질 수 있다는 점이 포인트다.

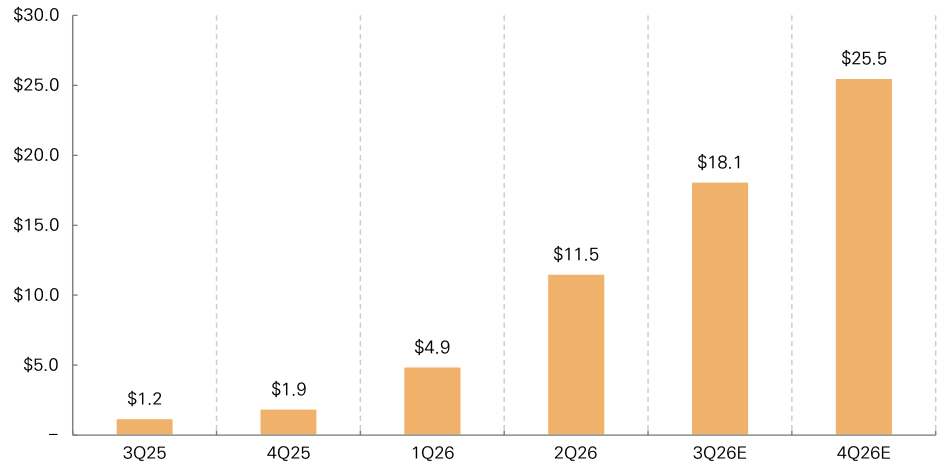
한편, 폐쇄형의 대표주자인 **Anthropic의 연환산 매출은 2025년 1월 US$1bn에서 2026년 7월 US$65~74bn으로 60배 넘게 늘었다. 비용이 매출을 앞지르던 시기는 지나가고 있다. 앞으로 이 진영을 압박할 것은 다음 모델의 훈련비가 아니라, 공개 모델이 값을 계속 끌어내릴 때 추론에서 남기는 이익률이다.** 이 부분은 2부에서 컴퓨트 원가와 함께 다룬다.

그림 10. Anthropic의 연환산 매출 런레이트 (단위: US$bn)
2026년 7월 말 연환산 매출 런레이트는 650억 달러를 초과한 것으로 공식 보도.
대체데이터 기반 외부 추정치 회사인 TickerTrends는 같은 기간 수치를 741억 달러로 추정.



자료: Bloomberg, Reuters, TickerTrends, Anthropic, 미래에셋증권 리서치센터

그림 11. Anthropic 분기 매출 추이 및 전망 (3Q26E와 4Q26E는 추정치, 단위: US$bn)
2025년 3분기 12억 달러에서 2026년 2분기 115억 달러로 불과 3개 분기 만에 9.7배 증가
API와 기업용 구독 매출의 동반 확대로 2026년 연간 매출은 600억~700 달러에 이를 것으로 추정



자료: Anthropic, Bloomberg, 미래에셋증권 리서치센터

6. 오픈소스에도 국경과 가격표가 있다

표 4. 미국 오픈웨이트의 등장으로 중국 오픈웨이트의 장점이 희석

항목	미국 공개 모델	중국 공개 모델	폐쇄형
가중치 공개	공개	공개	비공개
되팔 때 낼 돈	없음 / 적음	있음	있음(사용량 과금)
기업이 직접 쓸 때	서버값만	서버값만	토큰 과금
해당 성능지수상 비용	중간	설계발 원가 우위 등장	8월 초 최저 구간 진입
제재·규정 위험	현재로서 비교적 낮음	높음, 점점 더 높아짐	일정부분 규제 대상
반도체 종속	미국	중국 칩 병행	미국
AI PC 등 로컬 구동	일부 가능	일부 가능	불가
AI 기록 소유	엔터프라이즈(사용자)	엔터프라이즈(사용자)	계약별로 다름
토큰 / 매출 점유율	공개모델 합계 35.6% / 약 6%		49.8% / 94%

자료: 미래에셋증권 리서치센터

정리를 해보면, 오픈소스가 폐쇄형을 이긴다는 주장을 하고 싶은 것은 아니다. 서로 다른 세 가지 경제가 만들어지고 있다는 것이다.

서방권 공개 모델은 모델 자체에서 사용료를 걷기보다 생태계를 넓히는 데 목적이 있다. 메타는 광고가 있고, 엔비디아는 반도체가 있으며, 클라우드 사업자들은 서버 사용료를 받을 수 있다. 모델을 무료로 공개해도 그 모델이 돌아가는 다른 층에서 돈을 벌 수 있다.

중국 공개 모델은 다르다. 가격과 성능에서는 가장 공격적이지만, Kimi K3의 사례처럼 재판매 조건과 제재, 규정 준수라는 문턱이 붙는다.

폐쇄형은 토큰을 판매해 직접 매출을 만든다. 현재 토큰 점유율은 과반에 못미치지만 매출 점유율에서는 94%를 가져가는 이유다. 그러나 선두를 유지하려면 조 단위로 다음 모델을 계속 만들어야 하고, 공개 모델이 따라올 때마다 가격을 다시 내려야 하는 압박이 있다.

따라서 "오픈소스가 커진다"는 말만으로는 누가 돈을 버는지 알 수 없다. 더구나 모델을 고르는 기준도 성능 하나가 아니다. 가격이 싸도 라이선스가 맞지 않으면 외면받고, 성능이 좋아도 규정 준수 문제에 걸리면 기업은 사용할 수 없다. 반대로 그 모델이 미국산이라는 이유, 기업 데이터를 밖으로 내보내지 않는다는 이유로 'fair enough'하게 선택될 수 있다.

결국 **오픈소스에도 국경과 가격표가 있다는 것을 명심하자**. 가중치를 내려받을 수 있다는 사실만으로 국적과 계약, 제재와 규정 준수가 사라지는 것은 아니다.

III. 모델이 흔해지면 떠오르는 '가치'는 무엇일까

모델이 흔해진다는 건 AI 모델 하나로 돈 벌기가 어려워진다는 뜻이다. 그렇다면 부가가치를 창출하는 곳은 어딜까?

1. 쌓일수록 값어치 있는 학습의 기록

이전 장에서 살펴봤듯, 세계에서 가장 강력하다는 모델 Fable이 엔터프라이즈들의 AI 지출에서 더 많은 영역을 차지하지 못하는 이유 중 하나는 데이터 주권에 관한 것이다. Fable은 "AI와 주고받은 내용을 30일간 Anthropic이 보관해야 한다는 조건"이 걸려 있는 모델이기 때문이다. 모델이 아무리 좋아져도, 그 기록을 누가 갖느냐가 정해지지 않으면 팔리지 않는다. 사실 이 한 문장이 이 장 전체의 요약이다.

그렇다면 왜 기록(System of Record)이 그렇게 중요한지부터 생각해봐야 한다. AI를 쓰다 보면 기록이 남는다. 어떤 질문을 했고, AI가 뭐라고 답했고, 틀렸을 때 사람이 어떻게 고쳤는지의 기록이다. 이 기록으로 다시 AI를 가르치면 우리 회사 일에 딱 맞는 AI가 된다. 문제는 이 기록이 누구 손에 쌓이느냐다.

지난 한 달 간 이것의 중요성과 관련해서 여러가지의 실물적 증거를 찾을 수 있었다.

첫째, **AI 학습 데이터를 모으는 대표적 스타트업인 Mercor는 그러한 '기록'을 확보하려고 개별 도메인 회사와 계약할 때 최소 수억 원을 지불한다.**

그림 12. OpenAI·Anthropic·Google 같은 AI 연구소에 '훈련 재료와 시험장'을 공급하는 기업들
법률·회계·의료 전문가들을 고용해 평가 기준, 실제 업무수행 기록을 만들어 훈련 데이터로 판매
결국 전문인력 데이터를 모델 성능 향상에 사용 가능한 데이터와 평가 인프라로 가공하는 공급업체

회사	설립	총매출 연환산액	기업가치
대형 사업자 총매출 연환산액 10억 달러 이상			
Scale AI	'16	약 15억 달러 e	290억 달러 M
Surge AI 자력 성장	'20	약 30억 달러 †	250억 달러 †
Mercor 사업 전환	'23	20억 달러 이상	100억→200억 달러 †
Handshake 사업 전환	'13	약 10억 달러	35억 달러*
도전자 그룹 총매출 연환산액 1억 달러 이상			
micro1 사업 전환	'22	2.5억 달러 이상	약 25억 달러
Turing 사업 전환	'18	약 5억 달러	22억 달러
강화학습 환경 전문기업 '짐(gym)' 붐권			
Fleet 3분기 1.6억 달러 전망	'24	6,300만 달러	7.25억 달러
데이터 전문기업			
Arena 구 LMArena, 선호도·평가	'23	3,000만 달러	17억 달러
Truveta 헬스케어 사업 전환	'20	자료 없음	약 10억 달러
전환 중인 기존 사업자			
Snorkel AI	'19	자료 없음	13억 달러
Labelbox Aligner	'18	자료 없음	약 10억 달러*
합계 11개사		약 83.43억 달러+	약 779억~879억 달러

자료: Deedy Das, 외신, 미래에셋증권 리서치센터
주: e는 최근 공개연도 수치에서 외삽한 수치 / †는 투자 협의 중이라는 뜻/ *는 2022년의 오래된 평가액 / M은 Meta의 지분이 49%

둘째, 이달에는 그 값어치가 법원 문서로까지 공개됐다. 파산한 미국 저비용 항공사 스피릿 항공의 34년치 내부 운영 기록이 경매에 나온 것이다. 이메일, 사내 채팅, 업무 문서, 코드까지 전부다. 낙찰자는 바로 구글이었고 낙찰가는 US$10mn이었다. 차순위 입찰자가 US$7.5mn을 써낸 앞의 Mercor다.

그림 13. Google과 Mercor의 Spirit Airlines 비식별 데이터(개인식별정보를 제외) 확보 경쟁
Google이 1,000만 달러로 낙찰자에 선정, Mercor는 750만 달러로 예비 낙찰자에 지정

Successful Bidder	Alternate Bidder
Google LLC Total Consideration (USD): $10,000,000	Mercor.io Corporation Total Consideration (USD): $7,500,000

자료: 미국 뉴욕남부연방파산법원, 미래에셋증권 리서치센터

한때 **기업가치 US$6bn을 평가받던 회사가 34년간 실제로 어떻게 일했는지의 기록 전체가 US$10mn에 팔렸다는 뜻이다. 회사는 망했는데 그 회사가 일한 기록은 팔렸다는 점이 중요하다.** 사업의 성패와 무관하게, 실제 조직이 실제로 일한 흔적 자체에 시장가격이 붙기 시작했다는 뜻이기 때문이다.

셋째, 8월 3일 팔란티어 실적 발표에서 이런 말이 나왔다. "(AI를 엔터프라이즈가 활용할 때) 핵심이 되는 것은 기업 안에 있는 (raw) 데이터만이 아니다. **메타데이터, 추론 흔적, 사용 기록이며, 아직 이것을 통제할 장치가 없다. 그리고 그것이 기업의 내부 데이터보다 더 가치있다는 것을 이해해야 한다.**"

이러한 것은 엔터프라이즈 AI를 위한 '학습 기록'의 소유권이 더욱 중요한 가치가 됐다는 것이고, 이는 곧 기업들이 AI를 선택하는 결정적 사안이 된다.

2. 왜 이 기록은 모델에 흡수되지 않을까

그런데 AI 모델이 계속 좋아지면 결국 기업의 기록이라는 가치도 떨어질 수 있지 않을까? 이에 대해서는 사실 실리콘밸리에서도 논란이 있는 뜨거운 주제다. 그리고 이에 대해서는 사라 구오라는 벤처캐피탈리스트가 주장한 내용에 필자도 공감한다. 그녀가 든 현장감 있는 예시는 다음과 같다.

공장에는 모터를 만드는 방법이 작업지시서에 적혀 있다. 어떤 부품과 접착제를 쓰고, 몇 도에서 굳히며, 조립 전에 무엇을 닦아야 하는지까지 나와 있다. 처음 보면 누구나 지시서만 따라 하면 같은 제품을 만들 수 있을 것 같다. 그런데 생산을 시작한 지 3주쯤 지나자 일부 모터가 진동 시험을 통과하지 못했다. 불량은 주로 야간 근무조에서 나왔고, 경험이 적은 직원이 작업할 때 더 자주 발생했다.

원인은 작업지시서에 없었다. 밤에는 공장의 습도가 달라져 접착제가 제대로 굳지 않았고, 신입 직원은 부품을 너무 꼼꼼하게 닦은 나머지 접착을 방해하는 세척제 성분을 남겼다. 경험 많은 직원이라면 금방 알아챌 문제였지만, 처음 맡은 사람은 원인을 찾는 데 몇 주가 걸릴 수 있었다.

해결책은 간단했다. 야간에도 제습기를 계속 가동하고, 부품을 닦을 때 세척제를 한 번만 쓰면 됐다. 문제는 이런 노하우가 작업지시서에는 없다는 것이다. 경험자의 머릿속이나 어느 공장의 사고 보고서에만 남아 있을 뿐, 다른 공장이나 직원에게 제대로 전달되지 않는다.

기업의 진짜 경쟁력은 설계도에 적힌 지식만이 아니다. 현장에서 시행착오를 겪으며 쌓인, 아직 문서가 되지 않은 지식에도 있다. 위와 같은 사례는 수 백만 개로 존재할 수 있고 그 지식은 사람들 안에 저장되어 있다.

따라서, **모델이 아무리 좋아져도 흡수할 수 없는 지식이 있다. 글로 쓰인 적이 없기 때문이다**. 그것을 기계가 읽을 수 있는 형태로 만드는 일에 가격과 가치가 붙는 이유가 여기 있다. 그리고 이 일은 AI 모델을 교체한다고 하더라도 또 다시 하지 않아도 되는 일이다. 한 번 만들어 두면 어느 모델을 갖다 붙여도 쓰인다. 필자가 이 시리즈에서 "데이터와 선택권을 통제하는 자리"라고 강조하는 것 중 절반이 바로 이 영역의 비즈니스다.

여기서 한 가지 현상이 설명된다고 생각한다. AI를 쓰면 개인의 생산성은 확실히 좋아지는데, 회사 전체로는 AI를 써서 좀처럼 효과가 안나오는 것 같은 현상 말이다.

먼저 개인은 왜 빠를까? 본인이 스스로 아는 걸 그 자리에서 AI에게 바로바로 던질 수 있기 때문이다. "이 고객은 이런 걸 싫어하더라", "이 방식은 내가 작년에 해 봤는데 망했다" 같은 말을 대화하듯 AI에게 넣는다.

어디에도 적혀 있지 않은 지식이 실시간으로 AI에게 들어가는 셈이다. 그래서 AI가 그 사람의 두 번째 뇌처럼 작동한다.

하지만 회사 차원으로 스케일링을 하면 이게 작동하지 않는다. 기업이 가진 암묵지는 한 사람의 머릿속에 있지 않다. 수많은 임직원과 부서, 문서, 이메일, 업무 시스템에 흩어져 있다. 어느 공장의 숙련자가 아는 생산 요령, 영업 담당자가 체득한 고객의 성향, 재무 담당자가 알고 있는 예외 처리 기준이 서로 연결되지 않은 채 따로 존재한다.

대기업이 데이터를 많이 보유하고도 AI 활용에서 기대만큼의 효율을 내지 못하는 이유가 여기에 있다. 더 큰 근본적 이유는, “딱히 기록할 이유, 공유할 이유”가 없기 때문이다. 개인은 자기가 던진 지식으로 자기가 곧바로 편해진다. 그런데 직원은 자기가 아는 걸 회사 시스템에 넣어도 딱히 돌아오는 게 없다. 오히려 자기만 알던 것을 내놓는다고 생각하는 경향이 있다. 개인 단위에서는 이해가 맞아떨어지지만 조직 단위에서는 어긋난다는 것이다.

다른 문제는 지식을 아는 사람과 실제로 쓰는 사람이 다르다는 것이다. 개인은 자기가 던지고 자기가 답을 보고 그 자리에서 피드백을 주면서 고친다. 회사에서는 지식을 가진 사람과 AI의 답을 받는 사람이 다른 층에 있다. 틀려도 누가 고쳐야 할지 모른다.

온톨로지
회사 안의 사람, 물건, 공정, 계약 같은 것들이 서로 어떤 관계인지를 기계가 알아볼 수 있게 정리해 둔 가상의 지도.

이런 문제들 때문에 심지어 사내 클라우드를 깔아 놓고도 잘 안 쓰는 회사가 많다. 그릇은 있는데 부을 것이 없는 상태다. 여기가 엔터프라이즈 AI에서 가격이 붙는 자리라 판단한다. 흩어진 지식을 기계가 읽을 수 있는 형태로 모으고, 누가 무엇을 볼 수 있는지 정하고, 틀렸을 때 누가 고치는지를 정해 주는 일이다.

업계에서는 이것 중 하나를 온톨로지라고 부른다. 그리고 이 지도는 모델을 바꾼다고 다시 그리지 않는다. 한 번 그려 두면 어느 모델을 갖다 붙여도 쓰이기 때문이다.

3. 방어의 값

허깅페이스 해킹 사례를 보고, 필자는 '방어를 위한 자체 서버 수요'가 생겼다고 봤다. 외부 API가 해킹 포렌식 분석을 거부하자 자기 서버에 공개 모델을 깔아서 해결했기 때문이다. 그런데 오픈소스 모델을 손에 넣은 사이버 공격자들이 기업을 노리기 때문에, **기업이 스스로를 지키려고 오히려 최고 성능 모델에 더 돈을 쓰게 될 수도 있다.**

이에 따른 가장 정합적인 예측을 해보자면, 하이브리드 구조의 확산으로 판단한다. 평소 업무를 처리할 값싼 모델, 어려운 문제를 풀 최고 성능 모델, 그리고 외부 서비스가 멈춰도 작동할 자체 연산능력이다. 이 3층 구조가 얼마나 오래 갈지에 대해 최근 힌트도 나왔다. Anthropic의 연구자 숄토 더글러스는 "앞으로 2년 동안 모든 금융기관과 핵심 인프라 사업자가 스스로를 미리 해킹해 보는 작업을 해내야 한다"고 말했다. 그 작업이 끝나면 어떤 공개형 모델이 풀려도 걱정하지 않는다고 이유를 달았다.

투자자에게 이 말은 "2층 모델(보안과 위기 대응)"의 존재 목적이 해킹 사고가 난 뒤의 분석뿐만 아니라, 사고가 나기 전에도 자기를 공격해 보는 상시 업무가 된다는 말이 된다.

표 5. 엔터프라이즈 AI는 업무의 난도/민감도에 따라 세 가지를 동시에 확보하려 할 가능성이 높다

층	무슨 일에 쓰나	무엇을 쓰나
1층 · 일상 업무	문서 요약, 코드 보조처럼 대부분의 일	값싼 Good enough 모델
2층 · 보안과 위기 대응	침해 분석, 위협 탐지처럼 실패 비용이 큰 일	회사 전용 격리 환경에 폐쇄형 최고 성능 모델
3층 · 예비	위 두 층이 막혔을 때의 비상시	자기 서버에 깔아 둔 공개 모델

자료: 미래에셋증권 리서치센터

즉, 매일 돌아가는 수요로써 이 다층 및 하이브리드 구조의 확산은 필연적이라는 뜻이다. 여러 모델 사이에서 업무를 배분하는 관제 소프트웨어의 수요가 함께 늘어날 것으로 본다.

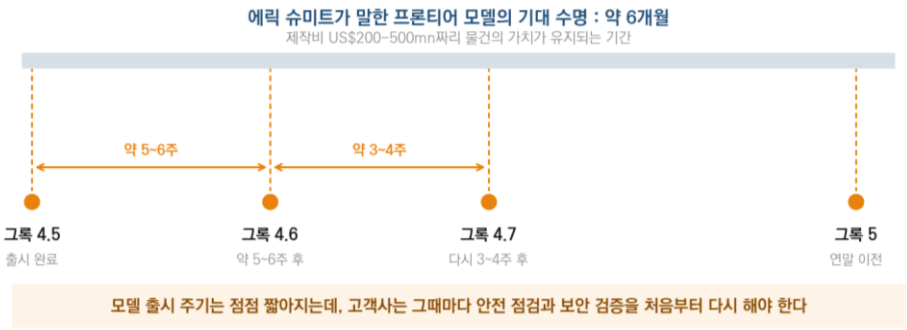
4. 좋은 모델보다 바꾸기 쉬운 구조가 오래 간다

좋은 AI 모델이 새로 나왔다고 기업이 곧바로 기존 모델을 교체할 수 있는 것은 아니다. 기업이 AI를 실제 업무에 넣으려면 검증부터해야 한다. 같은 질문에 일관된 답을 내놓는지, 고객 정보를 밖으로 내보내지 않는지, 회사가 요구한 형식으로 결과를 만드는지 확인해야 한다. 금융이나 의료처럼 규제가 강한 산업에서는 감사와 승인도 받아야 한다.

문제는 이 검증이 특정 모델과 특정 버전에만 유효하다는 점이다. 모델이 바뀌면 답변 내용뿐 아니라 오류가 발생하는 방식도 달라진다. 기존 모델에서 문제가 없었던 업무가 새 모델에서는 실패할 수 있다. 따라서 새 모델을 도입할 때는 그 모델이 연결된 업무를 다시 시험해야 한다.

그런데 모델의 교체 주기는 갈수록 짧아지고 있다. 에릭 슈미트는 프론티어 모델의 수명이 약 6개월에 불과하다고 말했지만, 요새 돌아가는 AI 판도를 보면 실제 출시 일정은 이보다도 훨씬 빠르다. SpaceXAI가 공개한 출시 일정을 보면 몇 주 단위로 신모델 Grok이 나온다는 것을 알 수 있다. (Grok 4.6은 한국 시각 8월 13일 새벽에 출시)

그림 14. SpaceXAI의 AI 모델 Grok의 출시 주기가 점점 빨라지고 있다



자료: 미래에셋증권 리서치센터

이러한 현실은, 앞 장에서 살펴봤듯, 결국 특정 모델에 업무를 묶어 두지 않고, 필요할 때 더 나은 모델로 바꿀 수 있게 해 주는 '관제의 영역'이 가치를 발생시킬 수 있다. 기업의 데이터와 권한, 업무 절차와 검증 기준은 그대로 두고 모델만 교체할 수 있다면 도입 비용이 크게 낮아지기 때문이다. 반대로 이 모든 것이 특정 모델에 맞춰져 있다면, 모델을 바꿀 때마다 데이터 연결과 보안 설정, 업무 절차를 다시 만들어야 한다. 이러한 가치를 최근 실적 발표 때 마치 세 회사가 짠 듯이 함께 강조한 것도 흥미롭다.

- 팔란티어 CEO: "이미 미 정부 사업에서 모델을 갈아 끼울 수 있는 제품이 있다."
- 아마존 CEO: "몇 년 안에 비슷하게 좋은 모델이 최소 여섯 개는 나올 것이다."
- 클라우드플레어 CEO: "모델을 가리지 않는 배포가 중요하다"

세 회사가 파는 제품은 다르지만 고객에게 하는 약속은 같다. '어느 모델이 승자가 되더라도 고객의 시스템을 처음부터 다시 만들 필요가 없게 해 주겠다'는 것이다. **모델의 수명이 짧아질수록 여러 모델을 비교하고 검증하고 교체할 수 있게 해 주는 중간 계층의 수명은 길어진다**.

사는 쪽에서도 논거가 되는 실제 사례가 나왔다. 필자도 겪었지만 대부분의 AI 유저들은 Anthropic의 최상위 모델인 Fable을 쓰다가, 안전 분류 장치가 작업을 자꾸 하위 모델인 Opus로 되돌려 보내는 것을 겪는다. 그리고 한 번 다운그레이드가 되면 그 대화 안에서는 되돌릴 방법이 없다. 그럴 바에는 처음부터 다른 회사 모델로 시작하겠다고 판단하는 사람이 늘어나고 있는 추세다. 모델이 나빠져서 옮긴 것이 아니다. 쓸 수 있느냐 없느냐가 요청마다 달라지는 것을 견디지 못해 옮긴 것이다. 그리고 옮기는 데 큰 비용이 들지 않았기 때문에 즉시 옮겼다.

승자는 계속 바뀐다. 그렇기 때문에 승자를 바꿔 탈 수 있는 구조에 값이 붙는다.

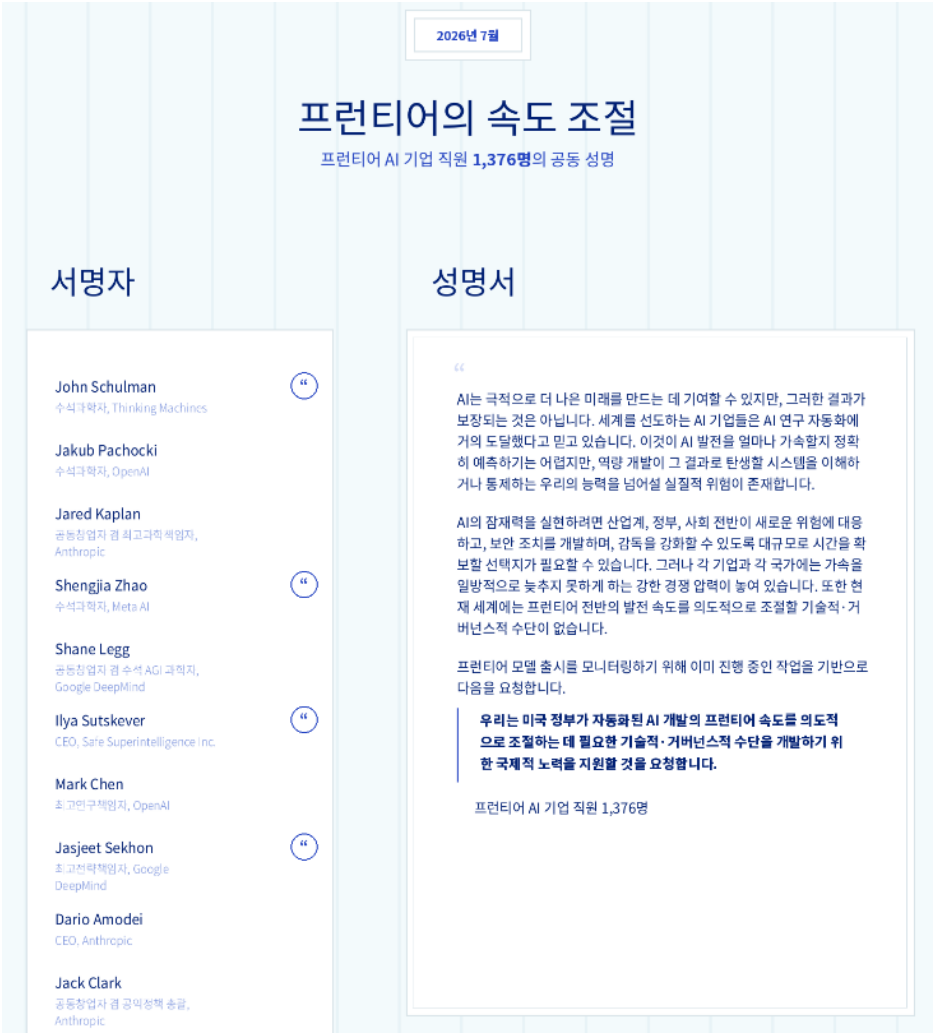
IV. 세 갈래의 규제, 그리고 1부 결론

1. 가능한 규제 목록

먼저 AI로 발생할 수 있는 사고를 막는 규제다. 7월 23일 미국 하원에서 양당 의원이 함께 법안을 발의했다. "AI가 통제를 벗어났을 때 강제로 멈출 수 있는" 절차를 다루며, 대상은 프론티어 모델을 만드는 폐쇄형 회사들이다. 허깅페이스 사고 발생 엿새 만에 나왔다.

다른 가능성 있는 규제는 AI의 능력을 재고, 속도를 조절하자는 규제다. 데미스 하사비스는 몇 주 전 금융업계의 자율규제기구를 본뜬 기구를 제안했다. AI 모델들이 기록하는 벤치마크 점수로 프론티어급인지 아닌지를 각각 정의하고, 어느 나라에서 만들었는지, 공개했는지 여부와 무관하게 적용하되 스타트업과 학계는 면제하는 구조다. 이와 비슷하게 지난 달부터는 프론티어 AI 회사 직원 수 천명이 AI의 개발 속도를 의도적으로 조절할 수단을 마련해 달라는 성명에 서명을 받고 있기도 하다.

그림 15. OpenAI·Anthropic·구글 등 AI 기업 직원들이 개발 속도를 조절할 제도적 장치를 요구
기업 간 경쟁 때문에 안전 검증을 생략하지 않도록 정부가 공통 규칙과 감시 체계를 만들자는 것



자료: pacingthefrontier, 미래에셋증권 리서치센터

마지막은 국가별로 가르는 규제다. 지난 달 미 재무장관은 미국 기업을 상대로 지식을 빼내는(증류 작업을 일삼는) 중국 기업에 제재를 위협했고 미국 기업의 중국산 공개 모델 사용을 제한하는 논의가 진행 중이다.

OFAC
Office of Foreign Assets Control의 줄임말로, 미국 재무부 산하 외국자산통제국. 자산 동결, 거래 금지, 자금 이동 차단 등을 통해 제재 대상의 금융·상업 활동을 제한하는 역할을 한다.

사실 이 대목에서 "발언자가 재무장관"이라는 사실은 결코 가볍게 넘길 일이 아니다. 중국 AI 기업의 증류 행위가 미국의 지식재산권 절취에 해당한다면, 이 문제는 기술 규제나 반도체 수출통제에서 끝나지 않을 수 있다. 재무부 산하 OFAC가 관할하는 자산 동결과 거래 제한까지 정책 수단에 들어올 수 있기 때문이다.

참고로, 현재 진행 중인, 이란을 겨냥한 미 재무부의 "Economic Fury" 작전은 이들이 어떤 방식으로 경제적 압박을 가하는지를 보여 준다. 재무부는 제재 대상만 지정하는 데 그치지 않고, 자금이 이동하는 은행과 환전소, 디지털자산 거래소, 선박과 해외 중개회사까지 따라간다. 핵심은 상품(원유 및 토큰)을 막는 것이 아니라 그 상품이 돈을 벌고 결제하는 통로를 싹 차단하는 것이다.

물론 중국 AI 기업에 이란과 같은 포괄적 제재가 적용된다는 뜻은 아니다. 현재 확인할 수 있는 것은 금융 제재 가능성도 정책 메뉴에 올랐다는 사실까지다.

2. 규제는 이미 다른 형태로 실행되는 중

앞의 세 갈래는 모두 법이나 제도의 이야기다. 그런데 실제로 지금 작동하고 있는 규제는 다른 모양이다.

앞서 언급했듯 지난 6월 Anthropic의 최상위 모델 두 개가 美 정부의 수출통제 지침으로 약 18일간 중단됐다가 다시 열린바 있다. 중요한 것은 다시 열린 뒤의 現모습이다. 지금 이 모델은 요청 내용에 따라 실시간으로 "판정"을 받아, 일부 요청은 자동으로 한 단계 아래 지능의 모델로 넘어간다. 법이 통과된 적은 없는데 규제는 이미 실행되고 있다. 국회를 거치지 않고, 회사가 자기 모델 안에 심어 놓은 판정 장치로 구현됐기 때문이다.

이때 투자자로서 챙겨야 할 포인트는 둘이다. 하나, 규제 위험은 법안이 통과되는 시점이 아니라 그 앞에서 이미 실적에 영향을 준다. 둘, 이 판정이 빡빡할수록 불편을 감수하기 싫은 사용자들은 결국 다른 모델로 옮겨 간다. 안전 장치가 곧 '전환 비용'이 되는 구조다.

3. 규제는 순서대로 오지 않는다

그런데 지난 12일에는 미 재무장관이 자국 기업의 오픈소스 모델 출시를 미국 혁신의 또 다른 승리라고 불렀다. 재무장관이 한낱 AI 모델 출시 하나에 트윗을 한다는 것부터가 신기한 일이다.

그림 16. 메타의 오픈소스 모델 발표에 대한 미 재무장관의 답글
오픈소스 모델이 규제해야 할 위험에서 '지켜야 할 국가 자산'으로 옮겨 가고 있음을 시사

자료: Scott Bessent, 미래에셋증권 리서치센터

그러나 그가 쓴 표현은 무게감이 있다. 개방형과 폐쇄형 가중치 모델 모두를 발전시키는 것이 미국의 리더십을 유지하는 길이라는 것이다. 이는 앞서 본 세 갈래의 규제 중 '국가별 규제'가 어느 방향으로 갈지를 보여 주는 시그널로 해석한다. **공개 모델이 규제해야 할 위험에서 지켜야 할 국가 자산으로 옮겨 가고 있는 것이다.**

표 6. 어떤 규제가 먼저 법이 되느냐가 중요한데, 현재로선 국가별 사용 제한이 유력

쉽게 말하면	상대적으로 유리	상대적으로 불리	이유
사고 책임·강제 중단 규제 (통제를 벗어난 AI를 멈추고 사고 책임을 묻는다)	공개 모델	폐쇄형 모델	모델을 직접 운영하는 프론티어 회사에 책임이 집중
성능 기준·개발 속도 규제 (일정 수준 이상 똑똑한 모델은 공개 여부와 관계없이 규제한다)	폐쇄형 모델	공개 모델	공개 모델도 규제 대상, 규제 비용을 감당할 대형 사업자가 유리
국가별 사용 제한 (미국 기업의 중국산 모델 사용을 제한한다)	**서방권 공개 모델**	**중국 공개 모델**	**기업 수요가 중국산에서 미국·유럽산 공개 모델로 이동**

자료: 미래에셋증권 리서치센터

세 가지 규제 중 무엇이 먼저 구속력을 갖느냐가 진영별 유불리를 가른다. 현재 가장 앞서 있는 것은 국가별 규제다. 그런데 최근 성능 기준 규제도 함께 움직이기 시작했다.

Anthropic의 CEO 다리오 아모데이는 며칠 전 X에서 열린 '댓글 토론' 논쟁에서, 현 행정부가 검토 중인 정책 방식을 지지한다고 밝혔다. 프론티어 모델은 배포 전에 시험을 받고, 공개 모델도 프론티어에 가까워지면 같은 시험을 받는 방식이다.

이 방식이 실제로 자리 잡으면, 공개 모델도 일정 성능을 넘으면 같은 절차를 밟아야 한다. 이는 규제 비용을 감당할 수 있는 대형 사업자가 상대적으로 유리해진다는 것을 의미한다.

다만 반대 방향의 힘도 세다. 유명 투자자 개빈 베이커는 앞서 본 허깅페이스 사고에서 프론티어 모델이 일으킨 문제를 결국 공개형 모델이 해결했다는 사실이 강한 규제 논의의 동력을 크게 꺾었다고 보고 있다. 베이커의 우군이자 '백악관 AI 정책 책임자' 데이비드 색스도 사전 승인 방식이 대기 줄을 만들어 미국을 늦출 것이라며 공개적으로 반대하고 있다.

결론적으로, 국가별 규제가 먼저 자리를 잡고 있고, 성능 기준 규제는 법이 아니라 배포 전 시험이라는 형태로 그 옆에서 이미 함께 굴러가고 있다.

4. 1부를 마무리하며…

이번 편의 결론은 '최고의 모델이 하나로 정해지지 않는다'는 큰 전제 위에 있다. 그 전제가 깨지면 결론도 깨진다. 그러나 현재로서는 이 시나리오의 가능성은 낮아 보인다. 두 달 사이에 AI 모델의 우위가 계속 뒤집히고 있고, 추가 훈련을 전혀 하지 않은 공개 모델이 실제 업무에서 프론티어를 이긴 사례가 보고됐으며, 세계 최대 클라우드의 최고경영자가 "최고 모델이 여섯 개는 될 것"이라고 공개적으로 말했다.

따라서 **지속적으로 가치가 불어나는 곳은 특정 모델이 아니라, 데이터와 선택권을 통제하는 자리일 것으로 보인다.** 그 자리는 기업 업무 안에서 현장 지식을 기계가 읽을 수 있게 만드는 것과 여러 모델을 한자리에 모아 놓는 것이 핵심이다.

그리고 이는 필연적으로 "엔터프라이즈 AI와 추론 컴퓨팅의 캄브리아기 대폭발"과 궤를 함께 한다. 오픈소스를 가장 강하게 옹호하는 마크 저커버그조차 "컴퓨트만은 크게 지어야 한다"고 말한 것처럼 말이다. 제본스의 역설이 일어날까 일어나지 않을까 고심해야 하는 상황이 아니라, "일어날 것이고 그래야만 한다"는 것이 AI guru들의 생각이라는 것을 유념하자.

실제로 AI를 쓰는 값은 크게 떨어졌는데, 그 AI를 돌리는 GPU의 임대료는 같은 기간 오히려 올랐다. 오픈소스를 가장 강하게 옹호하는 마크 저커버그조차 컴퓨트만은 크게 지어야 한다고 썼고, 반대편의 연구자들도 정작 희소한 것은 모델이 아니라 컴퓨트라고 말한다.

두 값이 왜 반대로 움직이는지, 그 결과로 반도체와 전기와 건물이라는 물건이 어떻게 담보로 잡혀 금융자산이 되어 가는지, 그리고 그 과정에서 돈은 누가 걷는지를 「은하지능전설 2부」에서 다룬다.

Compliance Notice

- 당사는 자료 작성일 현재 조사분석 대상법인과 관련하여 특별한 이해관계가 없음을 확인합니다.
- 당사는 본 자료를 제3자에게 사전 제공한 사실이 없습니다.
- 본 자료를 작성한 애널리스트는 자료작성일 현재 조사분석 대상법인의 금융투자상품 및 권리를 보유하고 있지 않습니다.
- 본 자료는 외부의 부당한 압력이나 간섭없이 애널리스트의 의견이 정확하게 반영되었음을 확인합니다.