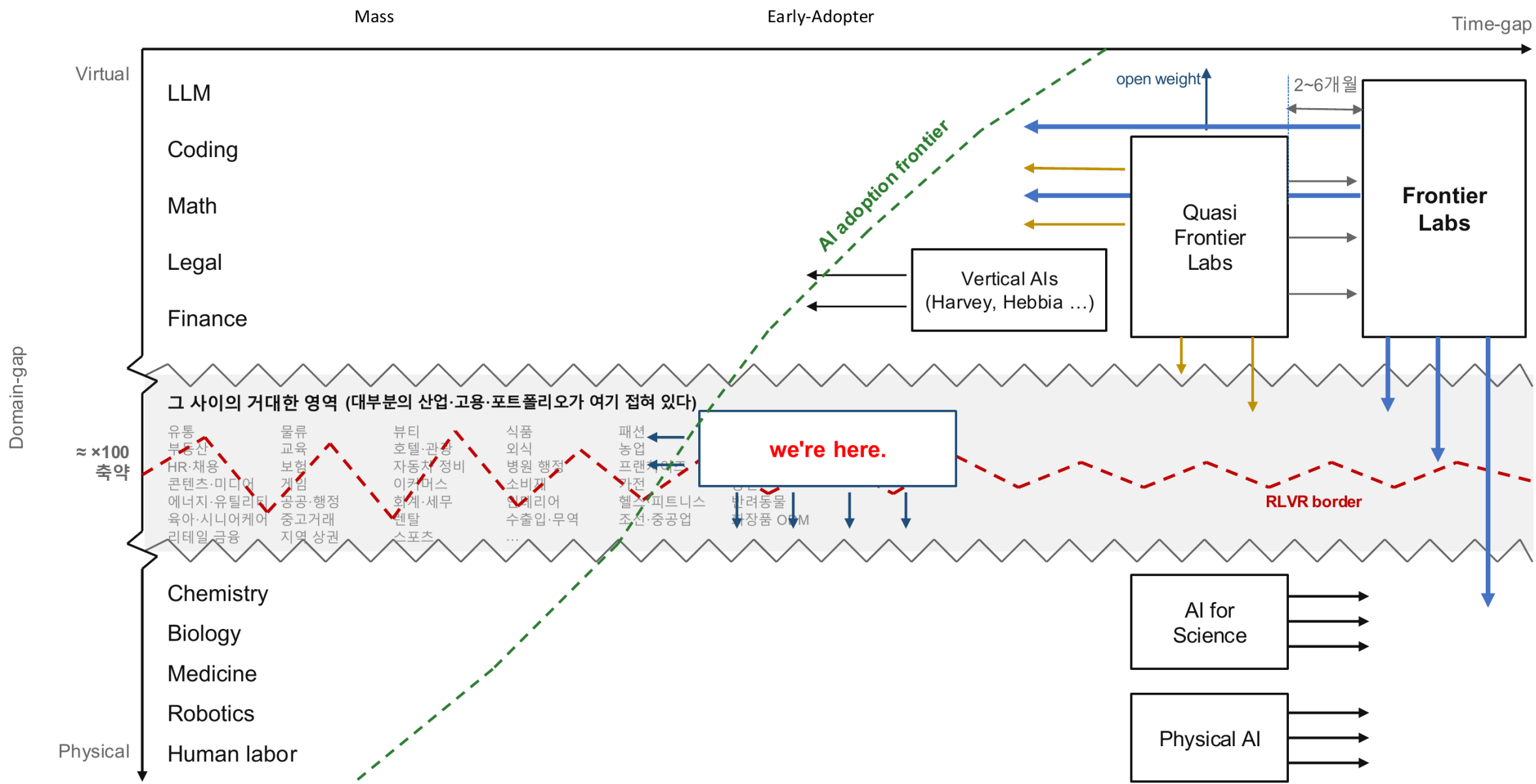


AI Frontier Korea · 2026년 10월 11일

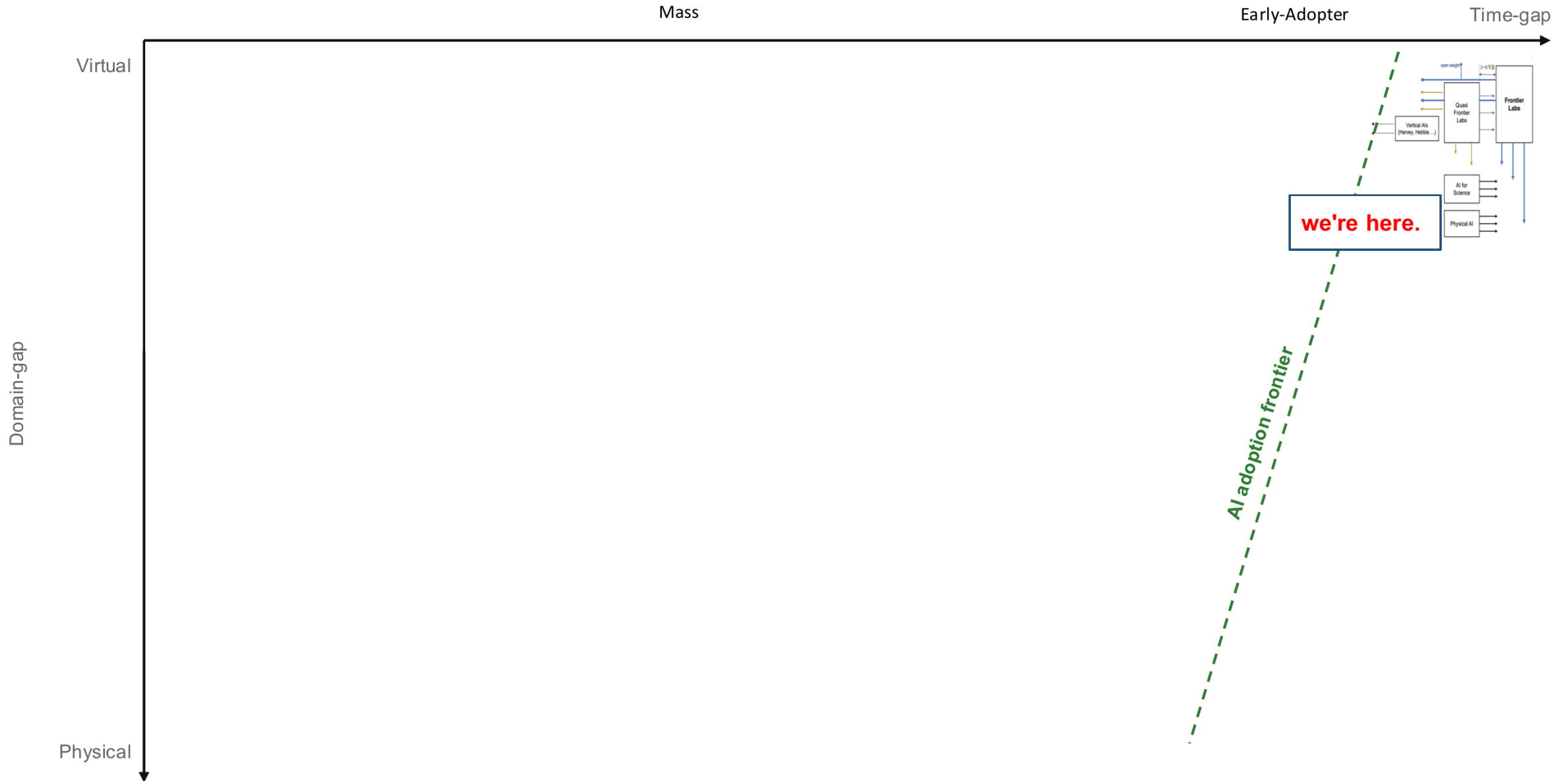
AI Frontier 생각실험

모델은 누구나 만들고, 독점은 없고, 소비자시장(consumer market)은 아직 열리지 않았는데
왜 모두가 수천억 달러를 까는가 — 과거의 산업에서 배울 수 있는 교훈은?

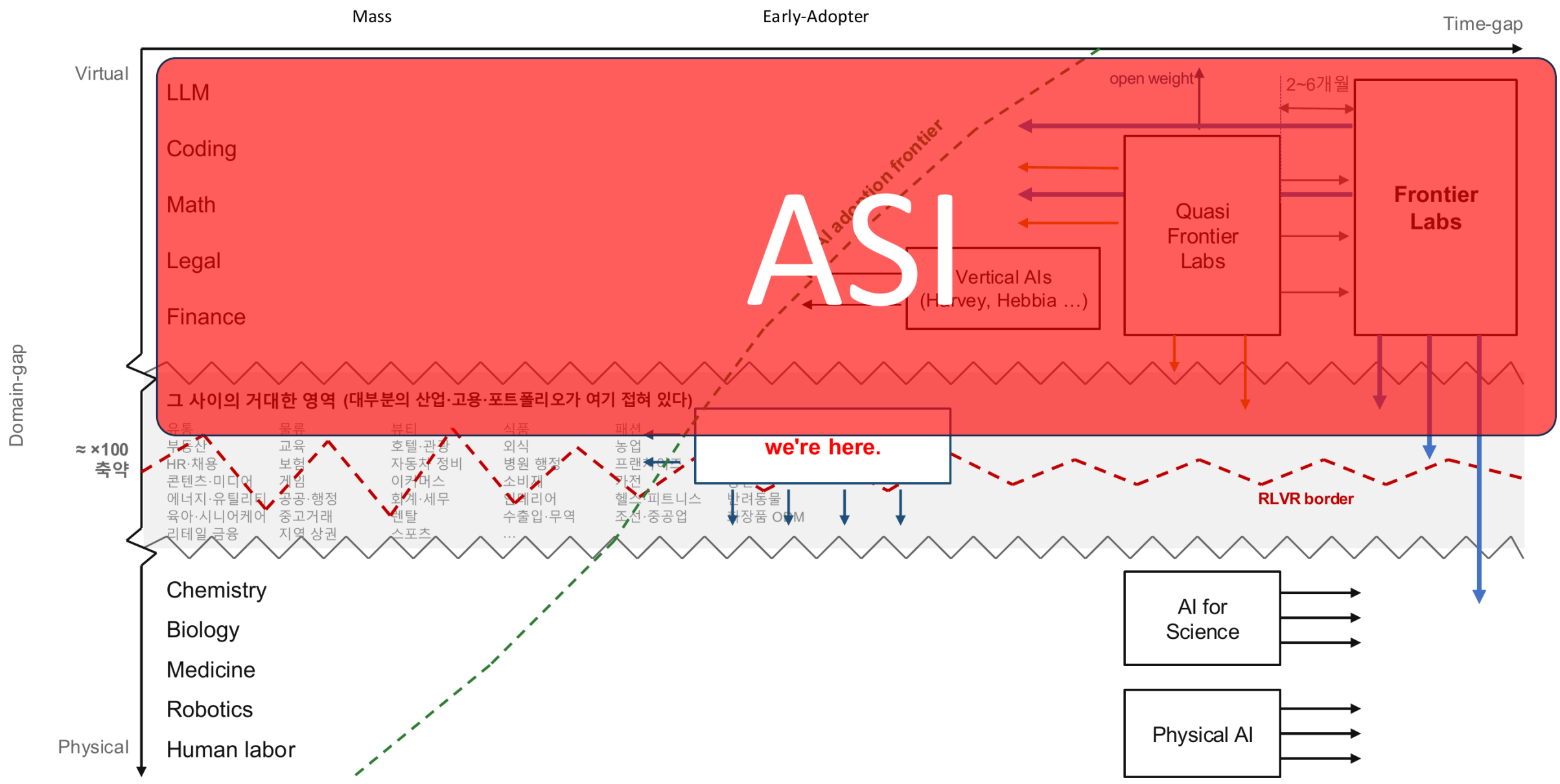
우리의 위치 : time-gap × domain-gap — 그 사이에 접혀 있는 거대한 기회



우리의 위치 : time-gap × domain-gap — 그 사이에 접혀 있는 거대한 기회



AGI 2026, ASI 2028 예측하던데 $\pi.\pi$



도망자 플레이북



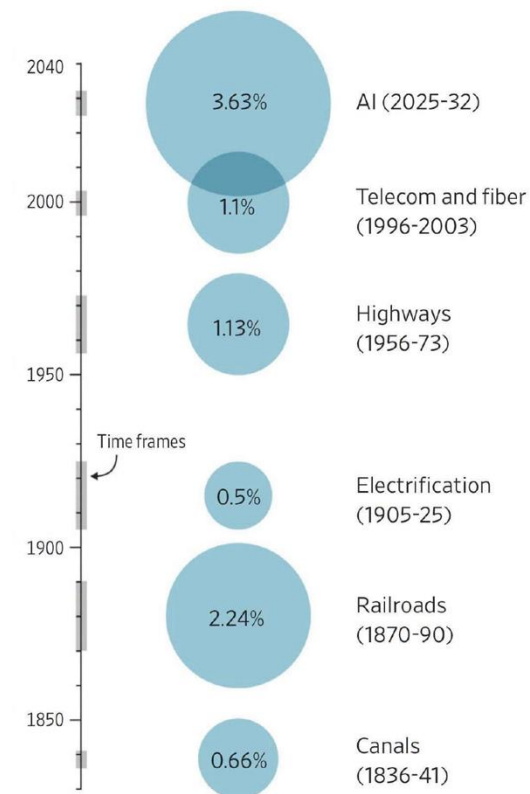
모델은 누구나 만들 수 있게 됐다.
독점은 일어나지 않는다.
소비자시장은 아직 열리지 않았다.

**그런데 왜 모두가 수천억 달러를 깔고 있는가. 그
리고 그 베팅은 회수될 것인가.**

2025-2032년 AI 인프라 투자는 연평균 GDP의 3.63%로 추정된다. 철도(2.24%)와 광통신(1.1%)을 넘는다.

출처: Stijn van Nieuwerburgh(Brookings)

Average annual infrastructure spending as a percentage of GDP



Note: AI spending is estimated.
Source: Stijn van Nieuwerburgh

깎리는 돈과 버는 돈

깎리는 돈

약 7,250억 달러

2026년 빅테크 4사 capex. 아마존 2,200억, 알파벳 1,950–2,050억, 마이크로소프트 약 1,750억, 메타 1,300–1,450억. 2025년 약 4,100억에서 77% 늘었다.

칩이 버는 돈

890억 달러 / 분기

엔비디아 데이터센터 매출(2026년 5–7월). 하이퍼스케일러 약 490억, AI 클라우드·산업·기업 약 400억. 매출총이익률 75%. 다음 분기 전체 매출 가이드스는 1,080억.

AI를 팔아 버는 돈

약 1,750억 달러 / 년

생성형 AI 산업 전체의 연환산 매출(중복 제거, 2026년 6월). 그중 OpenAI 400억 이상, Anthropic 약 650억(7월 말 보도). 나머지 모델 회사는 각각 연 10억 미만.

엔비디아의 두 분기 매출(962억 + 가이드스 1,080억)이 AI를 팔아 버는 산업 전체의 1년치와 비슷하다. 깎리는 돈은 버는 돈의 약 4배다. JPMorgan은 이 투자에서 적정 수익을 내려면 연 6,500억 달러의 추가 매출이 필요하다고 본다.

네 가지 관찰 ①

관찰 1

프런티어의 공식은 더 이상 독점 지식이 아니다

- 오픈웨이트는 프런티어를 평균 4개월 뒤에서 따라온다.
- 미국과 중국 최고 모델의 차이는 Arena 점수 2.7%다.
- 대신 자본 경쟁은 사상 최대다. Anthropic은 5월 650억 달러를 조달했고, OpenAI는 2030년까지 약 6,000억 달러의 컴퓨트 지출을 목표로 한다.

→ 기술이 해자가 아니라면, 무엇이 승자를 가르는가?

관찰 2

독점이 아니라 다원화가 일어난다

- 프런티어 조직은 6곳. 상위 4개 랭(Anthropic, xAI, Google, OpenAI)이 Arena 25점 안에 있다.
- GPT-4급 모델은 33개, 개발사는 12곳이다(2025년 6월).
- 매출 순위도 뒤집혔다. 보고 기준으로 Anthropic이 OpenAI를 넘었다.

→ 왜 독점이 안 생기고, 다원화는 언제까지 가는가?

네 가지 관찰 ②

관찰 3

소비자시장은 아직 열리지 않았고, 소수의 얼리어답터가 시장을 끈다

- 미국인 절반이 쓰지만 구독 결제자는 약 4.5%다. 그 안에서 상위 10%가 매출의 절반 이상을 낸다.
- 돈이 되는 건 엔터프라이즈 성격의 일이고, 그 중심은 코딩이다. Cursor 연환산 30억 달러, Claude Code 연 25억 달러.
- 미국 기업의 19.8%만 AI를 쓰고, 그중 66%는 사람을 보조하는 데만 쓴다.

→ 소비자시장은 왜 아직 열리지 않았고, 언제 열리는가?

관찰 4

소비자시장이 열리면 수요가 폭발한다는 낙관론 위에서 치킨게임이 벌어진다

- 빅테크 4사의 2026년 capex는 약 7,250억 달러, 2027년엔 1조 달러를 넘을 전망이다. 산업 전체 매출은 연 약 1,750억 달러다.
- 점점 빌린 돈으로 깎는다. SPV로 장부 밖에 뺀 부채가 1,200억 달러를 넘고, 앞으로 5년 부채 수요는 약 4조 달러다.
- 연준은 9월 금리를 올렸다. 그래도 멈추는 쪽이 프런티어 용량을 잃는다.

→ 수요는 이 베팅을 따라잡을 것인가?

기술이 공개되자, 경쟁은 자본 경주가 됐다

- **영국 철도:** 1829년 레인힐 트라이얼에서 증기기관차가 공개적으로 증명된 뒤, 기술은 곧 모두의 것이 됐다. 1840년대에 희소한 건 기관차가 아니라 노선 인가, 자본 모집 속도, 타이밍이었다. 그 경주의 챔피언 "철도왕" 조지 허드슨은 1849년 자본금에서 배당을 주던 게 드러나 무너졌다.
- **미국 철도:** 자본 조달의 챔피언은 남북전쟁 국채를 팔던 제이 쿡이었다. 노던 퍼시픽 채권을 일반 투자자에게 팔다가 1873년 파산했다.
- **광통신:** DWDM 장비는 누구나 살 수 있었다. 경쟁은 누가 더 빨리 돈을 모아 더 많이 까느냐였다. Global Crossing은 2002년 자산 255억 달러를 들고 파산했다.

끝까지 남은 쪽은 자기 현금흐름과 고객을 가진 쪽, 그리고 붕괴 뒤 헐값에 사들인 쪽이었다.

AI에서 프런티어는 유통기한 4개월짜리 재고다

- 공식은 공개재가 됐지만 프런티어는 공짜가 아니다. 오픈웨이트 격차는 2023–2025년 3개월에서 2026년 4개월로 오히려 벌어졌다. 닫힌 랩들이 연산으로 거리를 벌렸다.
- 그 4개월의 프리미엄은 크다. Anthropic의 추론 마진은 1년 새 38%에서 약 70%로 올랐다는 추정이 있다.
- 그래서 마케팅 타이밍, 유통력, 자본조달 속도가 중요해졌다. 상하기 전에 팔아야 하고, 다음 재고를 만들 돈을 그 사이에 모아야 한다.
- 자기 현금흐름이 있는 쪽(구글, 메타, 마이크로소프트)과 외부 자본에 기대는 쪽(OpenAI, Anthropic, xAI)을 나눠서 봐야 한다.

기술이 해자가 아닌 시장에서는 돈을 빨리 모으는 쪽이 앞서간다. 끝까지 남는 건 돈이 마를 때도 버틸 수 있는 쪽이었다.

돈이 풍부한 동안엔 모두가 지었고, 통합은 붕괴 다음에 왔다

- **영국 1844-46년:** 정부채 금리가 3.25%일 때 선두 철도회사는 10%를 배당했다. 1846년 한 해에 법안 272개, 9,500마일이 인가됐다. 분할납입 주식이라 10%만 내고 들어올 수 있었다. "안 지으면 남이 가져간다."
- **영국 1847년:** 철도 투자가 GDP의 6.7%에 달한 해, 흥작으로 금리가 오르고 추가 납입 요구가 몰리며 무너졌다. 투자자는 잃었지만 노선은 남았다.
- **미국:** 마일당 보조금으로 수요보다 앞서 지었다. 경쟁 노선이 여럿인 간선에서 운임 전쟁이 붙었고, 담합은 몰래 운임을 깎는 쪽 때문에 계속 깨졌다. 독점은 경쟁 노선이 없는 지선에만 있었다.
- **통합:** 1893년 공황 뒤 철도 투자는 GDP의 0.3%로 추락했고, JP모건식 구조조정으로 소수 대형 시스템이 됐다. 광통신도 2001-2003년 60곳 이상 파산한 뒤 살아남은 쪽이 헐값에 자산을 가져갔다.

AI의 다원화는 자본이 풍부한 단계의 모습이다

- 모델만으로는 독점이 안 된다. 가중치는 4개월이면 따라잡히고, 모델은 API 한 줄로 갈아탈 수 있다. 철도는 지리적 독점이라도 있었다.
- 독점이 생긴다면 모델 밖, 경쟁 노선이 없는 지선에서 생긴다. 사용자 접점(ChatGPT 주간 사용자 12억 명), 쌓인 맥락, 업무 흐름.
- 통합의 초기 신호는 이미 보인다. SpaceX와 xAI 합병(2월), SpaceX의 Cursor 인수 발표(6월), 금리 인상(9월).
- 모든 층이 다른 층으로 확장한다. 클라우드는 칩으로(TPU, Trainium), 랩은 앱으로(Claude Code, ChatGPT 광고), 앱은 모델로(Cursor Composer 2). 아무도 자기 층만으로는 안전하다고 보지 않는다.

다원화는 승자가 없다는 뜻이 아니라, 아직 돈이 마르지 않았다는 뜻이다.

3장 · 소비자시장은 왜 아직 열리지 않았고, 언제 열리는가?

소비자시장은 조금은 늦게 열렸다

- **첫 고객은 상인이었다.** 리버풀-맨체스터 철도(1830)는 원면 화물용이었다. 운하로 20시간이던 화물이 2시간에 갔고, 예상 밖 승객이 하루 500명에서 1,600명으로 늘었다.
- **전기는 40년이 걸렸다.** 발전소는 1881년에 생겼지만, 미국 공장 동력 중 전기모터 비중은 1899년 5% → 1919년 55% → 1929년 82%였다. 처음엔 증기기관 자리에 모터만 바꿔 달았고, 생산성은 공장을 다시 설계한 뒤에야 올랐다.
- **철도 위의 소비자시장은 제도 다음에 열렸다.** 시어스의 통신판매는 농촌 무료 배달(1896)과 소포우편(1913)이 깔린 뒤에 폭발했다.
- **인터넷 소비자시장은 대역폭이 헐값이 된 뒤에 열렸다.** 유튜브는 2005년, 넷플릭스 스트리밍은 2007년이다.

AI는 아직 증기기관 자리에 모터를 단 단계다

- **사용은 왔지만 지불은 안 왔다.** 대부분의 AI는 시간을 아끼는 도구다. 사람들이 돈과 시간을 쓰는 건 시간을 보내는 제품(소셜, 엔터테인먼트, 쇼핑)이다.
- **원가가 소비자시장의 사업모델을 막는다.** 파워유저 한 명 서빙에 월 수백~수천 달러. 웹2처럼 먼저 키우고 광고로 버는 전략이 아직 안 된다.
- **얼리어답터는 레버리지에 돈을 낸다.** 리버풀의 상인들처럼, 자기 일에 바로 돈이 되는 사람들이다. 코딩이 먼저 열린 건 테스트로 결과를 검증할 수 있어서다.
- **소비자시장이 열리는 조건:** (1) 단가가 무너져 광고·거래 모델이 작동한다. OpenAI 광고 연환산 10억 달러가 첫 신호다. (2) 에이전트를 믿고 맡길 제도가 생긴다. "에이전트는 정말 소비자 편인가." (3) 시간을 보내는 AI 제품이 나온다.

소비자시장은 공장을 다시 설계하고, 소포우편 같은 제도가 깔린 다음에 열린다.

광통신은 수요가 매년 두 배로 늘었는데도 망했다

- "인터넷 트래픽이 100일마다 두 배"는 1995–96년 2년간만 사실이었다. 그 뒤로도 트래픽은 매년 두 배로 늘었다.
- 그런데 광섬유 한 가닥에 여러 파장을 싣는 DWDM으로 공급이 더 빨리 늘었다. 2002년 깔린 광섬유의 2–5%만 쓰였고, 코닝의 2002년 매출은 예상 100억 달러 대신 40억 달러였다.
- 영국 철도도 같았다. 의회에 낸 계획대로라면 매출이 1844년의 열 배가 돼야 했다. 산업 매출은 그 뒤에도 늘었지만, 마지막에 지은 노선은 회수되지 않았다.

방향은 늘 맞았고, 기울기는 늘 과대평가됐고, 과실은 늘 다른 사람이 가져갔다.

지금은 모른다. 신호가 서로 다른 방향을 가리킨다

따라잡는다는 쪽

- 토큰 수요는 연 약 10배, 추론 공급 능력은 연 약 3.4배 는다 (Epoch). 연산 부족이 곧, 어쩌면 이미 왔다.
- GPU 1년 예약 가격은 반년 새 약 40% 올랐다. 엔비디아는 내년 약 70% 성장도 공급이 막은 숫자라고 말한다.

못 따라잡는다는 쪽

- 깔리는 돈이 버는 돈의 약 4배다. 토큰이 10배 늘어도 매출은 10배가 아니다. 같은 성능의 가격이 매년 5–10배 떨어진다.
- 빛을 갹는 건 토큰이 아니라 매출이다. GPU는 3–5년이면 감가 되고 금리는 오르고 있다.

지켜볼 변수는 하나다. AI로 버는 돈이, 깔아놓은 설비가 낡아가는 속도와 빌린 돈의 이자를 감당할 만큼 빨리 느는가. 감당하면 과잉 투자한 일부만 정리된다(1847년 영국). 못 하면 업계 전체가 법정관리와 통합을 거친다(1893년 미국).

결론

네 개의 답

1. 기술이 범용이 되면 승부는 자본과 유통으로 간다. 하지만 자본 경주의 승자가 최종 승자인 경우는 드물었다.
2. 다원화는 자본이 풍부한 동안의 현상이다. 통합은 돈이 마를 때 온다.
3. 소비자시장은 단가가 무너지고 제도가 깔린 다음에 열린다. 지금은 코딩과 얼리어답터만 먼저 도착했다.
4. 수요가 따라잡을지는 아직 모른다. 유료 매출이 감가상각과 이자를 넘는지 보면 된다.

그러면 가치는 싸진 지능 위에서 맥락과 신뢰를 쥔 쪽으로 간다.

가치는 희소한 곳에 쌓이고, 희소성은 위로 이동한다

칩: 지금이 가장 많이 버는 때

데이터센터 매출 분기 890억 달러, 마진 75%. 하지만 ASIC 출하가 GPU보다 세 배 빨리 늘고 마진 저점 가이드는 71–72%다. 비중은 정점 근처.

모델: 두 시장으로 쪼개진다

후행 역량은 가격이 매년 5–10배 떨어지는 운임 전쟁. 프런티어는 4개월짜리 재고를 빨리 파는 소수 과점. 랩은 하네스, 광고, 칩으로 넓힌다.

오케스트레이션: 맥락과 책임이 쌓이는 곳

지금은 가장 못 번다. 그런데 Cursor는 자체 모델 없이 연 20억 달러를 만들었다. 철도정산소와 소포우편의 자리다.

앱: 코딩만 먼저 도착했다

앱이 크면 랩이 먹는다(Midjourney). 그래도 단가가 무너지고 제도가 깔리면 가장 큰 몫은 여기로 온다. 시어스의 자리다.

그래서 우리는 무엇을 하고, 무엇을 지켜보나

플레이북은 네 답에서 나온다

1. 프런티어의 가치를 빨리 알아챈다. 4개월이면 범용이 된다. (1장)
2. 내 도메인이 안전한지 판단한다. 랩은 위로 올라온다. (2장)
3. 조금 늦게 오는 사람들을 대상으로 사업한다. 소비자시장은 아직이다. (3장)
4. 도메인 프런티어로 승천한다. 모델이 아니라 검증 기준과 맥락을 짚는다. (3·4장)

지켜볼 신호

- GPU 1년 예약 가격. 지금은 상승, 꺾이면 공급이 따라잡은 것.
- 유료 매출 증가 대 감가상각과 이자.
- 오픈웨이트와 프런티어의 격차. 지금 평균 4개월.
- 코딩 밖에서 결과를 검증하고 책임을 추적하는 제도.
- 앱 매출 성장률 대 랩 매출 성장률.

출처

역사

[Liverpool and Manchester Railway](#) · [The Peel Web](#)
[Railway Mania – Narrative](#) · [Victorian Web](#)
[Odlyzko, Collective Hallucinations](#) · [The Railway Mania](#)
[Cornell – Northern Pacific](#) [채권 판촉물](#)
[Light Reading – WorldCom](#) · [FDIC 2002](#)
[TechFinitive – Global Crossing](#) · [Laser Focus World](#)
[HSUS Table Dd848–853](#) · [Paul David\(HNN\)](#)
[Advisor Perspectives](#) · [36Kr](#)

지금

[Epoch – Compute crunch](#) · [Epoch – 오픈웨이트 격차](#)
[Epoch – 1e25 FLOP 모델](#) · [Stanford AI Index 2026 요약](#)
[Epoch – 매출](#) · [The Price of Progress](#) · [Wing VC](#)
[NVIDIA FY27 2분기](#) · [Tom's Hardware – ASIC](#) · [Spheron](#)
[U.S. Census BTOS](#) · [Cursor 통계](#) · [Beam](#)
[Crypto Briefing](#) · [Techloy](#) · [Icosa](#)
[FT 인용 – SPV 부채](#) · [Tom Tunguz](#) · [FOMC 2026.9](#)
[a16z Top 100 Consumer AI Apps 7판](#) · [패트릭 콜리슨 글](#) · [Anthropic Economic Index](#)