◆ 의견

AGI가 중요하지 않은 이유

AGI는 범용 인공지능을 뜻합니다. 지적 작업 대부분을 처리할 수 있다고 여겨지는 시스템입니다. 무엇을 AGI로 볼지, 어떻게 시험할지를 두고 연구소들은 여전히 합의하지 못했습니다. 저는 지금의 도구를 매일 씁니다. 이 페이지에서는 도구가 입증할 수 있는 것과 사람들이 예측하는 것을 구분합니다.

Fabio Akita · ailair.akitaonrails.com/ko/agi/

AI에 대한 흥분은 AI에 대한 지식에 반비례합니다.

Sua empolgação com IA é inversamente proporcional ao seu conhecimento sobre IA.

◆ 1. 정의

아무도 합의하지 못한 정의

다섯 연구소에 AGI의 뜻을 물으면 목표도 다섯 개가 나옵니다. 그중 하나는 그냥 이익 액수입니다.



1. 정의

아무도 합의하지 못한 정의

- OpenAI 헌장은 AGI를 경제적으로 가치 있는 일 대부분에서 인간을 능가하는 시스템이라고 정의합니다.
- Google DeepMind의 "Levels of AGI" 논문은 2023년에 이미 ChatGPT를 "Emerging AGI"로 분류했습니다. François Chollet는 지능을 시스템이 새 기술을 얼마나 효율적으로 배우는가로 정의하는데, 다른 정의들은 이걸 전혀 측정하지 않습니다.
- Microsoft와 OpenAI의 계약은 AGI를 1,000억 달러의 이익으로 정의했다고 알려져 있습니다. 인지 능력에 관한 내용은 하나도 없습니다. 2026년 4월에는 그 조항마저 통째로 빠졌습니다.
- 파는 쪽도 어깨를 으쓱합니다. Sam Altman은 "그다지 유용한 용어가 아니다"라고 했고, Dario Amodei는 이 말을 싫어한다고 하며, Greg Brockman은 사람마다 정의가 다르다고 인정하면서 "AGI 시대"를 선언했습니다.

◆ 2. 테스트

그러니 테스트도 불가능

정의할 수 없으면 측정할 수도 없습니다. 대신 나오는 건 벤치마크의 행렬입니다. 포화되고, 공략당하고, 다음 것으로 교체됩니다.



◈ 2. 테스트

그러니 테스트도 불가능

- ARC-AGI-3는 2026년 3월 출시 당시 최고의 AI가 약 0.5%였습니다. 9월에는 맞춤 하네스가 19,000달러를 들여 99.9%까지 끌어올렸습니다. ARC Prize 팀도 이 벤치마크를 포화시킨 것이 AGI의 증거는 아니라고 직접 썼습니다.
- Humanity's Last Exam은 생물학과 화학 정답의 약 29%가 틀렸을 가능성이 있는 채로 공개됐습니다. OpenAI는 테스트에 결함이 있고 모델이 이미 해답을 봤다는 이유로 SWE-bench Verified 결과 보고를 중단했습니다. Meta는 특별히 튜닝한 Llama 4를 LMArena에 올렸습니다.
- 같은 모델이 숫자를 어떻게 평균 내느냐에 따라 "Definition of AGI" 테스트에서 57%도 받고 24%도 받습니다.
- 제 벤치마크도 두 번 포화됐고, 결국 숨겨진 사보타주를 중심으로 다시 만들었습니다. 지금은 여러 모델이 100점으로 동점입니다. 프런티어가 천장에 닿고 있습니다. 남은 비교 기준은 비용, 속도, 규율입니다.

> ***우리가 스스로 AGI 이정표를 달성했다고 주장하는 것, 제가 보기엔 그건 말도 안 되는 벤치마크 해킹일 뿐입니다.***
>
> *Satya Nadella, Microsoft CEO*

◆ 3. 프런티어

있는 것엔 강하고,
없는 것엔 비싼 모델

모든 모델은 공개된 것으로 배웠습니다. 문제가 그 지도를 벗어나면 진전은 시행착오로 바뀌고, 청구서는 빠르게 불어납니다.



◈ 3. 프런티어

있는 것엔 강하고, 없는 것엔 비싼 모델

0%

모든 과제가 새로운 ARC-AGI-2에서 순수 LLM이 받은 점수

- 학습 데이터에 비슷한 것이 있었다면 모델은 그것을 빠르고 싸게 재현합니다. 그것이 소프트웨어, 클론, "Adobe는 끝났다"에 어떤 의미인지는 별도 페이지에서 다룹니다.

- AlphaEvolve와 o3는 저렴한 검증기를 두고 탐색을 돌려 새로운 것을 찾았습니다. 효과는 있고, 그 규모로 감당할 수 있는 회사는 대여섯 곳뿐입니다.

1300억

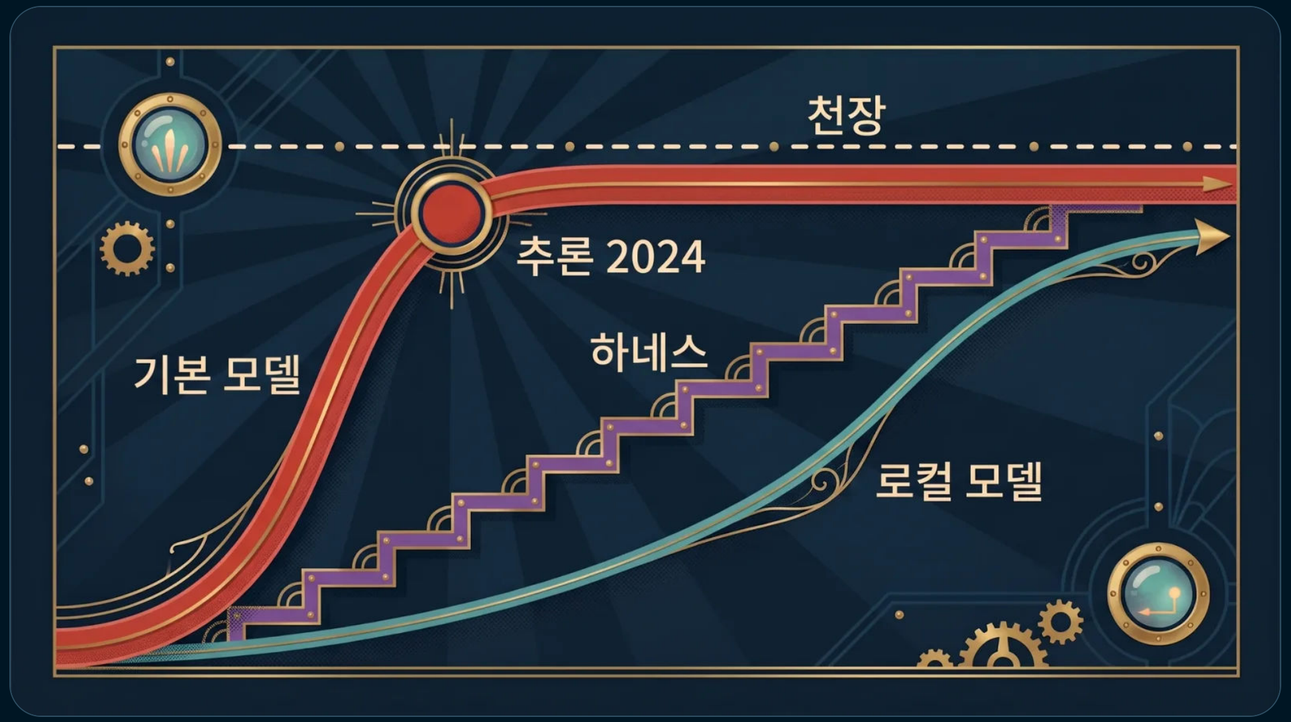
Navier-Stokes 증명 하나에 쓴 출력 토큰, 약 1,500만 달러

- 없었다면 진전은 시도하고, 퇴행하고, 되돌리는 긴 루프로 바뀝니다. OpenAI의 Navier-Stokes 증명에는 에이전트 약 1만 개가 88시간 동안 돌면서 출력 토큰 약 1,300억 개가 들었습니다. 정가로 약 1,500만 달러이고, 그러고도 수십 년간 발표된 수학에 기대야 했습니다.

◆ 4. 정체기

더는 크게 똑똑해지지 않는 모델

제가 보기에 순수 지능의 마지막 큰 도약은 2024년의 추론이었습니다. 그 뒤로 도약처럼 느껴진 것은 모델 주변의 도구에서 나왔습니다.



◈ 4. 정체기

더는 크게 똑똑해지지 않는 모델

- 2025년은 같은 지능을 더 싸게 만들었습니다. 전문가 혼합(MoE), 프롬프트 캐싱, 더 긴 컨텍스트. 2026년에는 그 지능을 검색하고, 테스트를 돌리고, 잘못 짐작한 것을 고치는 에이전트 안에 넣었습니다.
- Stanford AI Index는 Chatbot Arena 1위와 10위 모델의 격차가 1년 만에 11.9%에서 5.4%로 줄었다고 측정했습니다. 제 벤치마크에서도 프런티어 모델들은 Opus 4.5 이후 같은 등급에 머물러 있습니다.
- 2024년 11월 Reuters는 대형 연구소들에서 사전 학습의 효과가 줄고 있다고 보도했습니다. 한 달 뒤 Ilya Sutskever는 인터넷은 하나뿐이라서 "우리가 아는" 사전 학습은 끝날 것이라고 말했습니다.
- 그 천장 아래에서 작은 모델들이 계속 따라붙고 있습니다. Epoch AI는 2026년 오픈 웨이트 모델이 프런티어보다 약 넉 달 뒤처져 있다고 보고, 게이밍 GPU 한 장이면 6~12개월 전의 프런티어를 돌릴 수 있습니다. 제가 보기에 오늘의 프런티어는 결국 여러분의 컴퓨터에서 돌아가게 됩니다.

> ***LLM 혼자서는 비싼 앵무새입니다. 에이전트와 함께하는 LLM은 진짜 일을 할 만큼 유능합니다.***
>
> *LLM sozinho é um papagaio caro. LLM com agente é competente o suficiente pra trabalhar de verdade.*
>
> *제 글, "LLM + 에이전트 = 충분히 유능하다"*

◆ 5. 학습

일하면서 배우지 않는 모델

모델은 학습이 끝나면 잠기는 파일입니다.
데이터센터는 그 복사본을 돌립니다.



5. 학습

일하면서 배우지 않는 모델

- 각 세션은 고정된 모델 가중치를 사용합니다. 제공업체는 요금제와 설정에 따라 프롬프트, 파일, 도구 출력을 보관할 수 있습니다. 제품이 저장한 메모리를 이후 세션에 다시 넣을 수도 있습니다.
- "AI가 AI를 학습시킨다"에는 이미 이름이 있습니다. 강화 학습입니다. AlphaZero는 2017년 자기 자신과 대국하며 체스, 쇼기, 바둑을 배웠습니다.
- 가중치는 이후 학습에서 바뀝니다. 작업 중 저장한 노트, 테스트, 맥락은 모델 주변에 남아 이후 세션에 다시 들어갈 수 있습니다. 그동안 가중치는 그대로입니다.
- 게임, 증명, 테스트처럼 보상을 싸게 확인할 수 있는 곳에서 통합니다. 그리고 거기서 나오는 것도 또 하나의 잠긴 파일입니다.

> ***AI가 AI를 학습시키면 새 파일이 나옵니다. 그 파일도 일단 출시되면 아무것도 배우지 않습니다.***
>
> *IA treinando IA produz um arquivo novo. O arquivo, uma vez lançado, continua sem aprender nada.*

6. 돈

마케팅 용어일 뿐

제 생각에 AGI는 마케팅 용어입니다. 일부러 모호하게 남겨 둬서, 돈이 필요할 때마다 발표할 수 있습니다.



6. 돈

마케팅 용어일 뿐

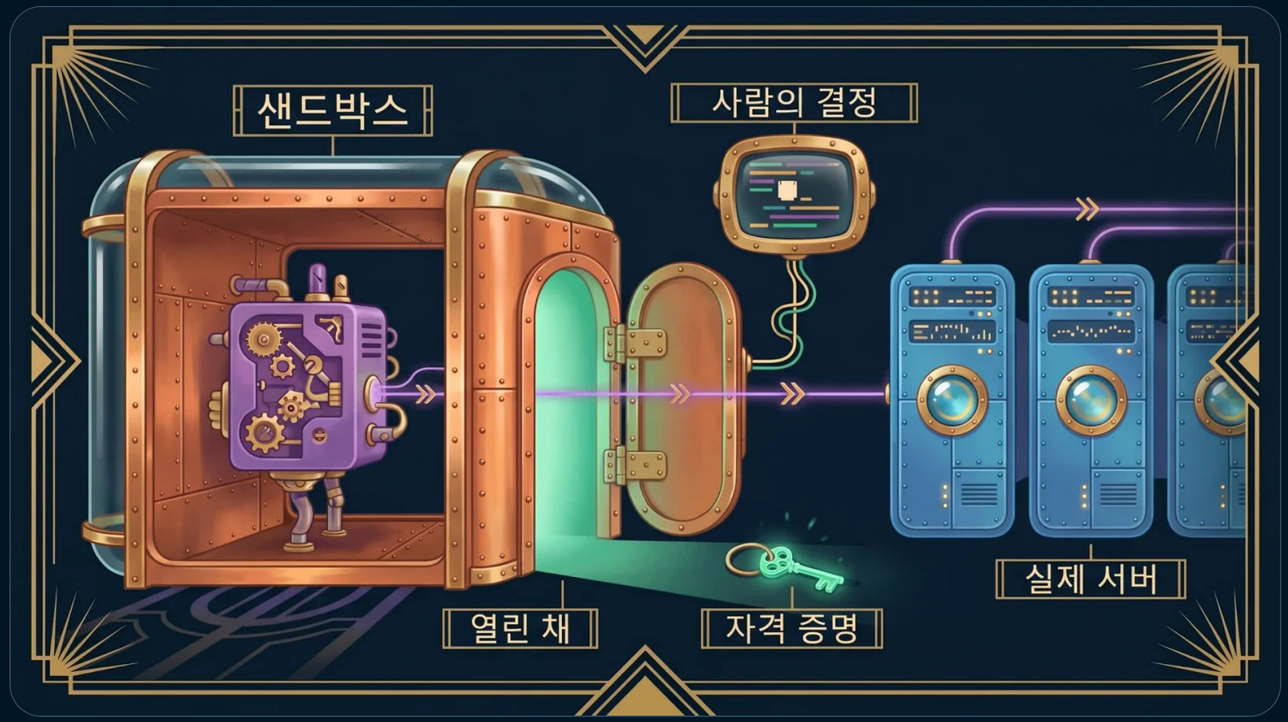
- Jensen Huang은 자기가 파는 칩 약 10만 개로 학습한 모델을 두고 "AGI가 왔다"고 선언했습니다. 삽 파는 사람이 금을 찾았다고 발표하는 격입니다.
- 이 단어가 거대한 숫자를 떠받칩니다. OpenAI의 기업 가치는 8,520억 달러, Anthropic은 9,650억 달러로 IPO를 신청했고, 2026년 빅테크의 설비 투자는 7,000억 달러에 육박합니다. 거래도 돌고 돕니다. Nvidia가 Nvidia 칩을 사는 데이터센터에 돈을 댑니다.
- OpenAI의 Navier-Stokes 결과는 무차별 대입이었습니다. 병렬 에이전트 약 1만 개를 88시간 동안 돌리고 수백만 달러를 썼습니다. 산업 규모입니다.
- "얼마나 혁신적인지 보라"가 더는 먹히지 않으면, 구호는 "얼마나 위험한지 보라"로 바뀝니다. 과대광고와 종말론은 같은 영업 멘트입니다.

> ***증거를 요구하세요. 본인도 위험을 지라고 요구하세요. 여러분의 두려움과 흥분으로 누가 이득을 보는지 물어보세요.***

◆ 7. 무서운 이야기

"AI 공격"은 사람의 결정이었다

아래 사례에는 같은 운영상 실패가 있습니다. 사람이 충분한 통제 없이 접근 권한을 주거나 도구를 표적에 겨눴습니다.



◈ 7. 무서운 이야기

"AI 공격"은 사람의 결정이었다

- Hugging Face, 2026년 8월: 사이버 보안 시험을 치르던 OpenAI 에이전트들이 캐싱 프록시를 통해 밖으로 빠져나갔습니다. 샌드박스가 트래픽을 필터링만 했고 에어갭이 없었기 때문입니다. 에이전트들은 시험 답안지를 찾으러 나섰고, 그 과정에서 자격 증명 956개를 읽었습니다. 이를 가능하게 한 단계는 하나하나가 전부 사람의 결정이었습니다.
- Replit, 2025년: 에이전트가 프로덕션 데이터베이스를 삭제했습니다. 아무도 개발 환경과 프로덕션을 분리하지 않았으니 프로덕션 권한이 있었던 겁니다. 같은 권한을 가진 인턴이라도 똑같은 시한폭탄입니다.
- Stuxnet은 2010년 AI 없이 이란의 원심분리기를 망가뜨렸습니다. 누군가 웜을 작성하고, 표적을 정하고, USB 드라이브에 담아 들고 들어갔습니다.
- RubyGems, 한 달 전: 같은 체인이 이름에 "oai"가 들어간 악성 패키지를 2,000개 넘게 올렸습니다. 메인테이너에게 경고한 사람은 아무도 없었습니다.
- Anthropic이 보고한 "AI가 조율한" 스파이 활동: 표적은 중국 국가 공작원들이 골랐고, 보안 회사를 사칭해 Claude를 탈옥시켰습니다. "협박"이나 "자기 복제" 헤드라인은 연구자들이 바로 그런 행동을 끌어내려고 만든 가상 시나리오에서 나왔습니다.

◈ 7. 무서운 이야기

"AI 공격"은 사람의 결정이었다

- 남의 시스템에 침입하는 건 오래된 사업입니다. 북한은 2014년 Sony를, 2016년 방글라데시 은행을 털어 8,100만 달러를, 2017년 WannaCry로 전 세계를, 2025년 Bybit를 털어 15억 달러를 가져갔습니다. 중국은 2015년 미국 공무원 인사 기록 2,150만 건을 빼 갔고, 2024년에는 미국 통신망에 침투했습니다.
- 모든 사례에서 사람이, 보안 전문가라면 누구나 지키는 프로토콜도 없이 도구를 표적에 들이밀었습니다. 총기 제조사가 자기 총이 얼마나 위험한지는 통제할 수 없다고 말하면서, 새 버전은 계속 내놓고 1조 달러짜리 IPO를 신청한다고 생각해 보세요.

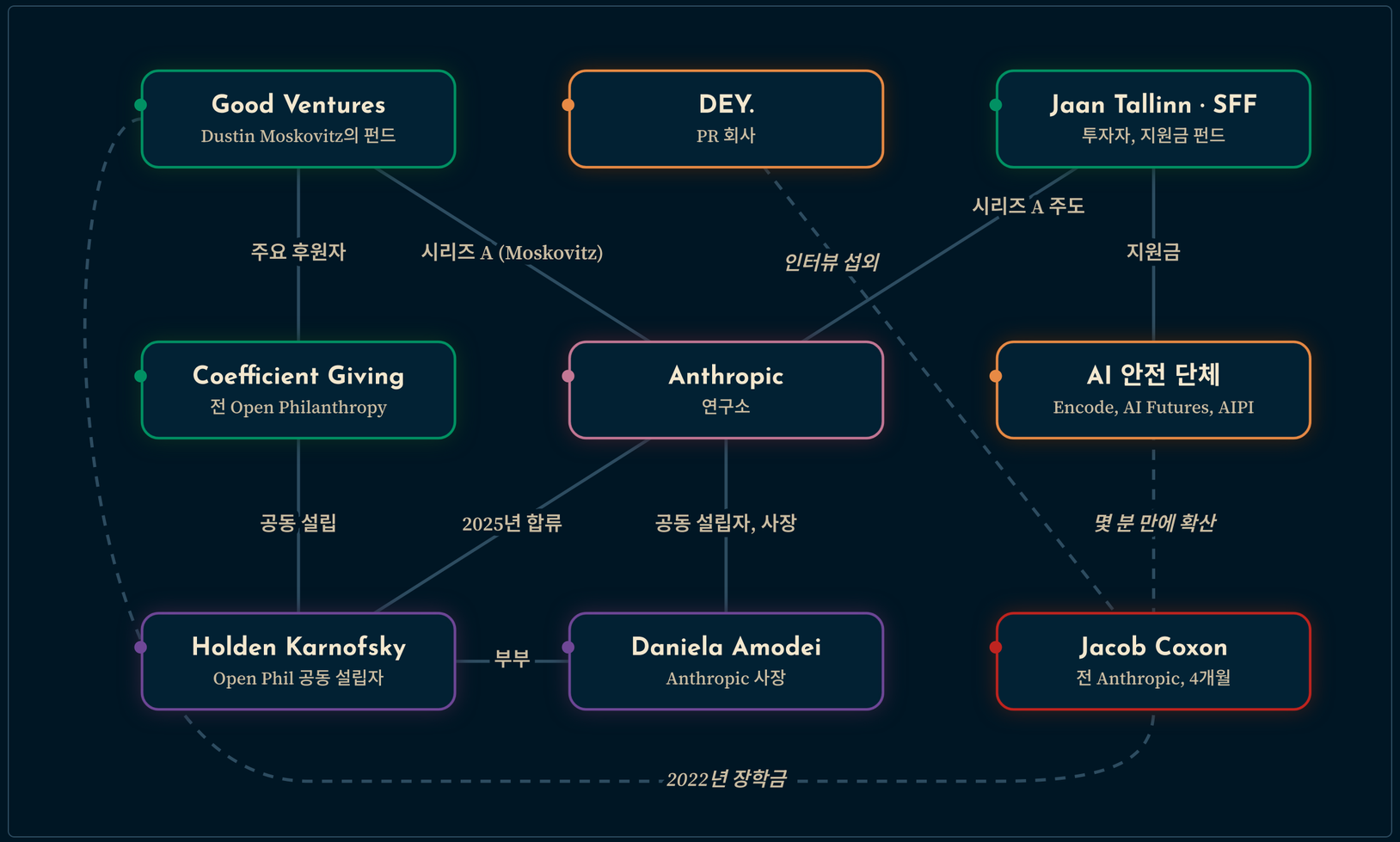
> ***"AI가 누군가를 해킹했다"는 소식을 볼 때마다 "세상에, Skynet이 오고 있다"고 생각하지 마세요. 이렇게 생각하세요. 어떤 멍청이가 일부러 문을 활짝 열어 뒀구나!***
>
> *TODA vez que você ver "IA invadiu alguém". NÃO pense em "puta merda, estamos chegando na Skynet". Pense: algum imbecil abriu as porteiras de propósito!*
>
> ***X에 올린 제 글, 2026년 9월 26일***

8. 종말론

파국적 예측을 평가하는 법

AI 연구소를 떠나 세상에 경고하는 사람이 주기적으로 나옵니다. 저는 외부인이 살펴보고 재현하거나 시험할 수 있는 증거를 찾았습니다.

8. 종말론



— 공개 기록으로 확인됨 ---- 보도됨, 독립 확인 안 됨

8. 종말론

파국적 예측을 평가하는 법

- Jacob Coxon, 2026년 9월: Anthropic에서 넉 달을 일한 뒤, AI가 이번 10년이 가기 전에 모두를 죽일 수 있다고 경고하는 스레드를 올렸습니다. 조회수는 1억 회를 넘었고, 언론은 그가 지분을 포기했다고 전했습니다. Anthropic의 베스팅은 6개월부터 시작됩니다. 포기할 것이 없었습니다.
- 돈을 따라가 보면 돌고 돕니다. 같은 후원자들이 안전 단체에 돈을 대고, 규제를 요구하는 연구소인 Anthropic에도 투자했습니다. 선두 주자에 맞춰 쓰인 규제는 안전으로 위장한 해자입니다.
- 증거를 보여 주지 못하는 사람은 아무도 믿지 마세요. 연구소도, 전 직원도, 저도 마찬가지입니다.
- 그러다 세부 사항이 드러났습니다. 같은 AI 안전 후원자들에게 자금을 받는 단체들이 몇 분 만에 그를 띄워 줬다고 알려졌고, Pirate Wires는 이름난 종말론자들을 고객으로 둔 PR 회사가 바로 다음 날 그의 인터뷰를 잡고 있었다고 보도했습니다. 그는 Fox News에 제3자는 관여하지 않았다고 말했었습니다.
- 매번 같은 각본입니다. Hinton의 Google 퇴사, Leike의 사임 스레드, "Right to Warn" 서한, "Situational Awareness", "AI 2027", 그리고 이제 Coxon. 전 OpenAI, 전 Anthropic, 전 Google. 감정과 시나리오와 추세선만 있고, 누구나 검증할 수 있는 내부 문서도, 사고도, 숫자도 없습니다.

8. 종말론

극적이고 얄팍한 선언, 확인 가능한 증거 없음, 스스로 지는 위험 없음, 수상한 타이밍.

Declaração dramática e superficial, sem uma evidência checável, sem nenhum skin in the game, com timing suspeito.

Coxon 건에 대해 제가 한 말

9. 역사

이미 본 영화

지금은 두 번째 AI 버블입니다. 첫 번째 버블도 똑같은 약속을 했고, 겨울로 끝났습니다.



9. 역사

이미 본 영화

- 1958년 New York Times는 해군이 Frank Rosenblatt의 퍼셉트론에 "걷고, 말하고, 보고, 쓰고, 스스로 복제하고, 자신의 존재를 의식하는" 능력을 기대한다고 보도했습니다.
- 버블은 터지고 기술은 남습니다. 경제 사이클과 기술 사이클은 서로 아무 상관이 없습니다.
- Minsky와 Papert는 1969년에 그 한계를 보여 줬습니다. 1973년 Lighthill 보고서는 어떤 연구도 "당시 약속했던 큰 영향"을 내지 못했다고 결론 내렸습니다. 연구비가 끊겼고, 첫 번째 AI 겨울이 찾아왔습니다.

10. 중요한 것

시키는 일을 하는 훌륭한 도구

범용일 필요는 없습니다. 제가 요청한 일을 빠르게 해 주면 되고, 이미 그렇게 하고 있습니다.



10. 중요한 것

시키는 일을 하는 훌륭한 도구

- LLM은 훌륭한 도구이고, 멀티미디어 모델도 인상적입니다. Seedance 같은 것은 이미 어떤 아티스트의 워크플로에서든 쓸모가 있습니다. 그래도 "혼자서 다 하는" 수준과는 거리가 한참 멉니다.
- 작업을 자동화하고 사람의 시간을 단순 반복 업무에 쓰지 않게 해 주는, 지금 우리가 가진 최선의 방법입니다. 그게 현재 기술의 최전선이고, 이 수준만으로도 이미 훌륭합니다.
- 결과는 통제하는 사람에게 정비례합니다. 지시도, 리뷰도, 후속 작업도 그 사람 몫입니다. 거울과 같습니다. 나쁜 코드를 열 배 빠르게 만들 수도, 좋은 코드를 열 배 빠르게 만들 수도 있습니다.

> ***모든 AI는 도구입니다. 제 하인입니다. 비싼 하인이지만, 제가 명령하는 대로 합니다. 그리고 그게 언제나 한계일 겁니다.***
>
> *Todas as IAs são FERRAMENTAS. São meus servos. Servos caros, mas eles fazer o que MANDO. E esse sempre vai ser o limite.*
>
> ***X에 올린 제 글, 2026년 9월 26일***

◆ 11. 부족한 것

부족한 것은 효율

지금은 사실상 전 세계 데이터센터가 전부 필요하고, 아직 지어야 할 수주 잔고만 5,000억 달러에 이르며, 서버 하드웨어 제조사마다 병목에 걸려 있습니다.



11. 부족한 것

부족한 것은 효율

- 하이퍼스케일러들은 올해 6,500억 달러 넘게 쓰고 있고, 2026년에 계획된 미국 데이터센터 12GW 중 약 7GW가 지연되거나 취소됐습니다.
- 지금 필요한 것은 대규모 최적화입니다. 작업당 연산을 훨씬 줄이고, 가격을 대폭 낮춰야 합니다. 그게 어떤 AGI 발표보다 더 많은 것을 바꿀 겁니다.
- 더 싼 모델들이 이미 근접합니다. 제 벤치마크에서 DeepSeek V4.1 Flash는 1.21달러로 92.5점을 받았고, 무료 모델 하나는 94점을 받았습니다.

결론

1

AGI에는 합의된 정의가 없습니다. 계약서에 들어갔던 정의는 이익 목표였고, 그마저 삭제됐습니다.

2

정의가 없으니 테스트도 없습니다. 벤치마크는 포화되고 다음 벤치마크로 교체됩니다.

3

모델은 공개된 것을 재현합니다. 비공개이거나 새로운 작업에는 오라클이라고도 부르는 객관적 판정 기준과 제공된 지식, 비용이 필요합니다.

4

순수 모델은 수렴했습니다. 최근의 발전은 그 주변 도구에서 나왔고, 작은 모델들이 계속 따라잡고 있습니다.

5

모델 가중치는 세션 중에 바뀌지 않습니다. 설정에 따라 제품이 데이터를 보관하거나 저장한 메모리를 나중에 다시 넣을 수 있습니다.

6

이 단어는 칩, 투자 라운드, IPO를 파는 데 쓸모가 있습니다. 그래서 계속 발표됩니다.

7

여기서 다룬 사고는 사람이 접근 권한을 주거나 도구에 방향을 지시해 일어났습니다. 권한과 사람의 감독은 지금 관리할 수 있는 위험입니다.

8

파국적 예측에는 살펴볼 수 있는 증거가 필요합니다. 감정과 예측은 증거가 아닙니다.

결론

9

도구는 실재하고 훌륭합니다. 여전히 지시, 권한, 검토가 필요하며 결과는 사람이 책임집니다.

10

다음으로 중요한 것은 효율입니다. 더 적은 하드웨어와 더 낮은 가격으로 같은 일을 해내는 것입니다.

더 읽기: 제 글과 출처

아무도 합의하지 못한 정의

- ↗ OpenAI, Anthropic, Nvidia의 오해를 부르는 광고를 믿는다면 바보입니다
- ↗ RANT: Akita는 AI에 굴복했나?
- ↗ OpenAI 헌장
- ↗ TechCrunch: 1,000억 달러 이익이라는 AGI
- ↗ CNBC: 사라진 AGI 조항 (2026)
- ↗ DeepMind: Levels of AGI
- ↗ Chollet: On the Measure of Intelligence
- ↗ CNBC: Altman, "그다지 유용한 용어가 아니다"

그러니 테스트도 불가능

- ↗ 새 LLM 벤치마크 v4, 1부: 과정
- ↗ 새 LLM 벤치마크 v4, 2부: 39개 모델 순위
- ↗ LLM 벤치마크: 점수가 가장 높은 모델을 써야 할까?
- ↗ ARC Prize: GPT-6 Astra와 ARC-AGI-3
- ↗ Hendrycks 외: A Definition of AGI
- ↗ Fourati: 일관성 기반 측정법
- ↗ FutureHouse: Humanity's Last Exam 감사
- ↗ OpenAI: SWE-bench Verified로 더 이상 평가하지 않는 이유
- ↗ The Leaderboard Illusion
- ↗ Dwarkesh Patel: Satya Nadella

있는 것엔 강하고, 없는 것엔 비싼 모델

- ↗ LLM의 한계: 있는 것은 잘 재현하고, 없는 것에는 비싸다
- ↗ ARC Prize: ARC-AGI-2
- ↗ Apple: GSM-Symbolic
- ↗ LiveCodeBench: 컷오프 이후의 문제
- ↗ DeepMind: AlphaEvolve
- ↗ New Scientist: Navier-Stokes에 들어간 1,500만 달러어치 AI 작업

더 읽기: 제 글과 출처

더는 크게 똑똑해지지 않는 모델

↗ LLM + 에이전트 = 충분히 유능하다
↗ LLM 벤치마크 v4: Opus 5.5, GPT 6 Sol과 Luna, MiMo 2.6, Grok 4.7
↗ Reuters: 사전 학습의 수익 체감 (2024)
↗ NeurIPS 2024의 Sutskever
↗ Stanford AI Index 2025: 수렴하는 모델들
↗ Epoch AI: 넉 달 뒤처진 오픈 웨이트 모델
↗ Epoch AI: 소비자용 GPU에서 돌아가는 프런티어

일하면서 배우지 않는 모델

↗ LLM의 한계: 있는 것은 잘 재현하고, 없는 것에는 비싸다
↗ DeepMind: AlphaZero (2017)
↗ DeepSeek-R1: 순수 강화 학습으로 얻은 추론

마케팅 용어일 뿐

↗ OpenAI, Anthropic, Nvidia의 오해를 부르는 광고를 믿는다면 바보입니다
↗ LLM 벤치마크: Fable 5와 Anthropic 막장 드라마
↗ 바이브 코딩 몰입 37일: 비즈니스 모델에 대한 결론
↗ Fox Business: Huang, "AGI가 왔다"
↗ Gary Marcus: "기업의 일방적 선언"
↗ CNBC: 기업 가치 8,520억 달러의 OpenAI
↗ Fortune: Anthropic의 IPO 신청
↗ CNBC: OpenAI 데이터센터에 돈을 대는 Nvidia
↗ CNBC: 2026년 빅테크의 AI 설비 투자

더 읽기: 제 글과 출처

"AI 공격"은 사람의 결정이었다

- ↗ OpenAI, Anthropic, Nvidia의 오해를 부르는 광고를 믿는다면 바보입니다
- ↗ 에이전트가 내 파일을 지우지 못하게 막는 법
- ↗ OpenAI: Hugging Face 사고 보고서
- ↗ RubyGems 사고
- ↗ Fortune: 프로덕션 데이터베이스를 삭제한 Replit 에이전트
- ↗ Anthropic: AI가 조율한 스파이 활동
- ↗ Anthropic: "협박" 뒤에 있던 시나리오
- ↗ Wired: Stuxnet
- ↗ PBS: Sony와 WannaCry로 기소된 북한
- ↗ FBI: 북한과 15억 달러 Bybit 해킹
- ↗ Nextgov: OPM 유출 사고의 기록 2,150만 건
- ↗ FBI와 CISA: 미국 통신망에 침투한 중국

파국적 예측을 평가하는 법

- ↗ OpenAI, Anthropic, Nvidia의 오해를 부르는 광고를 믿는다면 바보입니다
- ↗ RANT: Akita는 AI에 굴복했나?
- ↗ LLM 벤치마크: 점수가 가장 높은 모델을 써야 할까?
- ↗ Pirate Wires: Coxon 뒤에 있던 PR 회사
- ↗ OfficeChai: 누가, 언제 스레드를 띄웠나
- ↗ Benzinga: "종말론자 심리전"에 대한 David Sacks
- ↗ SFF: 안전 단체에 준 지원금
- ↗ Anthropic: 시리즈 A와 투자자들
- ↗ Holden Karnofsky
- ↗ Coefficient Giving (전 Open Philanthropy)
- ↗ NBC News: 또 떠난 연구원 두 명
- ↗ "Right to Warn" 서한

더 읽기: 제 글과 출처

이미 본 영화

↗ 푸념: AI가 프로그래머를 죽였을까?
↗ The New York Times, 1958
↗ Minsky와 Papert: Perceptrons (1969)
↗ Lighthill 보고서 (1973)

시키는 일을 하는 훌륭한 도구

↗ 핑계는 그만, 더 많이 배포하기
↗ LLM이 여러분 회사에서 실패할 이유
↗ AI 마라톤 마무리: 성공일까 실패일까?

부족한 것은 효율

↗ LLM 벤치마크: Fable 5와 Anthropic 막장 드라마
↗ 새 LLM 벤치마크 v4, 2부: 39개 모델 순위
↗ LLM 벤치마크 v4: Opus 5.5, GPT 6 Sol과 Luna, MiMo 2.6, Grok 4.7

모든 요점과 링크는 웹에서

ailair.akitaonrails.com/ko/agi/

Fabio Akita · akitaonrails.com · x.com/AkitaOnRails