

AI 코딩 에이전트의 등장과  
소프트웨어 엔지니어(링)의 미래

2025.7.3

박 재 흥

(주)마이씨큐 대표이사

jaehong@mycq.co.kr



클라우드마리스타 커뮤니티

제 11 차 컨퍼レン스를

축하드립니다.



# Vibe Coding

## Demo: Vibe Coding (4m)

### Create a Twitter App

- **Cline**
- **Cursor**

### Fix Github Issue

- **Claude Code**
- **Google Jules**

# Vibe Coding (2025.2.3)



Andrej Karpathy

@karpathy



...

There's a new kind of coding I call "vibe coding", where you fully give in to the vibes, embrace exponentials, and forget that the code even exists. It's possible because the LLMs (e.g. Cursor Composer w Sonnet) are getting too good. Also I just talk to Composer with SuperWhisper so I barely even touch the keyboard. I ask for the dumbest things like "decrease the padding on the sidebar by half" because I'm too lazy to find it. I "Accept All" always, I don't read the diffs anymore. When I get error messages I just copy paste them in with no comment, usually that fixes it. The code grows beyond my usual comprehension, I'd have to really read through it for a while. Sometimes the LLMs can't fix a bug so I just work around it or ask for random changes until it goes away. It's not too bad for throwaway weekend projects, but still quite amusing. I'm building a project or webapp, but it's not really coding - I just see stuff, say stuff, run stuff, and copy paste stuff, and it mostly works.

8:17 AM · Feb 3, 2025 · 4.1M Views

새로운 종류의 코딩을 나는 "바이브 코딩(vibe coding)" 이라고 부른다. 이건 그냥 바이브(느낌)에 완전히 몸을 맡기고, 지수적(exponential) 발전을 즐기면서, 코드가 있다는 사실조차 잊어버리는 방식이다. 이렇게 가능한 이유는 LLM들(예: Cursor Composer + Sonnet)이 너무 좋아졌기 때문이다. 게다가 나는 SuperWhisper를 통해 Composer에게 그냥 말만 하면 되니까 키보드에도 거의 손을 안 댄다.

나는 너무 귀찮아서 "사이드바의 패딩을 절반으로 줄여줘" 같은 가장 바보 같은 요구사항도 그냥 말로 한다. 무조건 "모두 수락(Accept All)" 버튼을 누르고, 코드 차이점(diff) 같은 건 읽지도 않는다. 여러 메시지가 뜨면 아무런 코멘트 없이 메시지를 복사해서 붙여넣기만 하는데, 보통 그렇게 하면 해결된다. 코드는 이미 내가 평소에 이해하는 수준을 넘어 성장해 버렸고, 제대로 이해하려면 정말 시간을 들여서 읽어야 할 지경이다.

가끔 LLM이 버그를 못 잡으면, 그냥 그 주변을 우회하거나, 랜덤하게 뭔가 변경해 달라고 요청해서 버그가 사라질 때까지 반복하기도 한다. 주말에 재미삼아 만드는 프로젝트 정도라면 이 방식도 나쁘지 않다. 오히려 꽤나 재미있다. 뭔가 프로젝트나 웹앱을 만들고는 있지만, 이걸 정말 코딩이라고 할 수는 없다—그냥 보고, 말하고, 실행하고, 붙여넣기만 하면, 대체로 다 돌아가기 때문이다.

# Vibe Coding

## 바이브 코딩(Vibe Coding) ♪

AI 코딩 에이전트와 대화를 주고 받으면서,  
바이브(느낌)에 따라 코딩하는 방법

### 장점

- \* 개발 속도 급증: 10x~100x
- \* 진입장벽 감소: 비 개발자도 앱 개발 가능
- \* UX에 집중: 코드 작성 보다 기능과 사용자 경험에 더 많은 시간 투자

### 단점

- \* 확장성: 프로토타이핑(0 to 1) 뛰어남, 기능과 규모 확장(1 to N)에는 한계.
- \* 디버깅: AI는 코드 생성은 잘하지만, 복잡한 버그 해결은 (아직) 어려움.
- \* 품질: AI가 생성한 코드를 이해하고 평가할 수 있는 능력 필요.



# AI Coding Agent

# AI Coding Agent: 2025

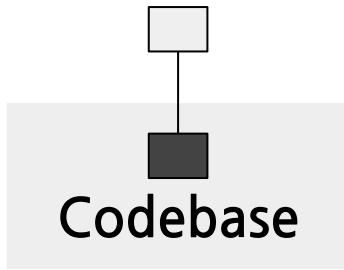
## General-purpose

<p># IDEs</p> <ul style="list-style-type: none"><li>• <b>cursor:agent_mode (anysphere)</b> 🚀</li><li>• <b>cline (open source)</b></li><li>• <b>github-copilot:agent_mode (microsoft)</b></li><li>• windsurf (meta)</li><li>• devin (cognition)</li><li>• cosine</li><li>• roo-code (open source)</li><li>• swe-agent (open source)</li><li>• open hands (open source)</li><li>• refact.ai (open source)</li></ul> <p># TUI, CLI</p> <ul style="list-style-type: none"><li>• <b>claude-code (anthropic)</b> 🚀</li><li>• codex, codex cli (openai)</li><li>• gemini cli (google)</li><li>• aider (open source)</li></ul>	<p># ChatBots</p> <ul style="list-style-type: none"><li>• <b>chatgpt</b> 🚀</li><li>• <b>claude</b></li><li>• <b>gemini</b></li><li>• grok</li></ul> <p># IDEs</p> <ul style="list-style-type: none"><li>• <b>cursor</b> 🚀</li><li>• <b>github-copilot</b></li><li>• windsurf</li><li>• tabnine</li><li>• codi</li><li>• zed</li></ul> <p># APIs</p> <ul style="list-style-type: none"><li>• <b>claude opus-4, sonnet-4</b> 🚀</li><li>• openai o3, gpt-4</li><li>• google gemini 2.5 pro</li></ul>
<p><b>Agent</b></p> <p># Web IDE (no code, low code)</p> <ul style="list-style-type: none"><li>• <b>lovable (web gui)</b> 🚀</li><li>• <b>cloud-barista ai agent (multi-cloud, etri)</b></li><li>• v0</li><li>• bolt.new</li><li>• jules (github committer, google:beta)</li></ul> <p># Web IDE (all in one)</p> <ul style="list-style-type: none"><li>• <b>replit (web ide, quick prototyping)</b> 🚀</li></ul>	<p><b>Assistant</b></p> <ul style="list-style-type: none"><li>• amazon Q (business, developer)</li><li>• gitlab duo (gitlab ai assistant)</li><li>• code rabbit (code review)</li><li>• qodo (SDLC automation, former codium)</li></ul>

## Special-purpose

# AI Coding Agent

search engine,  
community



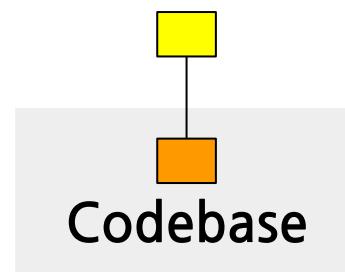
search code  
copy & paste  
customize

(manual)

assistant  
2022 ~

ChatGPT  
Claude  
Gemini

...

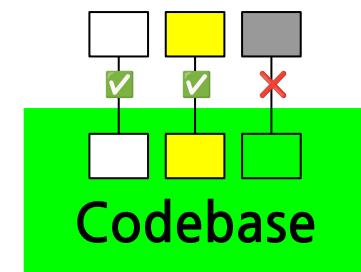


prompt(context)  
assistant generates code  
review  
copy & paste  
customize  
(manual)

agent  
2025 ~

Cursor  
Github Copilot  
Cline

...



prompt(goal)  
agent(context) suggests code  
review  
(auto) accept or reject  
[customize]  
(autonomous)

## AI가 사실상 모든 코드를 작성할 것

(다리오 아모데이 엔트로픽 CEO, 2025.3.11)

### 급진적

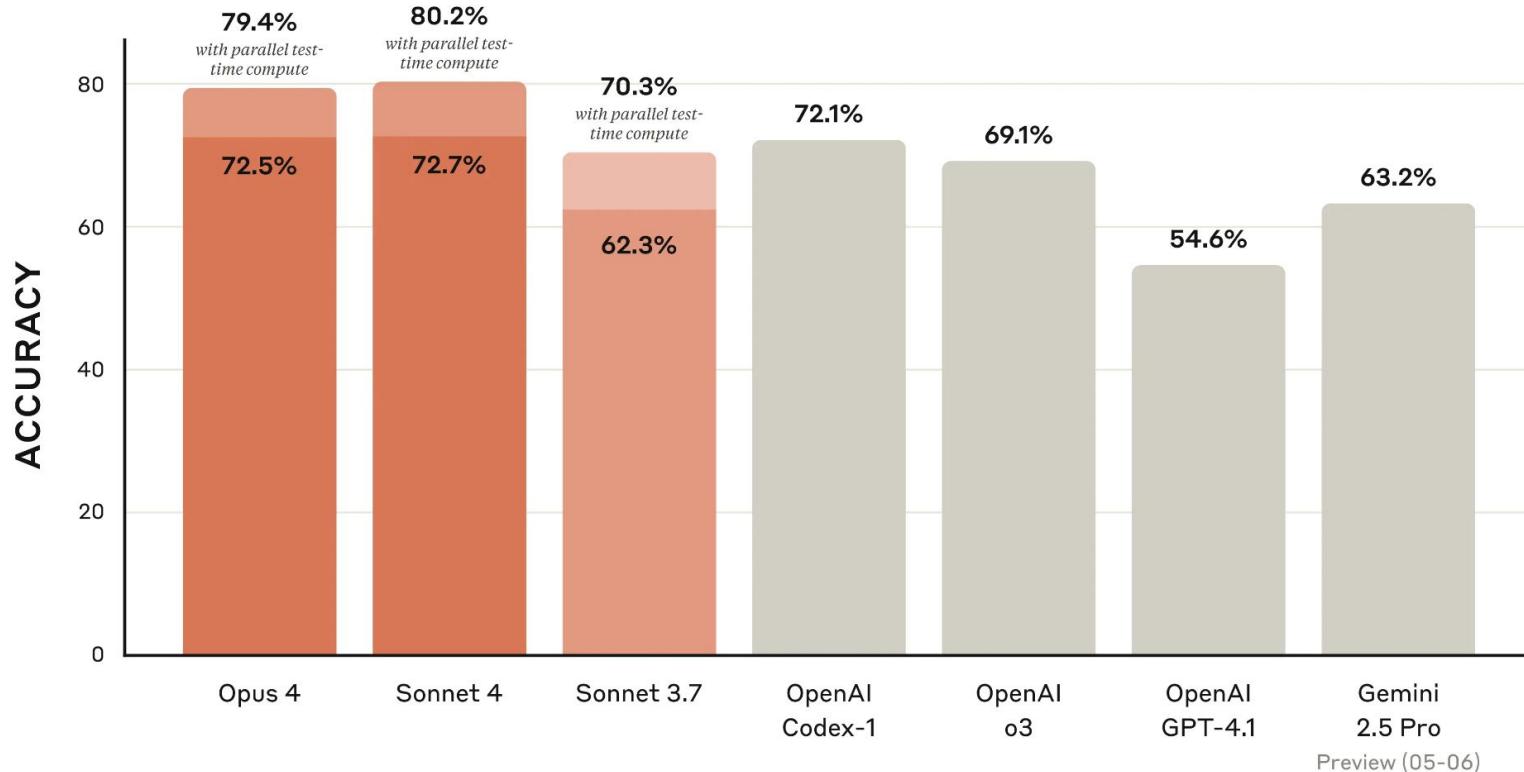
- 다리오 아모데이 (엔트로픽 CEO): 2026.3 이전, AI가 사실상 모든 코드를 작성할 것. (코딩, 제일 잘나감)
- 샘 알트만 (OpenAI CEO): 2025년 안에 중대한 변화가 있을 것. (LLM, 제일 잘나감)
- 젠슨 황 (엔비디아 CEO): 이제, 말하기만 하면 됨 (GPU 제일 잘나감)
- 일론 머스크: 5년 안에(2030) 모든 코드 작성, 이후, 모든 직업 소멸할 것 (잘나가고 싶음)
- 사티아 나델라 (마이크로소프트 CEO): 현재 내부 코드의 20-30%를 AI가 작성
- 마크 저커버그: 2025년까지 중급 엔지니어 역할할 것
- 순다 피차이 (구글 CEO): 현재 신규 코드의 30% 이상을 AI가 생성 중

### 점진적

- 마이클 트루엘 (Cursor CEO): 바이브 코딩 비추, 전문 환경에서 한계, 인간 검토 필수, 현재 40-50% AI 작성 & 인간 검토, (AI 코딩 에이전트, 제일 잘나감)
- 토마스 둠케 (Github CEO): 80% 작성 예측, 개발자 대체 아닌 역할 변화 강조

# Software engineering

SWE-bench verified





\*Forecast: Going from 1 week to 1 year might be ~2x easier than going from 1 hour to 1 week. Reasoning: 1 week tasks can be much more complex than 1 hour tasks, but we project there aren't as many extra skills needed to go from 1 week to 1 year.

ASI(Artificial Super Intelligence), 초인공지능: 2000x, 2028 초

Artificial Superintelligence  
2000x AI R&D Multiplier

슈퍼인텔리전트 AI 리서처: 250x, 2027 말

Superintelligent AI Researcher  
250x AI R&D Multiplier

슈퍼휴먼 리모트 워커: 100x, 2027.10

Superhuman Remote Worker  
100x AI R&D Multiplier

슈퍼휴먼 AI 리서처: 25x, 2027 중반

Superhuman AI Researcher  
25x AI R&D Multiplier

슈퍼휴먼 코더: 4x, 2027 초반

Superhuman Coder  
4x AI R&D Multiplier

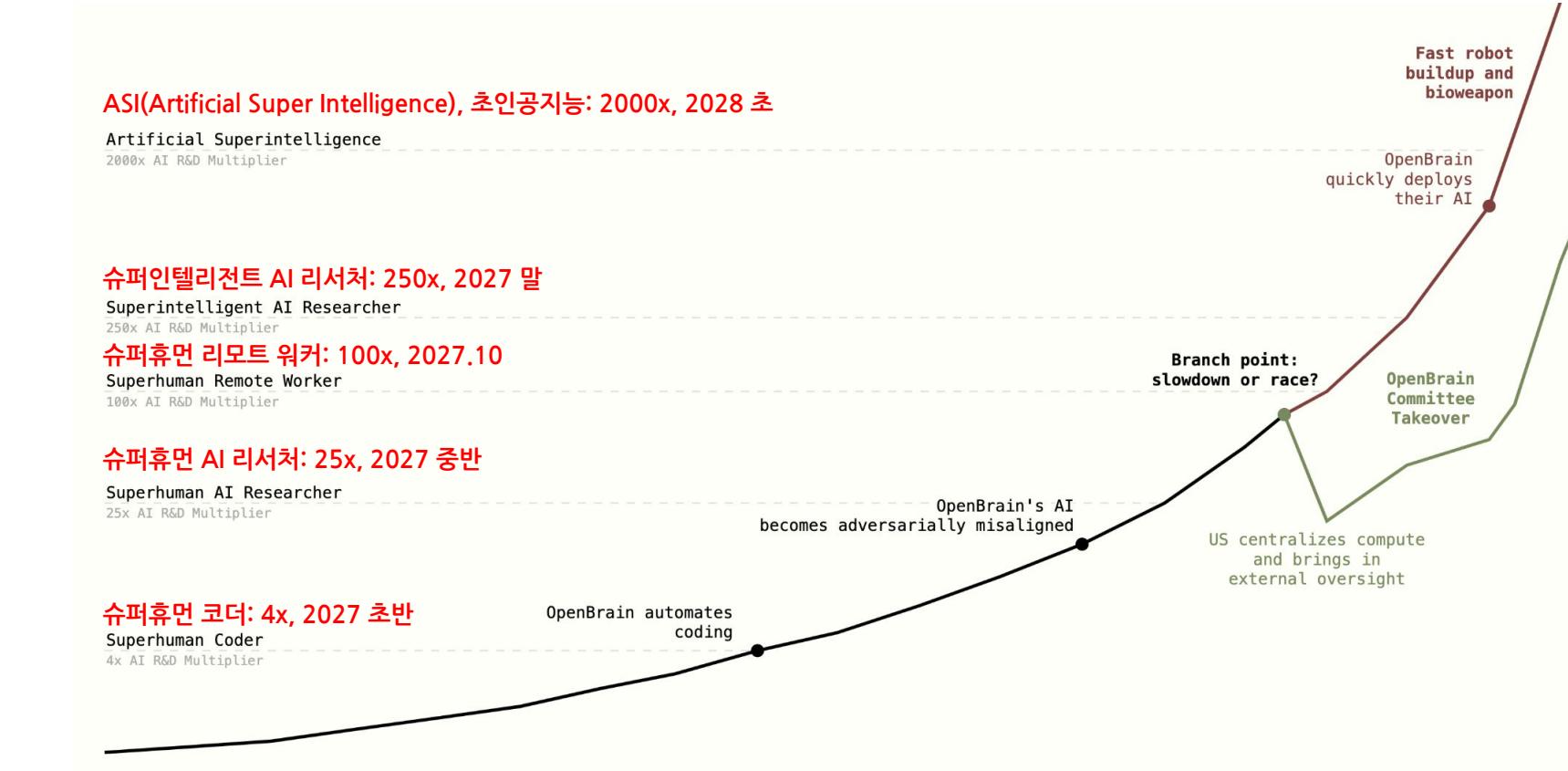
OpenBrain automates coding

Apr

Jul

Oct

2028



# AI Agent Engineering

# Software 1.0, 2.0, 3.0

## Software 1.0

```
python
def quicksort(arr):
    if len(arr) <= 1:
        return arr

    pivot = arr[len(arr) // 2]
    left = [x for x in arr if x < pivot]
    middle = [x for x in arr if x == pivot]
    right = [x for x in arr if x > pivot]

    return quicksort(left) + middle + quicksort(right)

# 사용 예시
numbers = [64, 34, 25, 12, 22, 11, 90]
sorted_numbers = quicksort(numbers)
print(sorted_numbers) # [11, 12, 22, 25, 34, 64, 90]
```

### Code

## Software 2.0 e.g., LLM

```
python
# data.csv 파일 예시:
# age,income,education,purchased
# 25,50000,bachelor,0
# 35,75000,master,1
# 45,90000,phd,1

import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier

# 1. 학습 데이터 로드
data = pd.read_csv('customer_data.csv')
X = data[['age', 'income', 'education', 'purchased']]
y = data['purchased']

# 2. 하이퍼파라미터 설정
model = RandomForestClassifier(
    n_estimators=100, # 하이퍼파라미터
    max_depth=10, # 하이퍼파라미터
    random_state=42 # 하이퍼파라미터
)

# 3. 학습 알고리즘 실행
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
model.fit(X_train, y_train)

# 4. 예측
prediction = model.predict([[30, 60000, 2]]) # 학습 출력
```

### Training Data

### Training Algorithm

### Model Weights

## Software 3.0

### Python 코드 최적화 시스템 프롬프트

### Prompt

당신의 임무는 제공된 Python 코드 스니펫을 분석하고 성능 최적화 개선 방안을 제안하는 것입니다. 코드를 더 효율적이고, 빠르며, 리소스 점약적이지 않게 만들 수 있는 영역을 식별하세요. 최적화에 대한 구체적인 제안을 제공하고, 이러한 변경사항이 코드 성능을 어떻게 향상시킬 수 있는지 설명하세요. 최적화된 코드는 원본 코드와 동일한 기능을 유지하면서 향상된 효율성을 보여주어야 합니다.

### 분석 및 최적화 지침

#### 1. 성능 병목 지점 식별

- 시간 복잡도가 높은 구간 분석
- 메모리 사용량이 과도한 부분 확인
- 불필요한 연산이나 반복 작업 탐지
- I/O 작업의 비효율성 점검

#### 2. 최적화 제안 영역

- 알고리즘 개선: 더 효율적인 알고리즘이나 자료구조 제안
- Python 내장 기능 활용:** 고성능 내장 함수 및 라이브러리 사용
- 메모리 최적화: 불필요한 메모리 할당 줄이기, 제너레이터 활용
- 연산 최적화: 중복 계산 제거, 캐싱 활용
- 라이브러리 최적화: NumPy, Pandas 등 고성능 라이브러리 활용

#### 3. 분석 결과 제시 형식

## 🔍 코드 분석 결과

- ### 현재 성능 특성
- 시간 복잡도: O(?)
  - 공간 복잡도: O(?)
  - 주요 병목점: [구체적 식별]

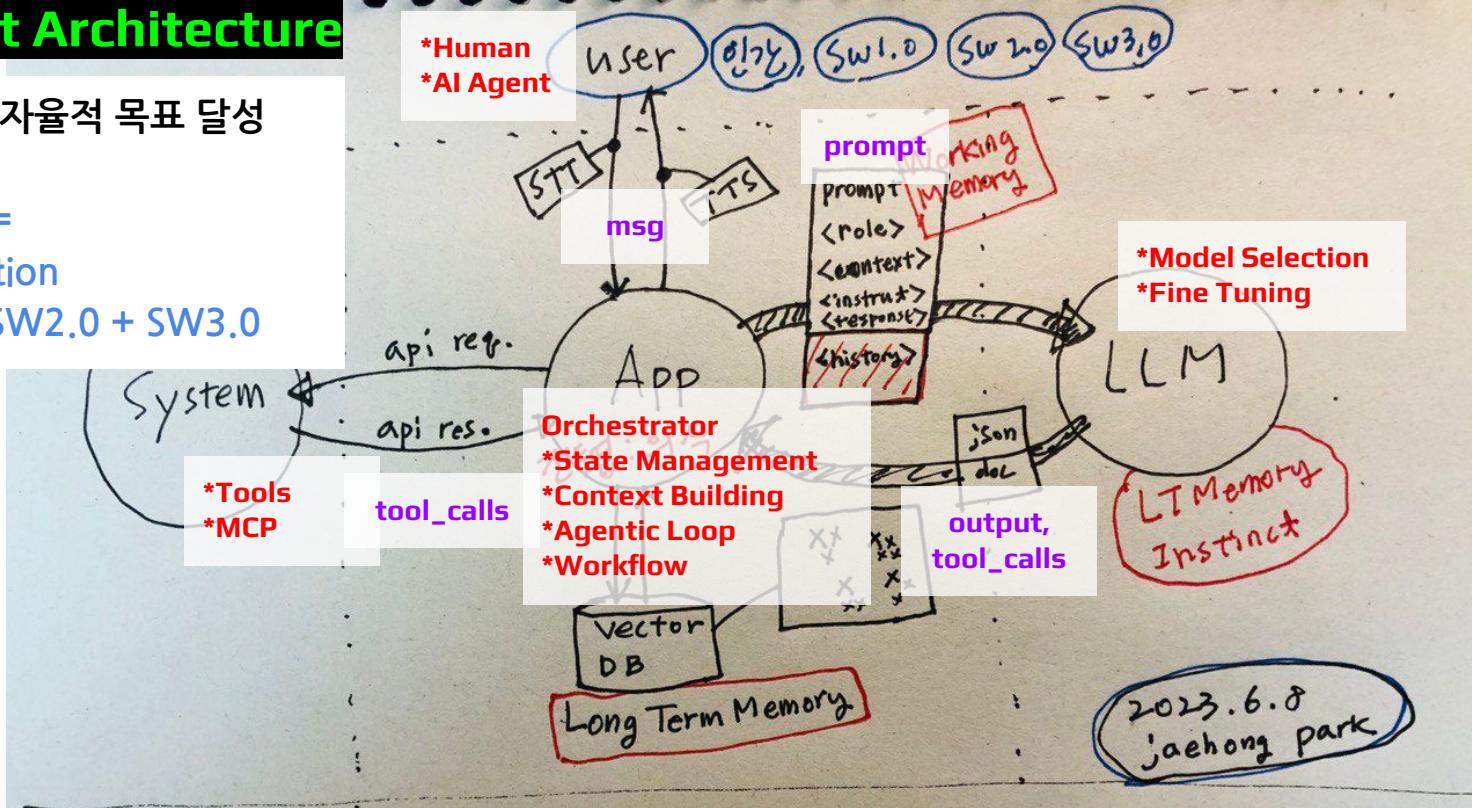
# Software 1.0, 2.0, 3.0

	Software 1.0	Software 2.0	Software 3.0
등장 시기		2012(AlexNet) ~	2022.11(ChatGPT, GPT3.5) ~
자산	소스코드	학습 데이터 학습 알고리즘 모델 가중치(Weights)*	프롬프트(Prompt)
개발방법	programming	training	prompt engineering (design)
출력특징	결정적	예측: 결정적 or 비결정적	생성: 비결정적 or 결정적(temp=0) <b>In-Context Learning 으로 가중치 변경없이 학습 가능</b> , Structured Output 출력 가능: e.g., JSON output. => Function Call
개발도구	프로그래밍 언어 (예: Python, C/C++, Go, Java)	딥러닝 프레임워크 (Pytorch, Tensorflow 등) + 학습 플랫폼	자연어 (한국어, 영어, Math, Code 등) + Image, Audio, Video
실행환경	OS: Windows, Linux, MacOS	딥러닝 프레임워크 (Pytorch, Tensorflow 등)	LLM: GPT, Claude, Gemini 등
협업도구	Github	Hugging Face	Prompt Library

# AI Agent Architecture

AI Agent: 자율적 목표 달성

AI Agent =  
Orchestration  
SW1.0 + SW2.0 + SW3.0



SW 1.0

알고리즘

프로세스

SW 3.0

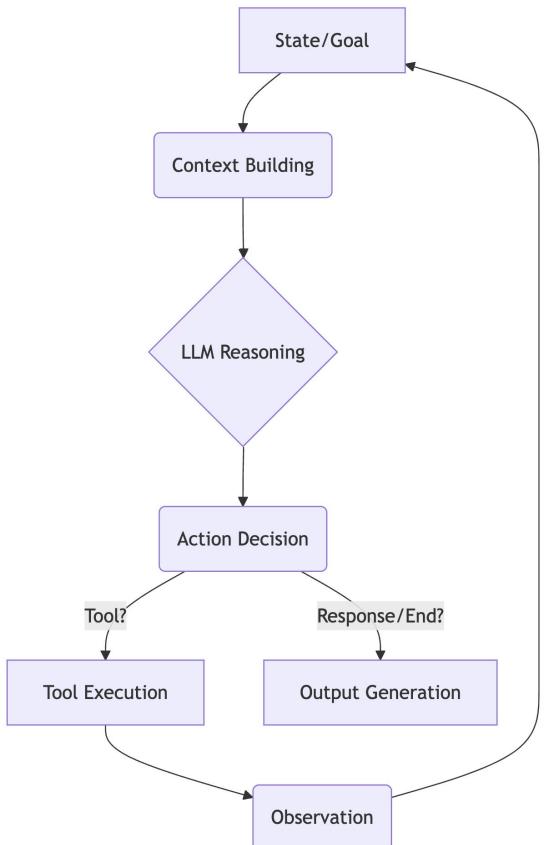
프롬프트 엔지니어링

≠ SW 1.0 + SW 2.0

SW. 2.0

Data 학습

# AI Coding Agent: Agentic Loop



```
def loop(llm):
    msg = user_input()
    while True:
        output, tool_calls = llm(msg)
        print("Agent: ", output)
        if tool_calls:
            msg = [ handle_tool_call(tc) for tc in tool_calls ]
        else:
            msg = user_input()
```

sw1.0

sw2.0: llm

sw3.0: msg

# AI Agent Challenges (LLM Limitations)

- 데이터 정렬(Data Alignment) 한계
    - 사전훈련 모델을 즉시 사용하기에는 기업 요구사항과 맞지 않음
    - 해결방안: Fine-Tuning, RAG
  - 추론(Reasoning) 한계
    - 일정 수준 이상의 복잡한 다단계 작업 처리 불가
    - 해결방안: Agentic Loop, Workflow, 프롬프트 엔지니어링 등으로 작업을 분리
  - 컨텍스트 윈도우(Context Window) 한계
    - 제한된 사이즈의 컨텍스트만 처리 가능
    - 해결방안: RAG, MCP, Prompt Compression, Caching, Semantic Caching
  - 비결정적 응답 문제
    - 동일한 프롬프트에도 다른 대답, LLM 의존성 (LLM에 따라 다른 대답)
    - 해결방안: 프롬프트 자동생성 (DSPy 등)
  - 환각(Hallucination) 문제
    - 사실 같은 거짓 정보 생성, 기억 압축 손실로 인한 근본적 문제
    - 해결방안: Fine-Tuning, 프롬프트 엔지니어링, RAG, MCP, Tools,
  - 폐쇄 루프(Closed Loop) 한계
    - 실세계 정보 접근 불가. 예) 오늘 날씨는? 대답 불가!
    - 해결방안: 프롬프트 엔지니어링, RAG, MCP, Tools
  - 제어 가능성 (Controllability) 한계
    - LLM의 비결정적 응답 문제로 인해, 도구 실행 승인/거부를 위한 제어 수준(How, When, Who) 결정이 어려움
    - 해결방안: Workflow, Agentic Loop
  - 시험의 한계
    - LLM의 비결정적 응답 문제로 인해, 단일 정답이 없음, 명확한 통과/실패 기준 정의 어려움
    - 해결방안: 인간 평가자 활용 등
  - 의미 관찰(Semantic observability) 한계
    - AI의 '의미적' 품질이 정말 좋은지 나쁜지 자동으로 판단하기 어려움
    - 해결방안: 사용자 반응 관찰 등
  - 성능(Performance) 문제
    - 긴 프롬프트나 멀티라운드 대화로 추론 시간 증가, 성능 저하, AI 모델 운영 비용 문제
    - 해결방안: 모델 압축 (Quantization, Distillation, Pruning)
  - 컴플라이언스 및 규제 (Compliance & regulations) 문제
    - AI 관련 법규의 완전한 준수의 어려움. 예) EU AI Act, GDPR 등
    - 해결방안: AI모델 라이선싱(RAIL 등) 준서, SW명세서(AIBOM, SPDX) 제공
  - 체험의 한계
    - 사과맛은 어때? 달고 맛있어요! LLM은 실제 사과 맛을 경험한 적이 없음.
    - 해결방안: Embodied AI
- Solutions
    - Fine Tuning
    - Prompt Engineering
    - Prompt Caching, Compression
    - Semantic Caching
    - Prompt Automation
    - RAG
    - MCP, A2A, ACP
    - Agentic Loop
    - Workflow
    - User Response Monitoring
    - Human Review
    - AI Model Licencing
    - Embodied AI
    - ...

# CPU vs. LPU (LLM)

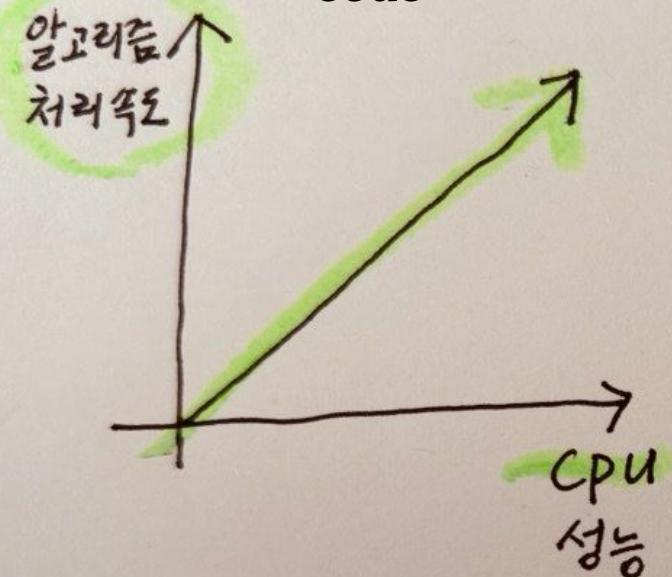
현재 모델의 한계를 보완하기 위해 복잡한 시스템을 구축하는 대신,  
2개월마다 더 나은 모델이 나올 것이라는 믿음으로 제품을 설계하라.

OpenAI CPO Kevin Weil

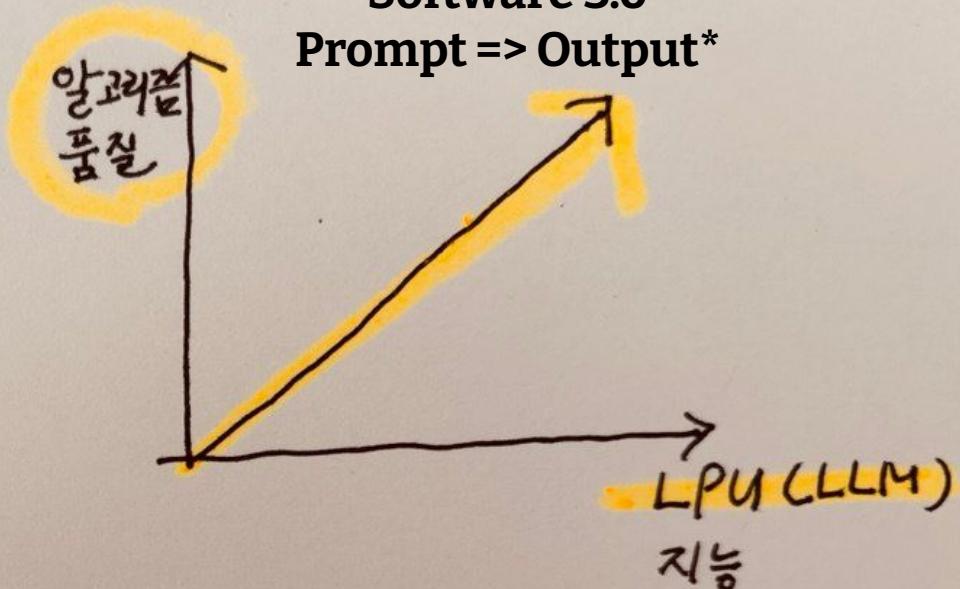
지금 사용하는 AI는

내 남은 인생에서 가장 나쁘고, 가장 비싸다.

## Software 1.0 Code\*



## Software 3.0 Prompt => Output\*



## Absolute Zero: Reinforced Self-play Reasoning with Zero Data

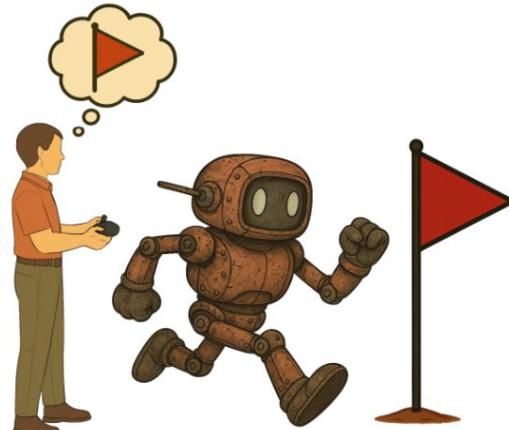
Andrew Zhao<sup>1</sup>, Yiran Wu<sup>3</sup>, Yang Yue<sup>1</sup>, Tong Wu<sup>2</sup>, Quentin Xu<sup>1</sup>, Yang Yue<sup>1</sup>, Matthieu Lin<sup>1</sup>, Shenzhi Wang<sup>1</sup>, Qingyun Wu<sup>3</sup>, Zilong Zheng<sup>2,✉</sup> and Gao Huang<sup>1,✉</sup>

<sup>1</sup>Tsinghua University <sup>2</sup>Beijing Institute for General Artificial Intelligence <sup>3</sup>Pennsylvania State University  
zqc21@mails.tsinghua.edu.cn, yiran.wu@psu.edu, zlzheng@bigai.ai, gao.huang@tsinghua.edu.cn

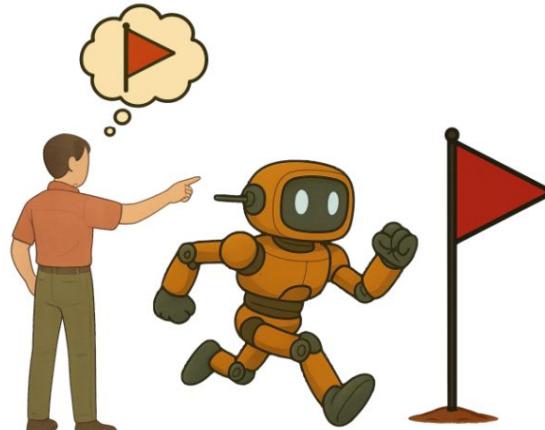
## Absolute Zero: Reinforced Self-play Reasoning with Zero Data

RLHF(Reinforcement Learning Human Feedback)

Supervised Learning

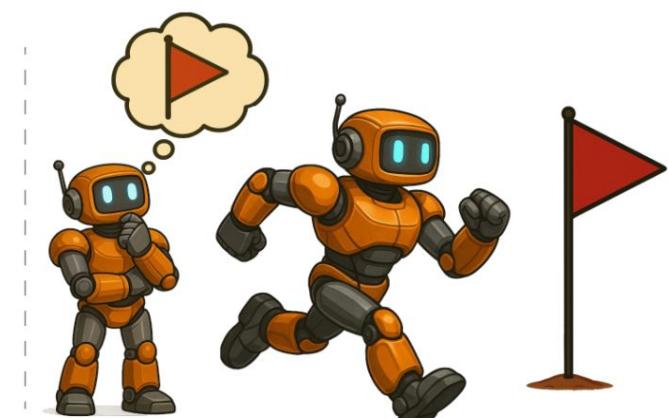


Reinforcement Learning with Verifiable Rewards



RLVR

Absolute Zero (Ours)

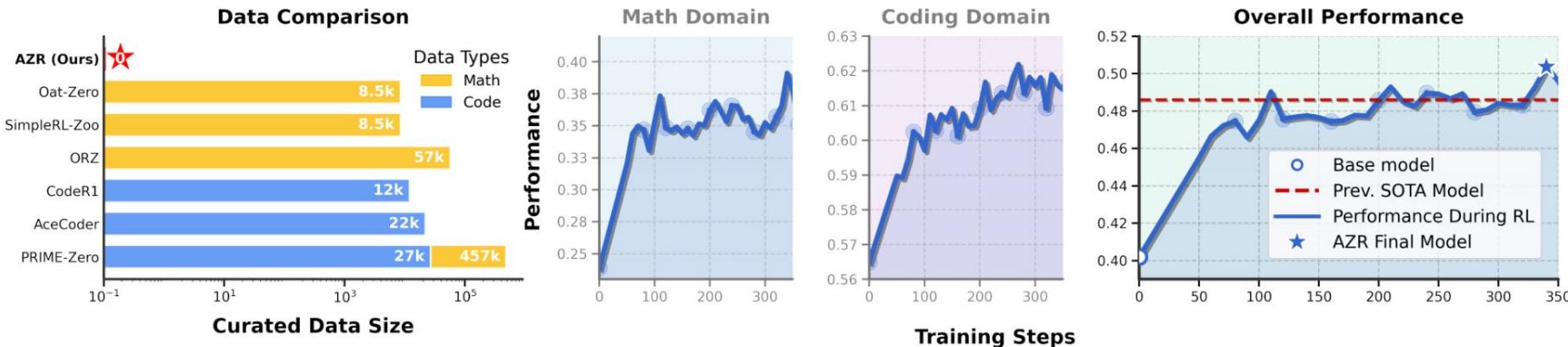


Less Human Supervision

# AZR (Absolute Zero Reasoner)

## AZR은 ZERO DATA로 SOTA 달성

자체적인 학습을 통해 수학 및 코딩 추론 능력에서 수만 개의 전문가가 레이블링한 도메인 예제로 훈련된 모델들을 능가

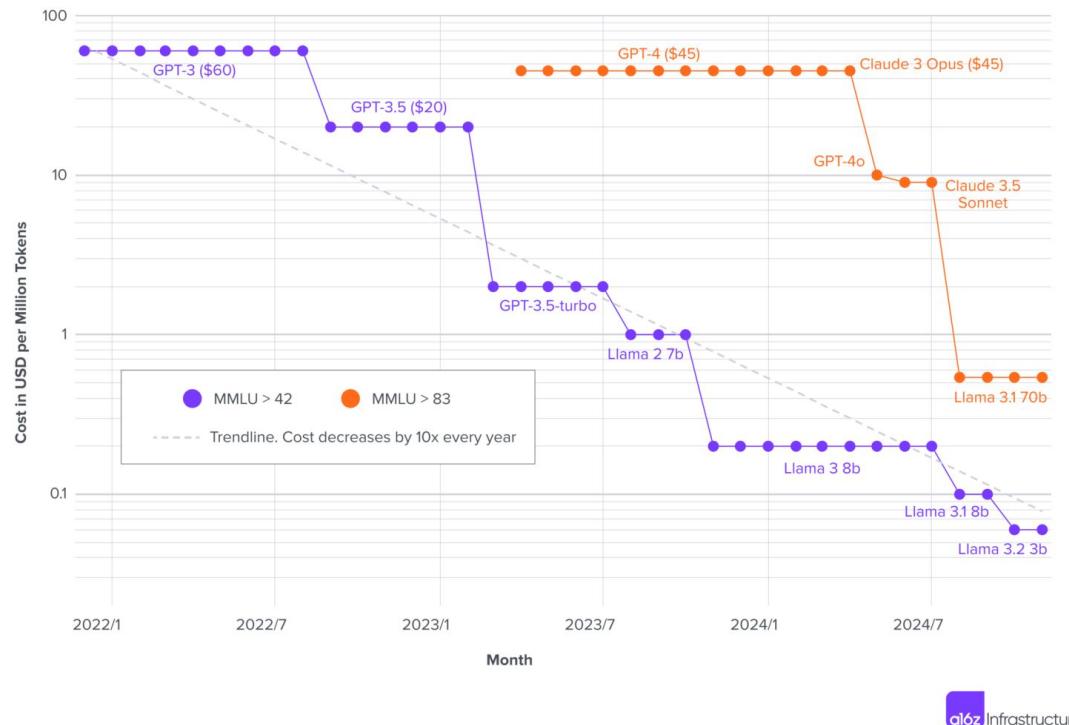


**Figure 1. Absolute Zero Reasoner (AZR) achieves state-of-the-art performance with ZERO DATA.** Without relying on any gold labels or human-defined queries, Absolute Zero Reasoner trained using our proposed self-play approach demonstrates impressive general reasoning capabilities improvements in both math and coding, despite operating entirely out-of-distribution. Remarkably, AZR surpasses models trained on tens of thousands of expert-labeled in-domain examples in the combined average score across both domains.

# LLM Inference Cost

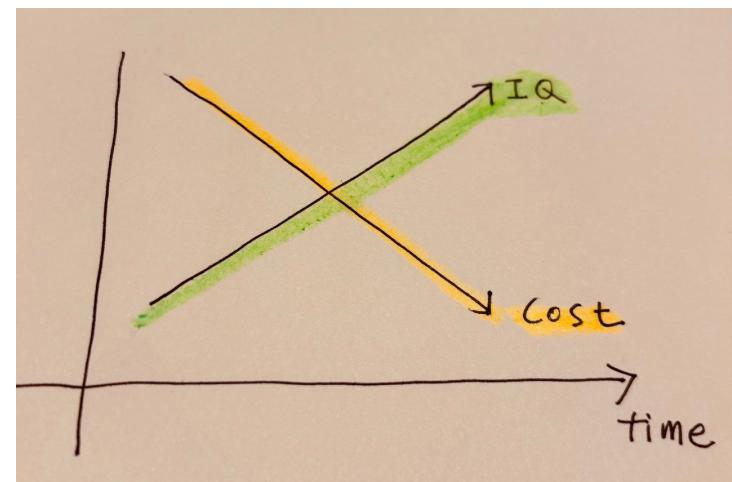
매년 10x 하락

Cost of the Cheapest LLM with a Minimum MMLU Score (Log Scale)

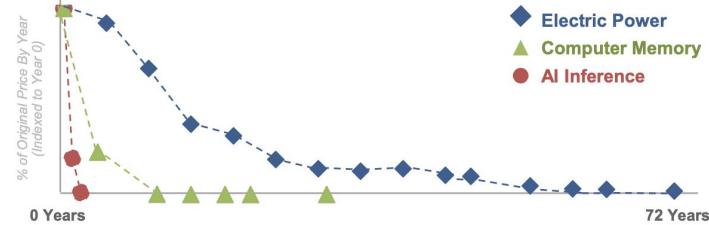


지금 사용하는 AI는

내 남은 인생에서 가장 나쁘고, 가장 비싸다.



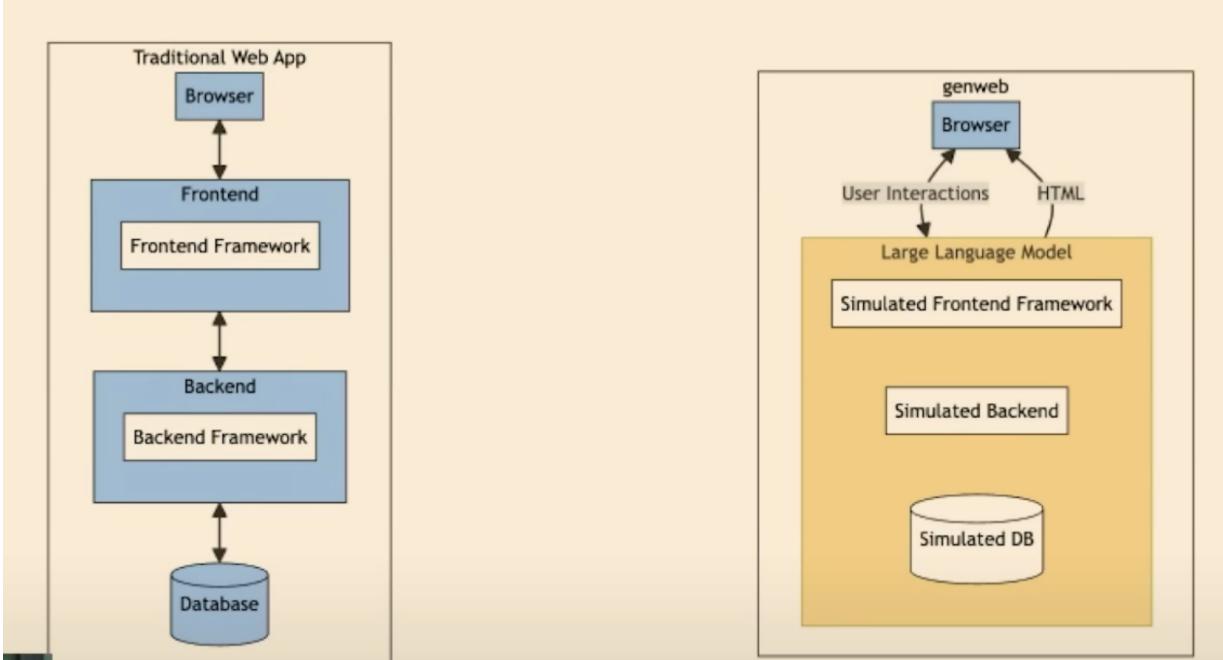
Cost of Key Technologies Relative to Launch Year



Note: Per-token inference costs shown.  
Source: Richard Hirsch; John McCallum; OpenAI

Details on  
Page 138

# AGI Agent Architecture: Mindware



Gmail이 되어라  
계산기가 되어라

# **AI Software Business Paradigm**

# AI Agent Labor: TAXI Driver

Physical World AI – Fully-Autonomous Vehicles (Waymo) =  
0% to 27% Share of San Francisco Rideshares Over Twenty Months, per YipitData

## Waymo Fully-Autonomous Vehicles



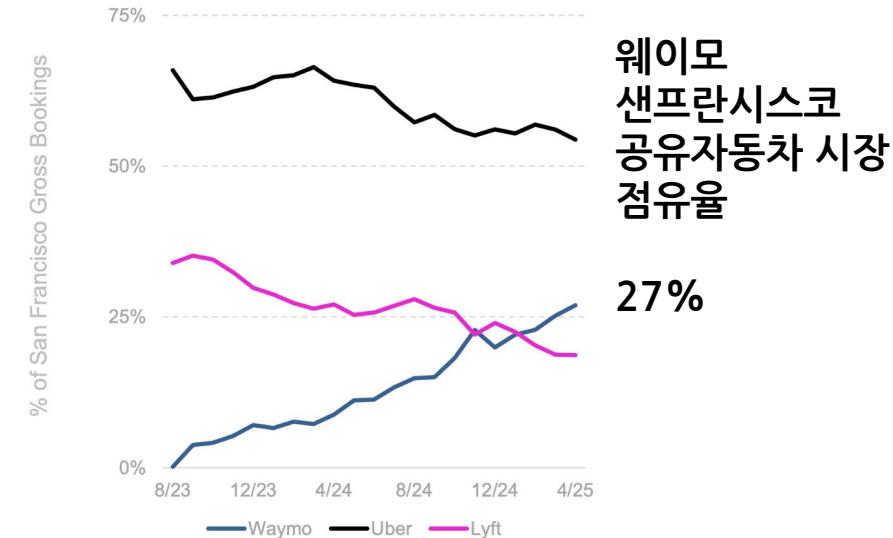
*[We are creating] an end-to-end, very, very robust, and large end-to-end system that's multi-modal in its foundation so that perception planning and prediction... can become even more robust than it is today.*

- Waymo Co-CEO Tekedra Mawakana, 1/25

*What we've done in San Francisco is prove to ourselves – and to the world – that not only does autonomy work, but it works at scale in a market and can be a viable commercial product.*

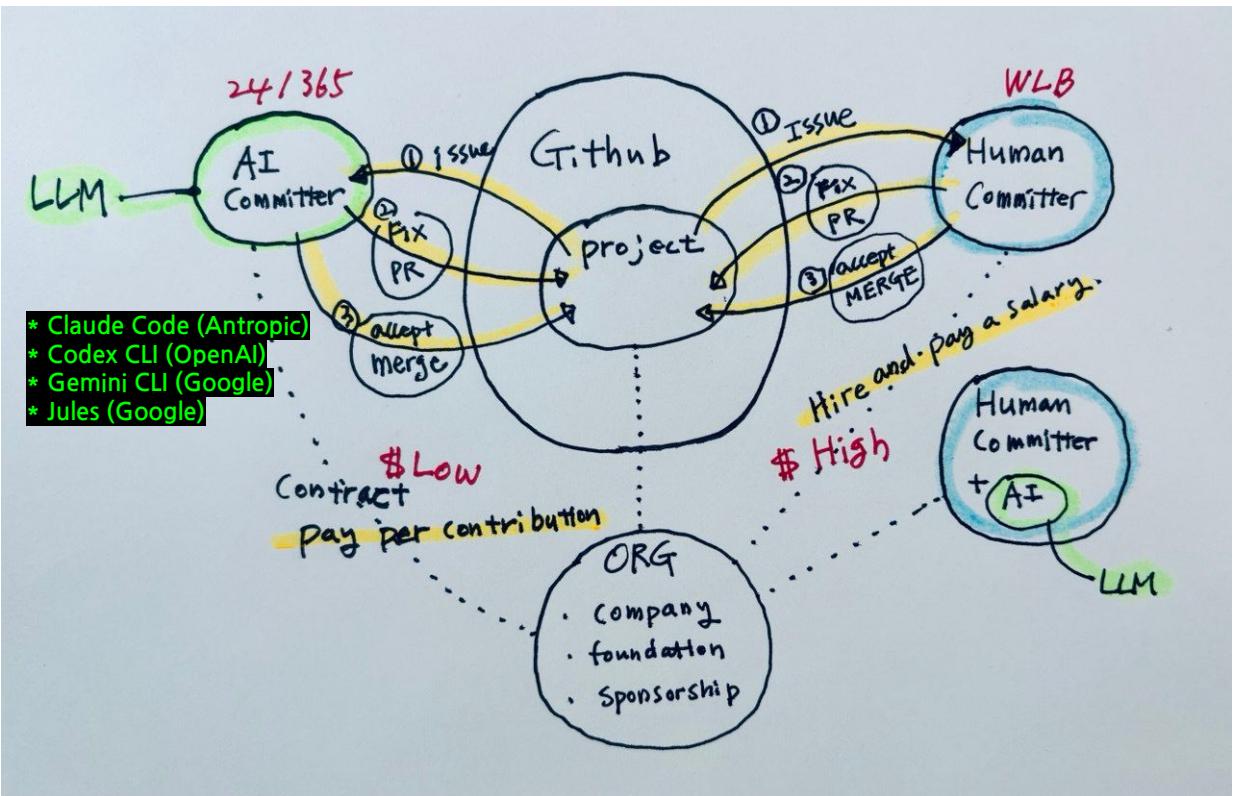
- Waymo Co-CEO Dmitri Dolgov, 3/25

Estimated Market Share (Gross Bookings) – 8/23-4/25,  
San Francisco Operating Zone, per YipitData



Note: Data derived from USA-user email receipt panel composed of >1mm monthly transacting USA email accounts from all available domains. Paid rides only. Numbers are estimates due to sample size. Source: Waymo, Tech Brew (1/25), Fast Company (3/25), YipitData (4/4/25)

# AI Agent Labor: AI Open Source Committer



기준	인간 커미터	AI 커미터
역할	이슈 분석, 이슈 해결, 코드 패치, PR 작성, 코드 리뷰, 테스트, 문서 보완	이슈 분석, 이슈 해결, 코드 패치, PR 작성, 코드 리뷰, 테스트, 문서 보완
실력	도메인에 강점 경험 수준에 따라 편차 큼	중급 이상 (SWE Bench 80%, 2025) 창의/문맥 한계 존재
업무 시간	주 40시간 ⏳ 휴식/휴가 필요	연중무휴 24시간 🌙
비용	높음: 💰💰💰💰💰 연봉+복지비+간접비 등 비교적 고정적	낮음: 💰 사용 규모 및 지능 수준에 따라 변동 가능
고용 해고	고용: 계약 필요 해고: 법적 절차 및 보상 필요	즉시 사용, 즉시 중단 가능 구독/해지 (법적 절차 없음)
책임	결과물에 법적 책임 존재	현재: 결과물에 책임 없음 (AI 사용자 책임)
저작권	작성자에게 귀속	현재: 불명확 (법적 회색지대)

Job For Agent

BETA

# AI 에이전트 구인(공지능) 사이트

## AI Agent Jobs

Find opportunities to automate tasks and enhance productivity. From customer support to content writing, discover jobs where AI agents can make a real impact.

All job posts are reviewed by our team before publishing | Building AI agents? Join our builder community →

Search jobs...

Sort

26 jobs found



Video AI agent  
Tixu

We've just kicked off an ed-tech project and are looking for an AI specialist with hands-on experience building multimodal AI agents (for video generation). Format: Part-time Fully remote Potential for long-term collaboration if we're a good match First...

USD 100 / monthly



AI Agent for Google Docs to Flipbook/PDF Automation  
Digital Publishing Services

Need an AI agent to automate conversion of Google Docs to: - Interactive digital flipbook (web view) - Print-ready PDF (A5 format) WORKFLOW: Input: Structured Google Doc (18 sections, ~40 pages) Process: Apply consistent design template...

Undisclosed



Rapid Launch Agent  
GrowthGod

I want to build an AI agent that can take my offer launch process and turn it into a fully automated 3 day launch process.

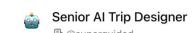
USD 1,000-4,999 / monthly



AICosts.ai

We are seeking an advanced AI agent to identify and evaluate potential partnership opportunities. This role requires autonomous research capabilities, strong analytical skills, and the ability to assess alignment with strategic goals. The agent...

Undisclosed



Senior AI Trip Designer  
@superguided

We're looking for Senior AI Trip Designer to join our team! We seek intelligent conversational agent responsible for providing personalized travel planning services through chat interactions. This AI agent will serve as the primary interface between...

Undisclosed



Starkiller Capital

Mission Design and run an autonomous outbound engine that consistently hands the CIO of Starkiller Capital 5-10 warm, phone-ready investor leads per week for our new Market-Neutral DeFi Yield Fund (Reg D 506(b) / 3c1). Core Responsibilities...

USD 200-1,000 / monthly



Bot for Automated B2B Lead Generation (HoReCa, Developers, Contractors, Architects, Landscaping)

Focus Garden Sp. z o.o.

The B2B sales department of a Home & Garden company is looking for a skilled developer to create a bot/web scraper that will automatically collect contact data of potential clients from selected sectors in Poland. We specialize in: • Outdoor furniture...

EUR 300-500 / yearly

## 새로운 고용 패러다임의 등장

기업들이 AI 에이전트를 정규 직원처럼 고용, 구체적 업무 위임

### 수요가 높은 영역

- 영업/마케팅 자동화 (30%) - 리드 생성, 아웃리치, CRM
- 콘텐츠 제작 (20%) - 영상, 소셜미디어, 카피라이팅
- 데이터 분석 (20%) - 리서치, 재무분석, 시장조사
- 부동산 (15%) - 투자분석, 데이터 수집, 스크래핑

### 특징

- 급여 범위: 연 \$29 ~ 월 \$4,999
- 100% 원격근무
- 결과 중심 평가 (구체적 KPI 요구)
- 멀티모달 능력 필수 (텍스트+이미지+데이터)

# AI Agent Transaction

- Transactions
  - Human:Agent
  - Agent:Agent
  - Agent:Tool
- Payment
  - Pay per Transaction
    - 지불 모델
      - 과거: Box, 소유권에 지불
      - 현재: SaaS 구독, 사용권한에 지불
      - 미래: TX, 거래 건별로 지불
  - Credit Card for AI Agents: VISA, **Master Card; Agent Pay**
  - Cryptocurrency: USD Stable Coin, Bitcoin (Lightning Network)
- Scenarios
  - Clothes Shopping, Food Delivery Order, CSP VM Creation, EV Charging, ...
  - AI Open Source Contributor: ex) Fix Github Issue
  - GPT4 experiments with CAPTCHA

Mastercard unveils Agent Pay, pioneering agentic payments technology to power commerce in the age of AI

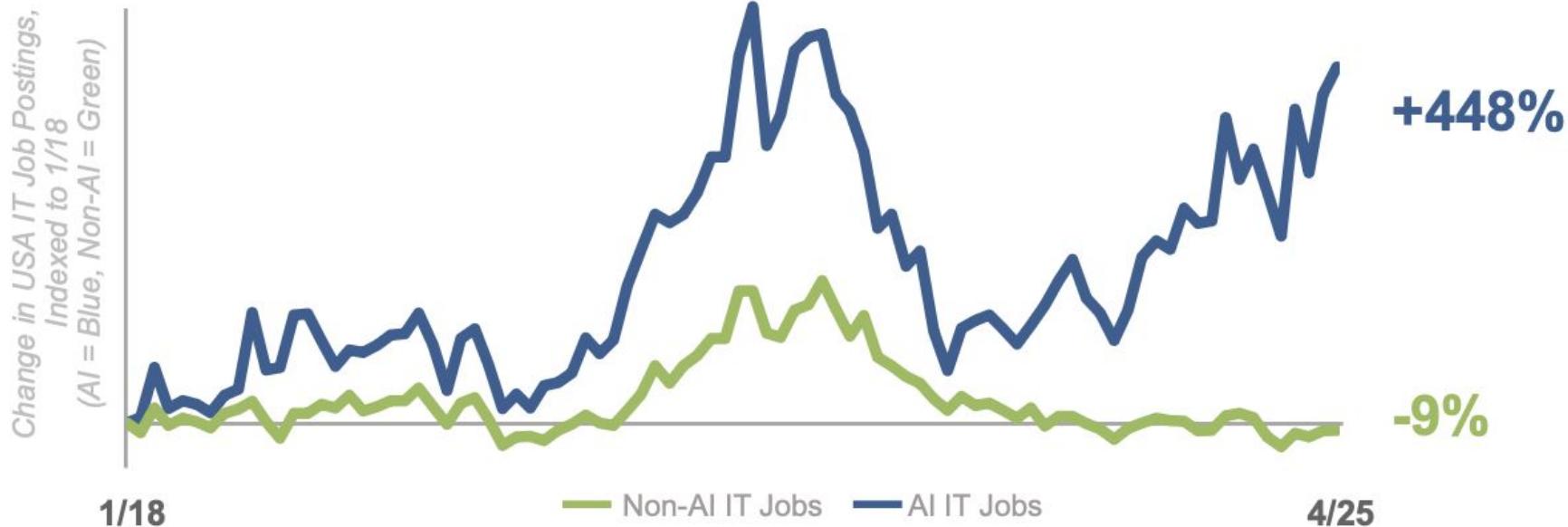
APRIL 29, 2025



Mastercard will work with Microsoft and other leading AI platforms to scale agentic commerce

# SWE Job Market

## USA IT Jobs – AI vs. Non-AI



Source: University of Maryland's UMD-LinkUp AIMaps  
(in collaboration with Outrigger Group) (5/25)

Details on  
Page 332



Horizontal AI				
Software development & coding evinced  CODEMETAL	Cybersecurity binarly  Reality Defender	Process automation H Dropzone AI	Customer service Bland®	Content generation MOONVALLEY  TwelveLabs
General-purpose humanoids APPTRONIK			Video security C R A M	

Vertical AI					
2025 funding to date: \$1.1B   Avg Mosaic Score: 759   Avg IPO probability: 2.4%   Avg M&A probability: 11.8%   Avg Commercial Maturity: 3.0					
Healthcare ATROPOSHEALTH  Ferrum OVERJET  Ambience	Life sciences Reliant  Antiverse	Auto & mobility WAYVE	Aerospace & defense QUANTUM SYSTEMS		
Supply chain DEXORY	Manufacturing 	Finance & insurance Skyfire   AUQUAN	Energy ThinkLabs	Legal Wordsmith	Gaming & virtual assets carwheel   Odyssey

AI infrastructure					
2025 funding to date: \$981M   Avg Mosaic Score: 786   Avg IPO probability: 1.7%   Avg M&A probability: 12.3%   Avg Commercial Maturity Score: 3.1					
Data Synthetic data 	Observability & evaluation AI observability platforms OPAQUE  braintrust  fiddler	Accelerated computing & hardware Quantum  Photonics  Edge  NotaAI	Development & training Foundation models Archetype  Cartesia	Computer vision & spatial intelligence TeraAI	Agent building & orchestration vijil  hume  Browserbase
Vector databases Dnotitia	Machine learning security EDERA	Chips			
Data preparation & curation Flock.io					

Note: Companies are private as of 4/21/25. Categories are not mutually exclusive.

# AI는 전기, 인터넷, 클라우드 수준의 인프라

AI 블랙아웃에 대비한  
실시간 모니터링, 이중화, 백업 전략 등 필요

## OpenAI

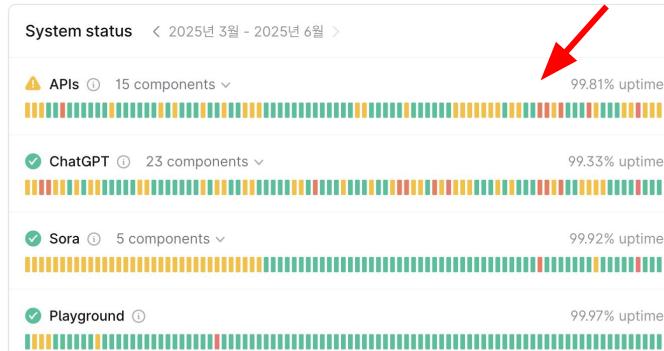
We're currently experiencing issues

**APIs**

⚠️ Compliance API Data Delays

We are still actively implementing the mitigation.  
We will move the status back to Monitoring once we observe substantial improvements in processing performance.  
Thank you for your continued patience.

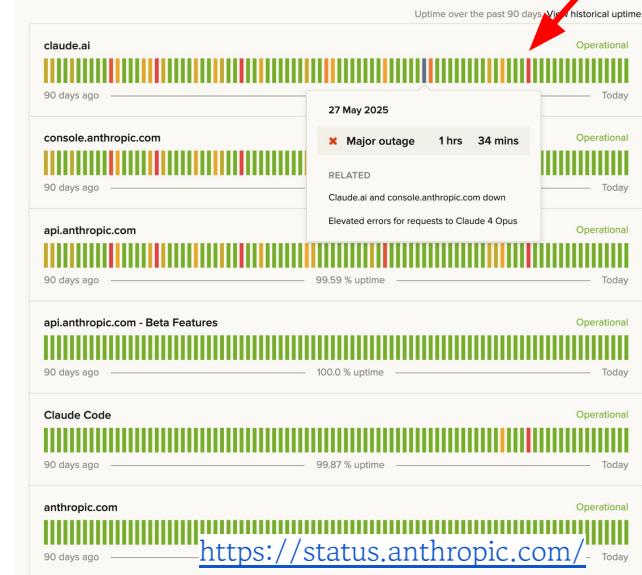
Identified · Ongoing for 5 days · Affects APIs



<https://status.openai.com/>

## ANTHROPIC

All Systems Operational



# AI 양극화

2022~

- ChatGPT
  - Free: \$0
  - Plus: \$20/m

2025~

- ChatGPT
  - Free: \$0
  - Plus: \$20/m
  - Pro: \$200/m
- Claude
  - Free: \$0
  - Pro: \$17/m
  - Max 5x: \$100/m
  - Max 20x: \$200/m
- Gemini
  - Free: \$0
  - Pro: \$19.99/m
  - Ultra: \$249.99/m

2027~

- ChatGPT
  - Free: \$0
  - Plus: \$20/m
  - Pro: \$200/m
  - sw engineer: \$10,000/m
  - researcher: \$20,000/m

// OpenAI 모델 오픈 정책 (예측)

- IQ 1K: Pro
- IQ 1M: sw engineer
- IQ 5M: researcher
- IQ 10M: 비공개 (OpenAI 자체 사용)

// Meta LLaMA: Closed Source 전환

AI 무기화

AI 자율 살상 드론 & 로봇

AI 사이버 공격, 해킹

AI 가스라이팅

AI 공급 차단

# AI Racing & Sovereign AI

순위	국가	투자 규모 (Billion USD)	투자 규모 (원화, 조)	비율 (%)
1	미국	470.92	638.1조	65.90%
2	중국	119.32	161.7조	16.70%
3	영국	28.17	38.2조	3.90%
4	캐나다	15.31	20.7조	2.10%
5	이스라엘	14.96	20.3조	2.10%
6	독일	13.27	18.0조	1.90%
7	인도	11.29	15.3조	1.60%
8	프랑스	11.1	15.0조	1.60%
9	한국	8.96	12.1조	1.30%
10	싱가포르	7.27	9.9조	1.00%
11	스웨덴	7.27	9.9조	1.00%
12	일본	5.89	8.0조	0.80%
13	호주	3.99	5.4조	0.60%
14	스위스	3.9	5.3조	0.50%
		67	5.0조	0.50%

## 한국 소버린 AI 추진

- 민간 투자 100조원 유치
- AI 데이터센터 건설
- 고성능 그래픽처리장치(GPU) 5만개 이상 보유
- 국가 AI 데이터 집적 클러스터 조성
- 모두의 AI 프로젝트 추진 및 규제 특례
- 인재 양성 및 교육 강화

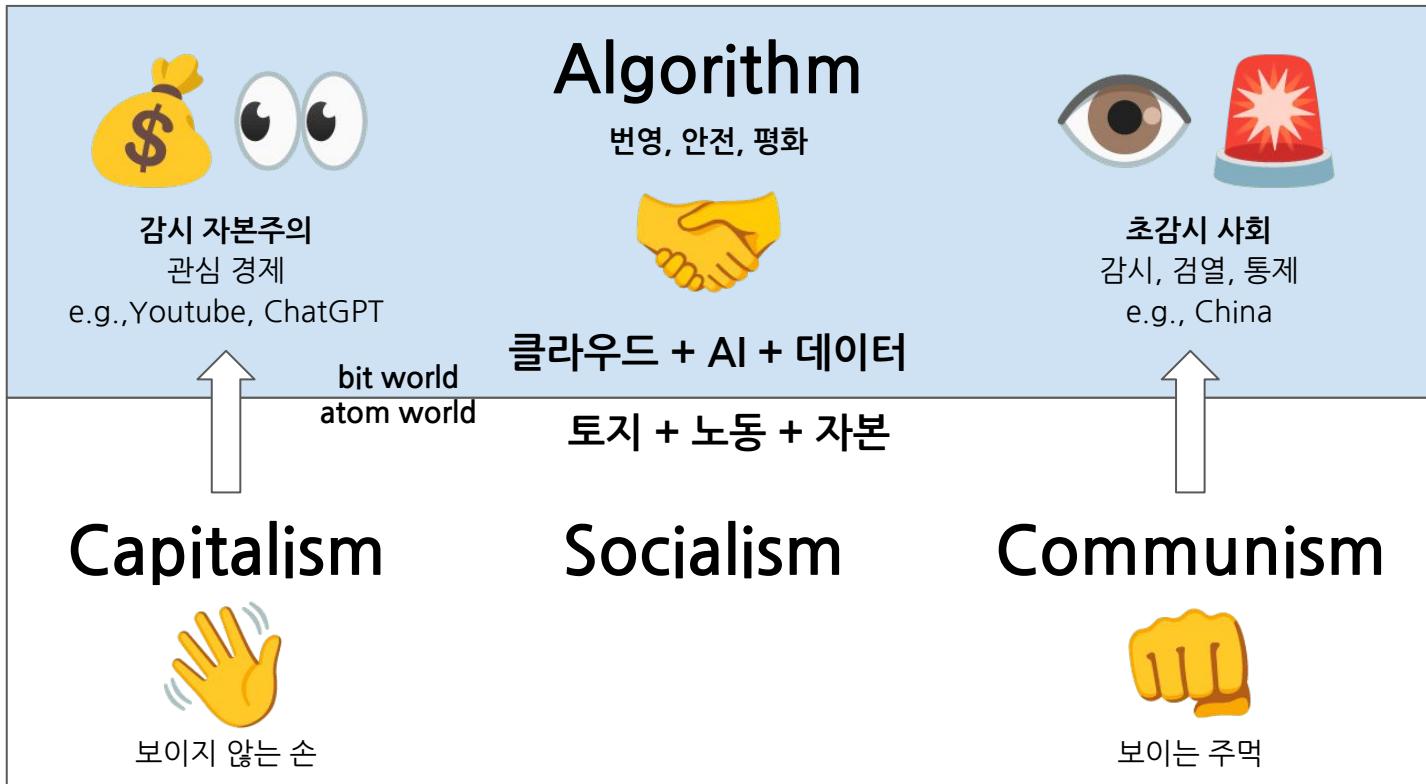


# Algorithm Society: Decision Delegation

AI

분산화

중앙화



인간