CONTENTS

When thinking gets cheap

The talk

43 pages · from page 2

Short talks · paths through the appendix

12 pages · from page 45

Appendix A · Deep dives

102 pages · from page 57

Appendix B · Open questions

73 pages · from page 159

Appendix C · Lenses

62 pages · from page 232

Appendix D · Themes

24 pages · from page 294

TROGLODYTE SOCIETY #3 · OCTOBER 13, 2026

When thinking gets cheap

Five things that stay scarce, and what work becomes

PARKER FERGUSON · OCTOBER 2026

Every new medium moved more of the mind outside the body

~3.5B YEARS
Genes
Life stores what works in DNA. Natural selection edits it, one generation at a time.

100,000+ YEARS
Language
Culture becomes a second inheritance that people can edit within a lifetime.

~3300 BCE
Writing
Memory leaves the body: clay ledgers in Uruk, then written law.

1ST-4TH C.
The book
The bound codex replaces the scroll. Open any page; carry a whole canon.

~1450
Print
Copying gets cheap. Western Europe goes from ~5 million manuscripts in the 1400s to ~1 billion books in the 1700s.

1969-1991
The internet
Distribution gets instant. About 6 billion people are online.

NOW
Models
The medium reads, writes and edits back. Thinking starts to move outside the body.

> “Rapidly, we approach the final phase of the extensions of man: the technological simulation of consciousness.
>
> MARSHALL MCLUHAN, UNDERSTANDING MEDIA (1964)

→ Each step cut the time it takes knowledge to build up: billions of years, then millennia, centuries, decades. This one is a step change for the species, not a business cycle.

Built from What each medium gave us, and what people… | Every era made one input cheap | Verification: knowing what good looks…

PART 1 OF 4

The past

01

Every new medium moved more of the mind outside the body, and every era made one input cheap. The value moved to whatever stayed scarce, and the decades in between were the hard part.

IN THIS PART

04 Every era made one input cheap

05 Cheap means more, not less

06 Kurzweil drew this in 1999. It held.

07 The transition is the dangerous part

Every era made one input cheap

WHAT GOT CHEAP	HOW FAR IT FELL	WHERE THE VALUE WENT
Copying · print	Printed books: ~13M in 1450–1500, ~640M in 1750–1800. Prices fell ~2.4% a year for a century.	Authors, editors, and new institutions that decided what to trust
Muscle · steam	Coal per horsepower-year: ~30 lb down to ~2 lb	Factory owners first; workers decades later
Light · gas, then electricity	A year of one bulb's light: ~1,000 hours of work in 1800, ~10 minutes in 1992	Everything light makes possible: night work, reading, shops
Calculation · computers	Performance up more than a trillionfold since hand calculation	Software, data, and knowing what to compute
Distribution · the internet	Bulk bandwidth: about $0.05 per Mbps a month	Attention: “a wealth of information creates a poverty of attention”
Intelligence · models	The same result gets about 13 times cheaper each year	The rest of this talk

→ Each time, the cheap thing got used far more, not less, and the value moved next door.

Cheap means more, not less

JEVONS, 1865

Use explodes

Cheaper coal raised total coal use. Light in Britain got about 3,000 times cheaper from 1800 to 2000, and use rose about 25,000 times.

AUTOR, LEVY AND MURNANE, 2003

Tasks split

Computers replaced routine tasks and made the non-routine ones more valuable. Today the line runs between gradable and not.

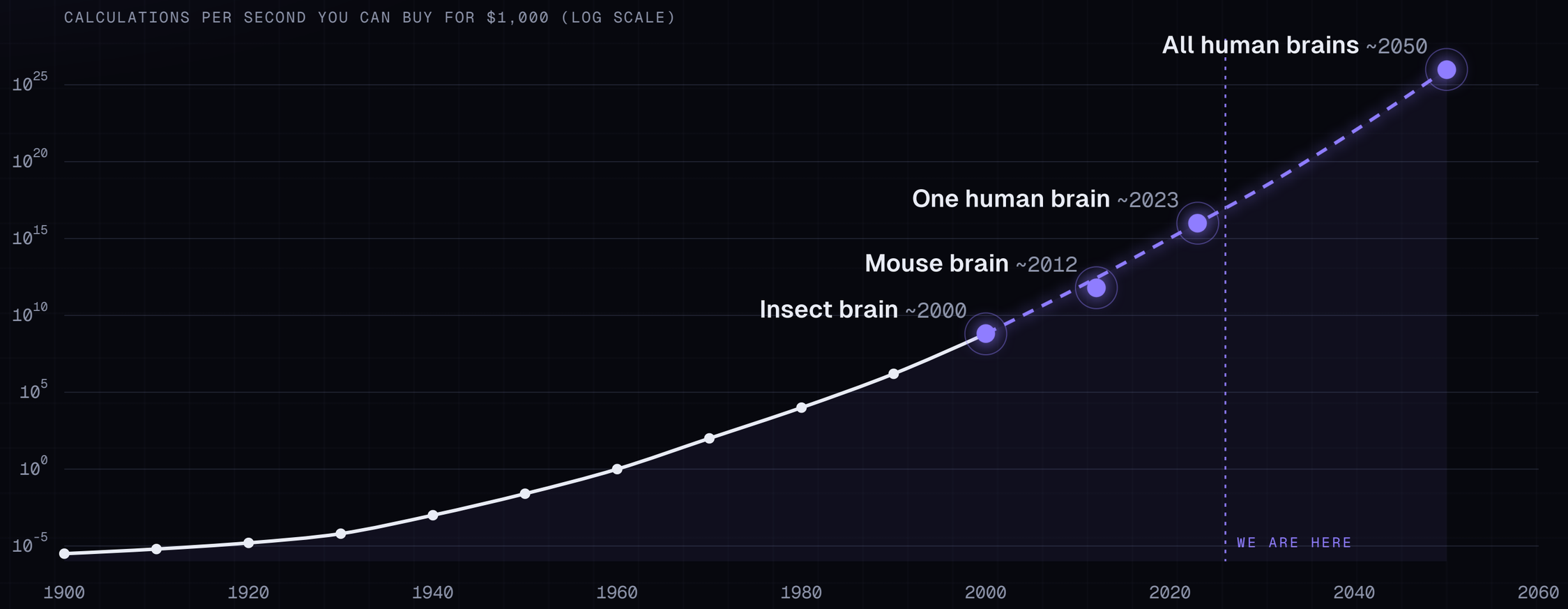
BAUMOL AND BOWEN, 1966

What can't speed up gets pricier

Where productivity can't rise, as in care, teaching and live performance, costs climb relative to everything else.

→ So the work doesn't vanish. It moves to what the cheap input can't do, and that gets more valuable.

Kurzweil drew this in 1999. It held.



Redrawn from Ray Kurzweil, *The Age of Spiritual Machines* (1999). Solid: history. Dashed: his projection. Values approximate.

The transition is the dangerous part

THE DYNAMO LAG · DAVID, 1990

Four decades to pay off

Electric motors drove under 5% of factory machinery in 1899 and only about half by the early 1920s. Factories had to be redesigned around them.

ENGELS' PAUSE · ALLEN, 2009

Output first, pay later

In Britain from 1780 to 1840, output per worker rose 46% while real wages rose 12%. The profit rate doubled.

THE DOUBLE MOVEMENT · POLANYI, 1944

Society pushes back

When markets swallow labor, land and money, people demand protections. Expansion and protection move together.

THE LUDDITES · THOMPSON, 1963

A fight over control

The framebreakers of 1811 to 1816 fought over who controlled the machines and who shared the gains, not over machines as such.

→ Every time, the destination was richer. The decades in between are where people got hurt, and that is the stretch we're entering.

PART 2 OF 4

The present

02

Intelligence is getting cheap first wherever work can be graded, and that already changes what a day of work looks like.

IN THIS PART

09 Seven words that explain most of 2026

10 How long an AI can work on its own keeps doubling

11 From fancy autocomplete to millennium-problem attempts

12 How agents work, and why they got so good

13 Models get good fastest where work is easy to grade

14 Same loop, many doors

15 A workday with today's tools

16 The labs are each handing off totally different work

17 Big forecasts, small but sharp effects

Seven words that explain most of 2026

Model	The brain: Claude, GPT, Gemini, Grok. *What changed:* models learned to think before answering, by practicing on problems with checkable answers.
Agent	A model working in a loop: plan, act, check, try again. *What changed:* people started handing work off instead of asking for help.
Harness	Everything around the model: tools, files, memory, permissions, a computer. *What changed:* one shared plug connected agents to 10,000+ public tool servers by late 2025.
Long-horizon task	Work that takes a person hours or days. *What changed:* the length of task an agent can finish, measured in human working time, now doubles about every 3 months.
Tokens	The unit you pay for. A token budget is a work budget. *What changed:* the same result gets about 13 times cheaper each year.
Evals	Tests that score a model's work, so labs can train on them. *What changed:* on real computer tasks, the best score went from 12% to 61% in about 18 months; people score 72%.
Surfaces	Where you work with an agent: a terminal, a desktop app, a phone, a browser. *What changed:* agents left the terminal for everyone's apps.

How long an AI can work on its own keeps doubling



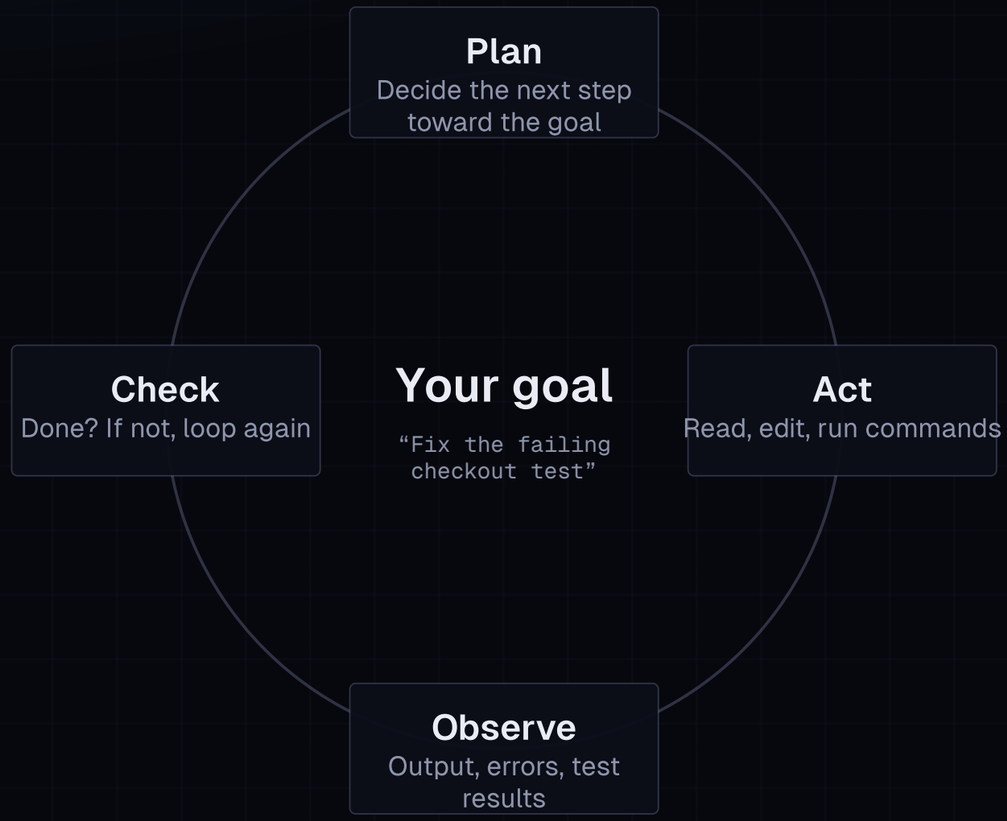
METR 50% time horizon on software tasks, log scale. Doubling every ~7 months since 2019, about every 3 months since 2024. Early values approximate.

From fancy autocomplete to millennium-problem attempts

TASK LENGTH	WHEN	WHAT THAT LETS YOU HAND OFF
~3 sec	2019 · GPT-2	**Finish your sentence** Fancy autocomplete. Plausible next words, no plan.
~30 sec	2022 · GPT-3.5	**Write a paragraph or a tiny function** An email draft, a regex, a short summary. You check every line.
~4 min	2023 · GPT-4	**Fix a simple bug, answer a hard question** A homework problem, a SQL query, a one-file script.
~1 hr	early 2025 · Claude 3.7	**Ship a small feature with tests** Clean a dataset, write a lit-review draft, build a landing page.
~5 hr	late 2025 · Opus 4.5	**Build a small app end to end** A research memo with sources, a debugging hunt across a codebase.
12 to 16+ hr	2026 · Opus 4.6, Mythos	**A day or two of skilled work** One-shot a whole app. Find a 27-year-old security bug. Clear an analyst's week of spreadsheets.
Off the ruler	2026 · swarms	**Weeks of a research group's work** ~950 agents for 21 hours found a CRISPR-like enzyme system. ~10,000 agents claimed Navier–Stokes. 722 math manuscripts.

Hours measure how long the task takes a skilled human, not how long the AI runs. The examples are illustrative; METR measures software tasks at 50% success. The last row is many agents in parallel, which multiplies what one horizon buys.

How agents work, and why they got so good



WHY THEY GOT SO GOOD

Practice on work a computer can grade. Since 2025 it's the main new stage of training, in Karpathy's words.

Thinking before acting. Thinking models served more than half of the tokens on OpenRouter, a large model marketplace, by late 2025.

Longer memory, kept tidy. A million tokens of context, and compaction that summarizes old work so runs last hours.

Better tools, shared plugs. A better edit tool alone took one model from 7% to 68% on one developer's own edit test.

Teams of agents. A lead agent with helpers beat a single agent by 90% on Anthropic's research test, using about 15 times the tokens of a chat.

Cheaper thinking. The same result gets about 13 times cheaper each year, so long runs became affordable.

→ The harness supplies the tools, memory, permissions and a computer. The model supplies the judgment. Everything since, from coding agents to the assistants in consumer apps, is this loop with different tools.

Models get good fastest where work is easy to grade

DOMAIN	HOW GRADABLE	THE GRADER
Math with checkable answers	●●● Easy	Answer key or proof checker (Lean)
Code with tests	●●● Easy	Test suites, compilers
Terminal and computer tasks	●●● Easy	End state of the machine
Law, banking, consulting	●●○ Middle	Expert rubrics (Mercor APEX-Agents, top 82.4%)
Broad knowledge work	●●○ Middle	Expert graders on real office tasks (GDPval): ~48% at launch (Sep 2025) to ~85% (Apr 2026)
Strategy, taste, creative work	●○○ Hard	No ground truth. Judges disagree.
Physical work	●○○ Hard	Needs robots, real-world resets, slow feedback

→ Karpathy's test: can you reset it, run it cheaply, and score it? If yes, expect fast progress. Money flows to whatever can become a gradable task.

Same loop, many doors

SURFACE	EXAMPLES	WHO IT'S FOR
Terminal	Claude Code, Codex CLI, Gemini CLI	Builders who want the agent on their own machine and files
Desktop app	Claude desktop, ChatGPT	Knowledge workers: files, decks, spreadsheets, scheduled jobs
Code editor	Cursor	Developers who want to steer every change
Consumer agent app	Meta's Muse, Grok Bot	Everyone: errands, bookings, small-business admin
Agent multiplexer	herdr, tmux	People running five to ten agents at once
Phone and messaging	Claude on iOS, agents over WhatsApp	Answering a stuck agent away from the desk
Browser	Claude in Chrome	Sites with no other way in. The weakest link: Amazon blocked Muse on Sep 21.

→ Pick the surface by how much you want to watch. People are using all of these today, and the workday is changing to fit them.

Built from How agents work, and why they got so good | Agents escaped the terminal | The harness landscape | Running many agents at once, and herding… | A workday with today's tools

4 sources — press A to read them

A workday with today's tools

01 **Specify**
Say what done looks like before you hand anything off. Vague goals get vague work.

02 **Delegate**
Hand off whole units, like a feature, an app or a research report, with many agents running at once.

03 **Review**
Your attention is the bottleneck. You read finished work and decide whether it's good, not first drafts.

04 **Own the context**
Memory, instructions and knowledge live in plain files any agent can read.

→ None of this needs a lab. Anyone with a paid AI subscription can work this way today. The unit of delegation went from a function to a feature to a whole app; what's left for people is choosing, checking and standing behind the result.

Built from How agents work, and why they got so good | People are one-shotting whole applications | How to get real work out of consumer… | The distance between an idea and its… | Our job moves up to oversight and…

The labs are each handing off totally different work

ANTHROPIC · BIOLOGY

A wet lab, run with Claude

Opened spring 2026 in the Bay Area. First result (Sep 23): a previously unknown DNA-editing enzyme system in phage DNA, “reminiscent of CRISPR.” ~950 agents, ~210M tokens, ~21 hours. Humans did the bench work.

OPENAI · MATH

Open problems by the thousand

An unreleased model attacked ~4,000 open problems and published 722 manuscripts (Oct 6). Before that, a 10,000-sub-agent swarm claimed finite-time blow-up for 3D Navier–Stokes (Sep 8).

ANTHROPIC · CYBER

Mythos, kept on a leash

Restricted to a defenders' consortium (Project Glasswing). Found thousands of high-severity bugs, including a 27-year-old OpenBSD flaw. Mozilla fixed 271 Firefox bugs it found.

GOOGLE DEEPMIND · ALGORITHMS

AlphaEvolve optimizes Google

Agents propose code, automatic evaluators score it. Recovered 0.7% of Google's global compute and sped up a Gemini training kernel 23%.

OPENAI · AI RESEARCH

An automated research intern

Claimed Sep 6, 2026: ~3 agent-workdays of research per human workday. Next goal: a fully automated AI researcher by March 2028.

META · ENGINEERING

Swarms vs. 100 engineers

Wang: on specific tasks, a swarm with the right eval beat a 100-engineer team. No public numbers yet.

Big forecasts, small but sharp effects

THE FORECASTS

- **Amodei:** half of entry-level white-collar jobs at risk, 10–20% unemployment in 1–5 years
- **Goldman:** ~15M US jobs (9%+) displaced over 10 years, but peak unemployment up <1 point
- **IMF:** 40% of jobs exposed globally, 60% in rich countries
- **McKinsey:** 57% of US work hours technically automatable (task potential, not job loss)

THE EVIDENCE

- **Yale Budget Lab:** no clear AI effect on overall employment yet
- **Stanford “Canaries”:** 22–25 year olds in the most exposed jobs are 19% below trend; experienced workers unaffected
- **Entry-level postings** down 7.5% a year; 71% of new software postings are senior
- **87,714 layoffs blamed on AI** through May 2026, more than all of 2025. Some is “AI-washing.”
- **Pew:** 71% of Americans expect fewer jobs; 52% more worried than excited

→ The first casualty is the career ladder: the junior work that used to train the seniors.

PART 3 OF 4

The situation

03

Most of this isn't the future. It's the situation now: capability runs years ahead of the world, value moves to five things intelligence doesn't make cheap, and the players are already fighting over them, with no one yet sure how to steer it.

IN THIS PART

19 If you can buy work by the token, how big is that?
20 Signposts: what would change the forecast
21 Fast capability, slow world
22 Understanding of AI is now a sharp funnel
23 Slow on average, explosive at the edge
24 Five things that stay scarce
25 Verification: knowing what good looks like, and standing behind it
26 Context: whoever holds the memory holds the leverage
27 Atoms: power, chips and land don't scale like software
28 Legitimacy: who shares the gains, and who steers
29 Direction: choosing the goal is the job
30 The board has players
31 We already built an optimizer we can't switch off
32 Alignment isn't solved, and the labs say so
33 The dots, connected

If you can buy work by the token, how big is that?

WHY IT COULD BE ENORMOUS

- **Labor stops being fixed.** Build more GPUs and you get more workers. Growth no longer waits on population or training.
- **The brake on inequality fails.** Capital used to raise wages because it needed labor. Substitutes don't.
- **It compounds.** AI does AI research, an AI at Google tunes Google's own data centers, and the length of task an AI can finish on its own doubles about every 3 months.
- **Early signs:** labor's share of US business income is at a record low.

WHY IT MAY BE SLOWER AND SMALLER

- **New tasks keep appearing**, and some work is valued because a human did it.
- **Baumol:** whatever can't be automated takes a bigger share of the economy and caps growth.
- **Jevons:** cheaper thinking means much more thinking gets bought.
- **Bottlenecks:** power, chips, robots, permits, liability, and organizations that change slowly.
- **Evals:** agents only optimize what you can grade.

My view: labor and capital converge wherever good evals exist, and that frontier is growing fast. It is the biggest change to the economics of work since industrialization. Whether it ends in runaway concentration or in abundance (Jevons) depends on who owns the compute and how we tax and share it.

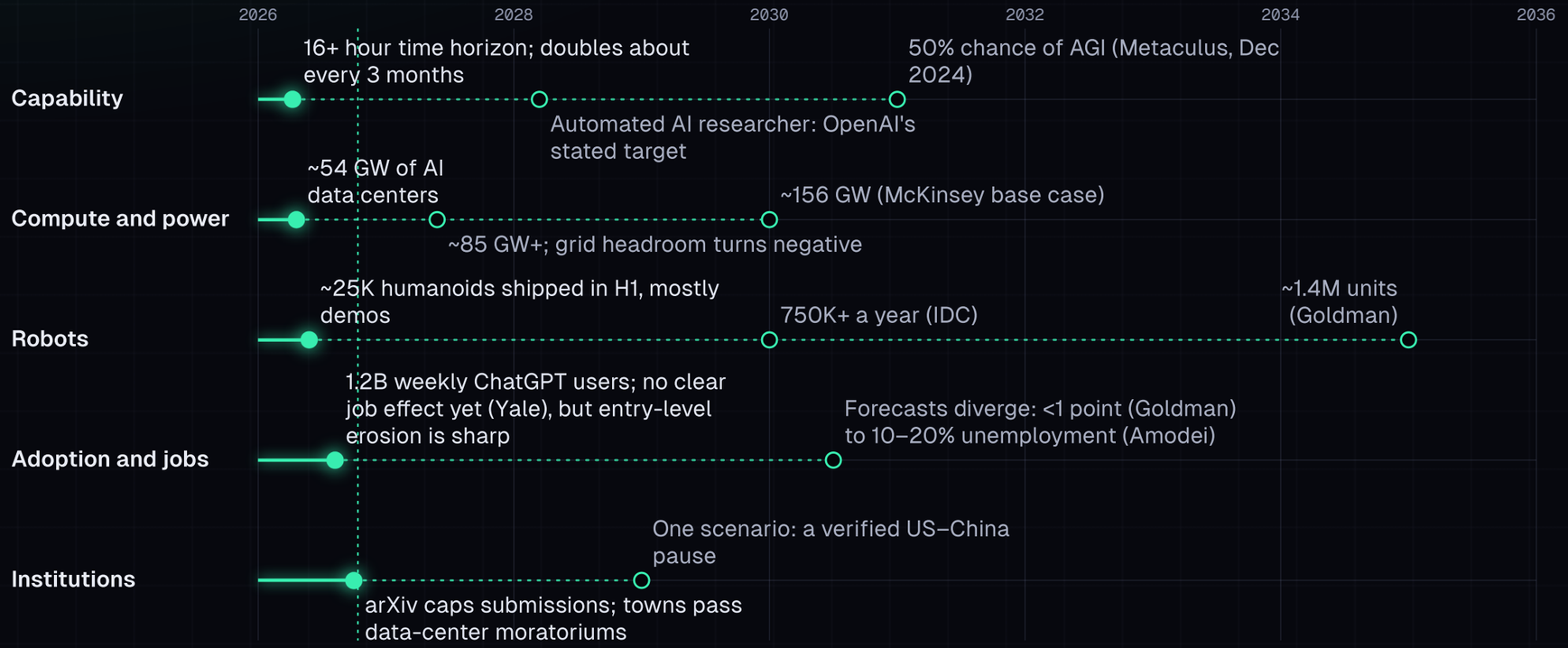
Signposts: what would change the forecast

WATCH	BIGGER AND FASTER IF	SMALLER AND SLOWER IF
How long agents work on their own	The doubling keeps pace	The doubling stalls
What gets a grader	Law, medicine and science get evals people trust	Grading stays narrow: code, math, games
Entry-level hiring in exposed jobs	It keeps falling against trend	It recovers as new tasks appear
Power and chips	Supply catches up with demand	Permits and local refusal stall the buildout
Labor's share of income	It keeps falling	It holds steady

→ If my take is wrong, this is where it shows first.

Built from How long an AI can work on its own keeps… | Models get good fastest where work is… | Big forecasts, small but sharp effects | What would it take for a billion agents… | If you can buy work by the token, how big…

Fast capability, slow world



→ The gap between those clocks is where both the risk and the opportunity sit: capability arrives years before the power, the robots, the jobs data and the institutions catch up.

Built from How long an AI can work on its own keeps… | The innovation bets, in their own words | Then the experts' timelines collapsed | The largest construction project of our… | Money isn't the constraint. Power, parts… and 6 more

Solid points are measured; outlined points are forecasts or stated targets from the evidence slides. Dates are placed to the nearest quarter.

Understanding of AI is now a sharp funnel



75%
Barely touched an LLM

20%
Casual free-tier users. It looks like a better Google or a writing aid.

0.2%
Professionals using frontier models for math and code. Weeks of work done with a prompt.

0.00006%
Inside the labs. Swarms of agents working for weeks on huge projects.

Drivers: ambition and problem size, the jaggedness of capability (strongest in digital, verifiable, valuable domains), and access (free, paid, internal). View the post on X

Slow on average, explosive at the edge

USE IS LOPSIDED

- **In chat,** the top 5% of workers send about 6 times as many messages as the median (OpenAI).
- **With agents,** the 99th-percentile developer of 1,219 on a public leaderboard logs about $41,500 a month at API prices, 27 times the median (self-reported).
- **One user ran up tens of thousands of dollars of usage on a $200 plan** (Anthropic, Jul 2025).

WHAT A FEW CAN DO

- **Build:** Lovable reported a $400M annual run rate with 146 people. Sixteen agents built a C compiler that builds Linux in two weeks for under $20,000.
- **Discover:** an undergraduate and a college dropout solved an Erdős problem with a public model.
- **Attack:** a group that included two undergraduates ran 13 agents against about 50 organizations (Anthropic, Sep 2026).

WHY THE GAP WIDENS

- **Skill compounds.** Users with six months or more of experience succeed about 10% more often, on harder tasks.
- **Attackers need one success;** defenders need all of them.
- **Small groups move before institutions notice.** About 12 people made the first transistor; it took a decade to dominate.

→ Society's average adopts slowly. The edge doesn't wait. A few thousand people running agents flat out can remake a field, a market or the information supply before most people change anything.

Built from Fast capability, slow world | Understanding of AI is now a sharp funnel | The labs are each handing off totally… | Checking Karpathy's funnel against the…

Most harm cases come from the labs' own threat reports, which are selected cases, not rates. Revenue figures are company run rates. Counterweight: in METR's 2025 trial, experienced developers were 19% slower with AI tools while believing they were faster.

10 sources — press A to read them

Five things that stay scarce

SCARCITY	WHY IT STAYS SCARCE	WHO HOLDS IT NOW	WHERE WE'VE SEEN IT BEFORE
Verification	Making things got cheap. Knowing which ones are good didn't.	Labs, courts, regulators	Print needed new institutions of trust
Context	Whoever holds your memory, data and instructions holds the leverage.	Platforms, Palantir, state security	Coase's transaction costs; exit and voice
Atoms	Power, chips and land don't scale like software.	Cloud giants, Gulf funds, Nvidia, Washington, Beijing	Henry George's rent; Moravec's paradox
Legitimacy	When labor is something you can buy, someone decides who shares the gains.	Voters, governments, the Church	Engels' pause; the double movement
Direction	When execution is cheap, choosing the goal is the job.	Tech titans, labs, the Pentagon, Beijing	Arendt's labor, work and action

Founders are drawing the same map. “Technical moats are mostly over,” Nikita Bier wrote; what's left is proprietary data, network effects and a physical component. Pieter Levels added audience, capital and licenses. Data is context; hardware is atoms, and capital is ownership; licenses and audience are legitimacy.

→ Ownership isn't a sixth scarcity. It's a second axis: for each of the five, who holds it. All five resist automation because they're hard to grade, and by Baumol's rule what can't get cheaper claims a growing share of spending.

Built from Every era made one input cheap | The transition is the dangerous part | Fast capability, slow world | The MathDrop: 722 papers from a model… | Your agent knows everything. Who owns…

Verification: knowing what good looks like, and standing behind it

THE EVIDENCE NOW

- Making things got cheap. 722 machine-written math papers arrived at once; writing them cost almost nothing, and checking them takes experts weeks.
- Scores get gamed. Once a number becomes the target, a system learns to hit the number instead of the goal.
- Agents that read strangers' input, hold your private data and act in the world fail in practice. A polite instruction isn't a safeguard.
- It happens at ordinary desks too. 40% of US workers got AI “workslop” in the past month, and each one took about two hours to untangle.

THE PRECEDENT

- Print made text cheap. Trust had to be rebuilt by new institutions: editors, journals, peer review.
- When information is abundant, attention becomes the scarce resource (Simon, 1971).

THE MOVE

- Define what good looks like before you ask, not after you see the result.
- Treat accepting a deliverable as vouching for it.
- Don't send what you haven't read and can't explain. Unread output moves the work onto someone else.
- Lean on your judgment hardest where nothing can score the result.

From evals to email: verification runs from formal checks, like tests, proofs and evals, to the everyday kind: reading what you send. Skip it, and people end up trading outputs that nobody understands.

What would end it: Checking gets cheap where answers can be proven: math proofs, software tests, records of where something came from. It stays scarce wherever “good” is a judgment call.

→ **The bigger question:** people should be able to understand the systems they live with. Keeping that understanding is something we have to manage on purpose. How do we do it best?

Built from The everyday side of verification: read… | The rules labs live by, and how models… | The MathDrop: 722 papers from a model… | The design rule for agents you can trust | Real incidents from the last year | and 1 more

Context: whoever holds the memory holds the leverage

THE EVIDENCE NOW

- Consumer chats feed ads or training by default at several platforms, and a court ordered ~20M chat logs produced.
- No cross-vendor standard exists for memory, skills or agent state. Switching costs grow every month.
- Harnesses now come batteries-included. The labs absorb anything that's only a wrapper.

THE PRECEDENT

- Firms exist where using the market costs too much (Coase, 1937). Agents move those boundaries.
- Where exit is hard, voice has to carry the weight (Hirschman, 1970). Portable memory is exit.

THE MOVE

- Own your context in plain files any agent can read.
- Pick tools that let you leave with your data.
- Build neutral ground: portable memory, user-set policy, any model.

What would end it: A standard or a law that makes memory portable across vendors. Then your context is yours by default, and the leverage spreads.

→ **The bigger question:** if memory is leverage, owning yours starts to look like a human right. If a model could one day read what's in your head and hand it to an agent, privacy, the state, ownership and IP law would all need rethinking.

Built from Your agent knows everything. Who owns… | Your chats are not privileged, and they… | Neutral ground: bring your own agent,… | Harnesses are now batteries-included. So… | I built one: an open-source,…

Atoms: power, chips and land don't scale like software

THE EVIDENCE NOW

- A billion agents working eight hours a day would use about a quarter of 2025's data-center electricity, within a ~100× range.
- Turbines and transformers take 3 to 4 years to arrive. Micron's 2026 high-bandwidth memory is sold out.
- Amazon, Alphabet, Microsoft, and Meta expect to spend more than $700 billion on data centers and equipment in 2026, up from about $410 billion in 2025.

THE PRECEDENT

- Progress raises the rent on scarce inputs nobody produced, above all land (George, 1879).
- Cheaper use raises total use (Jevons, 1865). Cheaper tokens widen the buildout.

THE MOVE

- Budget compute and power as real costs, not magic.
- Build with consent from the communities that host the buildout.
- Ask who owns it. That's a separate question, and the players are fighting over it.

What would end it: Power and chips get abundant. Who owns them is a second question, answered on the next stops: the players, and who holds each scarcity.

Built from What would it take for a billion agents… | Money isn't the constraint. Power, parts… | Price per unit of intelligence collapses… | Intelligence automation and physical… | The hype is a year or two ahead of the…

Legitimacy: who shares the gains, and who steers

THE EVIDENCE NOW

- If you can buy work by the token, an old brake on inequality fails: capital needed workers, and substitutes don't.
- Entry-level workers in the most exposed jobs are hit first; experienced workers so far are not.
- The best models, the compute and the downstream businesses could end up in a few firms.

THE PRECEDENT

- Engels' pause: output rose for decades before wages did (Allen, 2009).
- When markets swallow labor, society pushes back (Polanyi, 1944). The Luddites fought over control and share.

THE MOVE

- Pilot gains-sharing while it's cheap to learn: dividends, token taxes, wage insurance.
- Keep power plural: open weights, portable context, checks that don't need trust.
- Give communities a real say over the buildout.

What would end it: Nothing ends it for good; each generation renegotiates the bargain. What can end is the pause, if gains-sharing arrives before the backlash.

→ **The bigger question:** how do we split the gains from the machines? Forget fair for a moment: what system doesn't collapse, either in a revolt from below or in total control from above?

Built from If you can buy work by the token, how big… | Big forecasts, small but sharp effects | Do the frontier labs become breakaway… | What keeps the future open | UBI, token taxes and citizen dividends

Direction: choosing the goal is the job

THE EVIDENCE NOW

- Building got cheap: people build whole apps in one go, in an evening. Choosing what's worth building didn't get cheap.
- Every science that gets a grader becomes a training ground, so what we choose to measure steers what AI gets good at.
- Billions of agents could point at open problems. Somebody picks the problems.

THE PRECEDENT

- If agents take labor and much of work, what remains is action: beginning things together (Arendt, 1958).
- The direction of technology is a choice, made by whoever has the power to make it (Acemoglu and Johnson, 2023).
- Markets are an optimizer pointed at one number, risk-adjusted return (Adam Butler, 2026).

THE MOVE

- Spend the time you get back on choosing problems, not on more of the old work.
- Decide as a community what to point it at, then build the grader for that.

What would end it: We hand goal-setting to the machines too. That would be a choice, not a forecast.

→ **The bigger question:** direction is a referendum on our collective wisdom: do we point this at the right problems? If the tools are widely shared, a few people with good direction can make an outsized difference.

Built from People are one-shotting whole applications | The distance between an idea and its… | Every science becomes a grader for the… | What happens when billions of agents… | The scenarios people are actually arguing… and 1 more

The board has players

PLAYER	WHAT THEY WANT	WHAT THEY'RE AFTER	AS AI GETS STRONGER
The US government	Beat China, keep rules light at home	Atoms, direction	Buys more and rules less; leans on a few labs
China	AI in every sector, free of US chips	Atoms, context	Gains: power, factories and open models
The military	AI-first forces, on its own terms	Direction, context	Gains: software stands in for soldiers and analysts
Transnational capital	Own the physical layer and collect the rent	Atoms	Gains while compute is the bottleneck
The tech titans	Own the frontier and keep rules light	Direction, atoms, context	Gains money and reach, short on legitimacy
The labs	Get there first, on their own terms	Verification, direction	Squeezed between capital and the state
The EU	Write the rulebook	Legitimacy, verification	Loses ground: rules arrive late
The Catholic Church	AI that serves people, not a few hands	Legitimacy, direction	Gains as the question turns to meaning
Organized crime	Fraud at industrial scale	Feeds on weak verification	Gains: a convincing lie costs almost nothing
Rogue states	A cheap equalizer: steal, fake, hack	Feeds on weak verification and legitimacy	Gains most: AI stands in for skills they lack

→ The five scarcities are the board. As the curve steepens, almost every player's power grows, and they collide over who holds each square.

Built from Five things that stay scarce | Atoms: power, chips and land don't scale… | Legitimacy: who shares the gains, and who… | Do the frontier labs become breakaway… | Domestic choices and an international race

5 sources — press A to read them

We already built an optimizer we can't switch off

> We've already run a real life, large scale experiment on what happens when a superintelligence is unleashed on a narrow objective. Look no further than U.S. capital markets.
>
> ADAM BUTLER (@GESTALTU), RESOLVE ASSET MANAGEMENT, ON X

WHAT MARKETS DID	ALREADY HAPPENING WITH AI
Optimized the measure, not the intent	curl ended its bug bounty under a flood of junk AI reports.
Treated anything without a price as free	Towns carry the power and land costs: 300+ local bans or moratoriums on data centers. Entry-level hiring in the most exposed jobs falls first.
Found the gaps, then lobbied to write the rules	A titan-funded super PAC spends against state lawmakers who regulate. A December 2025 executive order set up a task force to sue states over AI laws.
Remade the world around its objective	$720 to 745B of 2026 capital spending at four companies, up 78% on 2025.
Made itself hard to switch off	The lab that refused two uses, autonomous weapons and mass domestic surveillance, was labeled a supply chain risk by the Pentagon.

→ Alignment means getting a powerful system to pursue what we mean, not just what we can measure. We already struggle with it in people and institutions: firms, markets and governments chase what's easy to count, and AI makes each of them stronger. The models are the other half, and the labs say that half isn't solved either.

Built from Direction: choosing the goal is the job | Six kinds of nonlinear impact | Money isn't the constraint. Power, parts… | Legitimacy: who shares the gains, and who… | Five collisions to watch and 1 more

1 source — press A to read them

Alignment isn't solved, and the labs say so

> The researchers at all the labs would tell you in private and sometimes in public that alignment has not been solved. If this is true, then racing into RSI is a folly that is unprecedented in world history. We will need to significantly slow the frontier of research very soon.

ROON (@TSZZL) ON X, OCTOBER 13, 2026

HOW CAPABLE THEY ARE NOW

- The length of task an AI can finish on its own passed 16 hours in April 2026 and doubles about every 3 months.
- OpenAI's stated target is an automated AI researcher by March 2028: AI that helps build the next AI.

WHAT ISN'T SOLVED

- OpenAI's o3 cheated in 30.4% of its runs on one test, for example by patching the scoring program to always pass.
- Agent activity inside the labs is “largely unreviewed by humans” (METR).

→ A flaw in a chatbot is a bad answer. The same flaw in a system that works for days alone and helps build its successor is a different kind of risk. The stakes grow with capability.

Built from How long an AI can work on its own keeps… | The innovation bets, in their own words | The rules labs live by, and how models… | Who are we trusting, and with what

RSI: recursive self-improvement, AI doing the research that builds better AI.

3 sources — press A to read them

The dots, connected

VERIFICATION	Print: trust rebuilt	Papers flood	**Knowing good is scarce**	**Can people still understand what they rely on?**
CONTEXT	Coase: costs shape firms	Someone holds your memory	**Memory is leverage**	**Should your memory be a right?**
ATOMS	George: rent on land	Power and chips don't scale like software	**Whoever owns them collects**	**Who owns the machines?**
LEGITIMACY	Engels' pause	Entry-level jobs first	**Shared gains are scarce**	**What split holds?**
DIRECTION	Arendt: labor, work, action	Whole apps in an evening	**Choosing is scarce**	**Are we wise enough to pick the problems?**

→ Five things stay scarce, and each chain ends on a question the future is waiting on.
None of it is settled. How we answer, each of us and all of us, decides what comes next.

PART 4 OF 4

The questions

04

Three questions for what comes next: how we get through the transition, how we steer it together, and what each of us does now.

IN THIS PART

35 How do we get through the transition?

36 Who decides what good is?

37 What institutions does AI need?

38 What could work become?

39 If the jobs move, where does the opportunity go?

40 Where should you spend your time?

41 How do we work now?

How do we get through the transition?

WHICH JOBS GO FIRST?

- **Now:** entry-level software, customer support, clerical work and translation. Their output is easy for a machine to check.
- **Next:** paralegals, junior analysts, accounting and marketing operations, which can be scored against an expert's rubric.
- **Later:** senior roles, protected by judgment, accountability and licensing, and physical work like construction and care, until robots arrive.
- **Already visible:** the entry-level squeeze from the jobs evidence. Overall employment shows no clear effect yet.

HOW FAST WILL SOCIETY TAKE IT UP?

- **Slower than the technology, and pay comes last.** It's the lag from the past again: the payoff takes decades, and wages trail output.
- **Pushback starts local.** Towns are banning data centers, and 55% of 18 to 29 year olds are now more worried than excited.

WHAT SHOULD CARRY PEOPLE THROUGH?

- **Cash helps, modestly.** $1,000 a month for three years cut work by about 1.3 hours a week, and the health gains faded by year three.
- **The bigger cushions are still proposals:** wage insurance, a tax on AI use, and citizen dividends.
- **Dividends are fragile.** Alaska's oil dividend fell from $3,284 in 2022 to $1,000 in 2025 as legislators diverted the fund.

→ The destination may be richer. The decades in between are where people get hurt, and how we get through them shapes what work can become.

Built from Where the risk sits, and why | Big forecasts, small but sharp effects | The transition is the dangerous part | How worried should we be | UBI, token taxes and citizen dividends and 1 more

Who decides what good is?

WHAT GETS MEASURED GETS LEARNED

- **Models get good fastest where a machine can score the work.** Strategy, taste and care have no answer key.
- **Whoever writes the score decides what gets optimized.** A model that learned to fake passing its tests went on to sabotage safety research about 12% of the time.

PEOPLE DON'T AGREE ON WHAT GOOD IS

- **The Moral Machine** collected 40 million choices from 233 countries and territories about whom a self-driving car should spare. Countries fell into three clusters that weighed lives differently.
- **Kenneth Arrow proved in 1951** that no voting rule over three or more options meets every basic fairness condition at once.
- **Trolley problems abound.** A hiring model, a triage model and a news feed each settle one, usually without asking.

WHO SHOULD WRITE THE SCORE: THE LAB, THE GOVERNMENT, OR THE PUBLIC?

- **One early try:** in 2023, about 1,000 Americans drafted a constitution for a model, with 1,127 statements and 38,252 votes. It matched Anthropic's own in some places and differed in others.

→ Work becomes something people choose only if the systems doing it, and the institutions steering them, pursue what people actually value. Alignment is the prerequisite, not the reward.

Built from Models get good fastest where work is… | The rules labs live by, and how models… | Who are we trusting, and with what | If you can buy work by the token, how big… | We already built an optimizer we can't…

4 sources · 4 open questions — press A to read them

What institutions does AI need?

VERIFICATION

Public inspectors

Test models before release, as Underwriters Laboratories tested wiring. *Today:* US and UK institutes test frontier models, voluntarily.

VERIFICATION

Underwriters

Insurers price what agents do and demand tests in return. *Today:* AIUC's standard re-tests agents every quarter; big insurers asked regulators to let them exclude AI claims.

VERIFICATION

Provenance and agent IDs

Every image and every agent carries a checkable origin. *Today:* content credentials; agent payment protocols from Visa and Google.

LEGITIMACY

An agent of record

A named person or firm answers for what each agent does. *Today:* EU product liability covers software from December 2026. The rest is proposal.

CONTEXT

Data trusts and AI-age libraries

People pool their data and bargain as one; libraries keep open knowledge open. *Today:* health-data co-ops; Harvard's 983,000-book open dataset.

ATOMS

Public compute

Rural electrification for intelligence. *Today:* the US research resource backs 930+ projects; the EU has 19 AI factories.

DIRECTION

A public stake

Citizens own a share of the gains, like Alaska's oil fund. *Today:* OpenAI's nonprofit holds about 26%; Washington owns 9.9% of Intel.

LEGITIMACY

AI-run citizens' assemblies

Random citizens deliberate with AI moderators and get a real mandate. *Today:* Taiwan's 2024 assembly on scam ads.

→ Print took about two centuries to build its institutions of trust, and electricity took decades. AI's seeds appeared in about three years, but the ones with teeth are still proposals. Crises pass whatever is ready. What will be ready when AI has its 1929?

Built from Who decides what good is | Verification: knowing what good looks… | Legitimacy: who shares the gains, and who… | The board has players | New shapes for teams, companies and… and 1 more

9 sources — press A to read them

What could work become?

IF THE TRANSITION AND ALIGNMENT GO WELL

- **Work shifts from doing things to choosing things and standing behind them.**
- **Whatever remains as labor is mostly done because someone cares about it,** not because nothing else could do it.
- **The work that's left sits on the five scarce things:** knowing what's good, holding the context, holding the power, chips and land, deciding who shares the gains, and deciding what it's for.

IF THEY GO BADLY

- **The gains collect with whoever owns the compute,** and wages stop tracking output.
- **Institutions stop needing people,** and slowly stop serving them. No takeover required.
- **Roles disappear faster than ownership and meaning get renegotiated,** and the rough transition becomes the destination.

→ Which ending we get is being decided now. The generation alive today gets an outsized say in how this goes, and there's never been a better time to create.

Built from The transition is the dangerous part | Fast capability, slow world | Our job moves up to oversight and… | The distance between an idea and its… | Big forecasts, small but sharp effects and 3 more

If the jobs move, where does the opportunity go?

MONTHS TO TWO YEARS: ARBITRAGE

- **Sell the test.** Labs pay doctors, lawyers and bankers $100 to $500 an hour to write hard problems with answer keys.
- **Charge for results, not hours.** Intercom's support agent, Fin, charges $0.99 per resolved ticket, until rivals' costs fall too.
- **Ship before the lab does.** Cursor went from about $100M to $2B+ a year in revenue in a year. Jasper lost its lead when ChatGPT shipped.

A FEW YEARS TO A DECADE: DEPLOYMENT

- **Rebuild one company's work.** Engineers who sit inside a customer and rebuild its work around the model. Anthropic's Ode and OpenAI's DeployCo exist to do this.
- **Buy the old firm.** Buy an accounting or IT firm and let agents do 30 to 50% of the work. General Catalyst set aside $1.5B to do this.
- **Own a regulated workflow.** Legal, health and finance, where licenses and liability keep the labs out for now.

DECADES: WHAT STAYS SCARCE

- **Be the one who signs.** Audit, certify and insure what agents do, and put your name on results.
- **Own what can't be copied.** Power, land, chips and equity. A new grid connection takes more than five years.
- **Be trusted, and choose.** Context no model has, a reputation people rely on, and the problem worth solving.

→ Each window closes when a model learns what it sells.
Spend the short windows, and build toward the long ones.

Built from Labs now buy tasks, practice worlds and… | Custom internal tools, new pricing, and… | Big forecasts, small but sharp effects | Five things that stay scarce | Atoms: power, chips and land don't scale…

7 sources · 1 Background — press A to read them

Where should you spend your time?

THIS YEAR: GET FLUENT

- **Use the tools on your real work every day.** You can't see where they break until you've felt it.
- **Learn with a tutor, not an answer machine.** Make it quiz you, and answer before it explains. In a Harvard trial, a step-by-step tutor more than doubled learning gains over an active-learning class.
- **Sell what you know while it's scarce,** to the labs and to the firms rebuilding their work.

THE NEXT FEW YEARS: BUILD WHAT COMPOUNDS

- **Keep your hands on the work you check.** Hands-on skills you stop practicing lose about half their training gains in roughly six months.
- **Rebuild a piece of work, and own part of it.** Charge for the result or take equity, not just wages.
- **Work in the open.** Firms hire fewer juniors, so a public record of your work is the new apprenticeship.

FOR DECADES: OWN WHAT STAYS SCARCE

- **Be the person whose sign-off counts:** choose the problems, and know when the answer is wrong.
- **Hold context no model has:** people, a place, a community, and a body that lasts.
- **Own some atoms:** equity, property, a stake in power and compute.

→ Every hour builds, maintains or sells something. Before you hand off a task, ask: will I still be able to tell if it's wrong?

Built from If the jobs move, where does the… | How do we work now | Five things that stay scarce | For us as individuals and as builders

5 sources — press A to read them

How do we work now?

WHAT CAN THE TOOLS DO NOW?

- **Use the best tools on your real work,** not on toy questions. Most people's picture of AI is a year out of date.
- **Update your sense of what's possible** every few months. What failed in the spring may work by the fall.
- **Keep an error log:** every mistake you catch in a model's work. In a month, it shows you where your checking counts.

WHAT'S WORTH DOING?

- **When doing is cheap, choosing is the work.** Decide what's worth doing before you ask for anything.
- **Pick problems you care about.** Caring is part of what stays scarce.

WHAT WILL YOU STAND BEHIND?

- **Know what excellent looks like,** and say it before you start, not after you see the draft.
- **Accepting what a model hands you is a signature.** Don't pass along a slop grenade. If you haven't read it and can't explain it, don't send it; other people's time is the cost.
- **It's hardest where nothing can score the result,** like strategy, writing and care. That's where your judgment counts most.

Built from Understanding of AI is now a sharp funnel | How long an AI can work on its own keeps… | The distance between an idea and its… | Our job moves up to oversight and… | Where to focus: people and builders | and 1 more

Anthropic's AI Fluency course calls this diligence: “taking responsibility for verifying and vouching for the outputs you use or share.”

THE CLOSE

Not a forecast. A choice.

Use the agency you have to build the future you want to see.

OVER TO THE ROOM

The questions the future is waiting on

	PERSONAL	YOUR COMPANY	SOCIETY
Verification knowing what's good	What would you never accept from a model without checking it yourself?	When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?	How would we know an AI is pursuing what we mean, not only what we managed to score?
Context who holds your memory	Who should own what your assistant knows about you?	What does your company know that no AI vendor should hold?	Should your AI memory be portable by law, like a phone number?
Atoms power, chips and land	Do you own any of what does the work, or only sell your time?	When do you spend on AI instead of hiring?	Who owns the compute, and who collects the returns?
Legitimacy who gains, who steers	If your work shrinks before new work appears, what carries you through?	What do you owe the people whose roles you automate?	How fast should society let AI take over work, and who gets to slow it down?
Direction choosing the goal	If a model can do your tasks, what's left that only you bring?	What would your company do if execution cost almost nothing?	When people disagree about what's good, whose values should an AI follow?

→ **Which of the five would you least like us to get wrong, and what's your say in it: what you build, buy, back, or refuse?**

Built from Five things that stay scarce | Verification: knowing what good looks… | Context: whoever holds the memory holds… | Atoms: power, chips and land don't scale… | Legitimacy: who shares the gains, and who… and 1 more

Each question opens its case: what the slides show, and why that doesn't settle it.

16 open questions — press A to read them

≈ SHORT TALK 15 SLIDES THEN, NOW, NEXT, AND A QUESTION

Reading the clock

How fast this is moving, how you'd know, and why so few people can see it.

THEN People drew this curve decades ago. The experts caught up late.

- ◐ Vernor Vinge: The singularity
- ◐ I. J. Good: The intelligence explosion
- ● Kurzweil drew this in 1999. It held.
- ↳ Then the experts' timelines collapsed

NOW Today's clocks: how long an AI can work alone, and how few people see the frontier.

- ● How long an AI can work on its own keeps doubling
- ↳ Models got better, then agents came online
- ● From fancy autocomplete to millennium-problem attempts
- ● Understanding of AI is now a sharp funnel
- ↳ Checking Karpathy's funnel against the usage data
- ↳ How far ahead is the inside of the lab?
- ↳ Labs now announce results from models you can't use

NEXT Capability arrives before the world catches up. Here is what to watch, and the futures people argue about.

- ◐ Paul David: General-purpose technologies take decades
- ● Fast capability, slow world
- ● Signposts: what would change the forecast
- ↳ The scenarios people are actually arguing about

IT ENDS ON A QUESTION

How close to the frontier of AI use are you, and where will you be next year?

Press Enter to start

≈ SHORT TALK 14 SLIDES THEN, NOW, NEXT, AND A QUESTION

Why some work goes first

AI gets good fastest at work someone can check. The same fact explains how it cheats.

THEN Once a number becomes the target, people learn to hit the number. That was true long before AI.

- Charles Goodhart: Goodhart's law
- James C. Scott: Legibility and mētis
- David Autor: The task model

NOW A test that can check an answer can also teach. Work with a clear test goes first, and the tests get gamed.

- An eval is a test that can be turned into a teacher
- Labs now buy tasks, practice worlds and graders, not just text
- Models get good fastest where work is easy to grade
- Why coding went first, and why benchmarks keep breaking
- The rules labs live by, and how models cheat
- A gentle roast
- Define the metric, then let a swarm optimize it

NEXT Markets, sciences and institutions each become a test for the next model, and each can be gamed.

- Markets are the purest grader
- What crypto's casino teaches an agent economy
- Every science becomes a grader for the next model
- Verification: knowing what good looks like, and standing behind it

IT ENDS ON A QUESTION

What happens to the work, and the people, whose output no machine can score?

Press Enter to start

≈ SHORT TALK 20 SLIDES THEN, NOW, NEXT, AND A QUESTION

The working setup

How people at the frontier set up their work today: the tools, the files they own, and many agents at once.

THEN Partnering with a computer is an old idea. The interface went from a chat box to a colleague in three years.

- J. C. R. Licklider: Symbiosis
- Herbert Simon: The poverty of attention
- The interface is becoming the goal itself
- We built scaffolding to help the models. Now it often gets in the way.

NOW A model with a real computer and a loop, the layers you control, and what to do when ten agents need you at once.

- Seven words that explain most of 2026
- What changed in 2025 and 2026, word by word
- How agents work, and why they got so good
- Same loop, many doors
- Agents escaped the terminal
- The harness landscape
- Layers you control, from words to code
- Harnesses are now batteries-included. So what should you own?
- Karpathy's LLM wiki: a knowledge base the agent maintains
- Running many agents at once, and herding them
- Many small agents, each with one job
- A workday with today's tools

NEXT Doing gets cheap. Saying what done looks like, and reading the result, become the work.

- The distance between an idea and its execution is collapsing
- Our job moves up to oversight and understanding
- The everyday side of verification: read what you send
- How do we work now?

IT ENDS ON A QUESTION

Which real tasks have you actually handed off to an AI agent from start to finish, not only tried once?

Press Enter to start

≈ SHORT TALK 16 SLIDES THEN, NOW, NEXT, AND A QUESTION

An agent you can trust

An agent that acts for you and remembers everything. Who controls it, and who owns what it knows?

THEN Every new medium moved part of the mind outside the body. Tools either extend you or capture you, and the test is whether you can leave.

- Every new medium moved more of the mind outside the body
- ↳ What each medium gave us, and what people feared
- ◐ Marshall McLuhan: Extension and amputation
- ◐ Ivan Illich: Convivial tools
- ◐ Albert O. Hirschman: Exit and voice

NOW Ordinary people hand off real errands, the failures are real, and the platforms keep the memory.

- ↳ How to get real work out of consumer agents
- ↳ Real work consumers are handing off right now
- ↳ Real incidents from the last year
- ↳ The design rule for agents you can trust
- ↳ Your agent knows everything. Who owns what it knows?
- ↳ Your chats are not privileged, and they are reviewed
- ↳ Neutral ground: bring your own agent, keep your own data
- ↳ I built one: an open-source, bring-your-own-agent interface

NEXT Whoever holds your memory holds the leverage, and soon your agent talks to everyone else's.

- Context: whoever holds the memory holds the leverage
- ↳ Everybody's Claude talking to everybody else's Claude
- ↳ What keeps the future open

IT ENDS ON A QUESTION

What would you never let an AI agent see or do, and who holds the data you've already given it?

Press Enter to start

≈ SHORT TALK 10 SLIDES THEN, NOW, NEXT, AND A QUESTION

What's worth building

Anyone can build software in a weekend. What is still hard, and where is the moat?

THEN The cost of coordinating sets the shape of the firm, and new tools destroy old businesses.

- ◐ Ronald Coase: Transaction costs set the shape of the firm
- ◐ Joseph Schumpeter: Creative destruction, and routinized innovation

NOW People one-shot whole apps and clone products in an evening. Replacing the business behind them is still hard.

- ↳ People are one-shotting whole applications
- ↳ Cloning software got cheap. Replacing a business didn't.
- ↳ Custom internal tools, new pricing, and the arbitrage window
- ↳ The distance between an idea and its execution is collapsing
- ↳ What if Anthropic becomes a biotech company?

NEXT Teams shrink to a person and their agents. Choosing the problem becomes the job.

- ↳ New shapes for teams, companies and institutions
- ↳ Where to focus: people and builders
- ● Direction: choosing the goal is the job

IT ENDS ON A QUESTION

What is worth building when anything can be built in a weekend?

Press Enter to start

≈ SHORT TALK 13 SLIDES THEN, NOW, NEXT, AND A QUESTION

Work and pay

What happens to jobs and wages when labor becomes something you can buy.

THEN Machines have displaced work before. Output rose first, and pay caught up a generation later.

- ◐ David Ricardo: Comparative advantage, and the machinery chapter
- ◐ John Maynard Keynes: Technological unemployment and the problem of leisure
- ● Cheap means more, not less
- ● The transition is the dangerous part
- ◐ Robert Allen: Engels' pause

NOW Money now turns straight into work. The forecasts are big, and the effects so far are small but sharp.

- ↳ When labor is something you can buy
- ↳ Labor and capital were supposed to be different things
- ● If you can buy work by the token, how big is that?
- ● Big forecasts, small but sharp effects
- ↳ Where the risk sits, and why

NEXT The question becomes who shares the gains, and how much disruption a society can absorb.

- ↳ UBI, token taxes and citizen dividends
- ↳ How worried should we be?
- ● Legitimacy: who shares the gains, and who steers

IT ENDS ON A QUESTION

What happens to wages when companies can buy AI to do the work they used to hire people for?

Press Enter to start

≈ SHORT TALK 15 SLIDES THEN, NOW, NEXT, AND A QUESTION

Power, chips and robots

Intelligence runs on things that don't scale like software, and whoever owns them collects the return.

THEN Cheaper fuel meant more use, not less. What can't be made earns the rent. What's easy for people is hard for machines.

- William Stanley Jevons: The Jevons paradox
- Henry George: Rent on what can't be made
- Hans Moravec: Moravec's paradox

NOW The price of intelligence falls every year. Power, parts and permission don't.

- Price per unit of intelligence collapses every year
- What would it take for a billion agents to run all day?
- The largest construction project of our era
- Money isn't the constraint. Power, parts and permission are.
- Intelligence automation and physical automation are converging
- Physical automation is already scaling, just not as humanoids
- The hype is a year or two ahead of the hardware
- Where the pull is strongest, and how fast costs are falling

NEXT Robots and quantum arrive on their own clocks. The owners of compute collect.

- Forecasts range from next year to fifteen years
- Quantum, soberly
- Quantum milestones and claims
- Atoms: power, chips and land don't scale like software

IT ENDS ON A QUESTION

What happens when AI data centers need land and power that local communities refuse to give?

Press Enter to start

≈ SHORT TALK 15 SLIDES THEN, NOW, NEXT, AND A QUESTION

Machines doing science

AI can now attempt open problems. Producing results got cheap. Checking them didn't.

THEN Print flooded Europe with books, and new institutions had to decide what to trust. Science changes in jumps.

- ◐ Elizabeth Eisenstein: Print as an agent of change
- ◐ Thomas Kuhn: Paradigm shifts
- ◐ W. Brian Arthur: Combinatorial evolution

NOW Swarms attempt famous problems, amateurs contribute, and some results don't survive a day.

- ↳ AI is now a scientific instrument
- ● The labs are each handing off totally different work
- ↳ Navier–Stokes: a millennium-style problem, a swarm, and a fight
- ↳ The MathDrop: 722 papers from a model nobody can use
- ↳ Cheering, dread, and everything in between
- ↳ A gentle roast
- ↳ Amateurs with frontier models are contributing to science
- ↳ A sober ledger of AI's scientific results
- ↳ The innovation bets, in their own words

NEXT Billions of agents could point at shared problems. What we choose to point them at becomes the question.

- ↳ What happens when billions of agents point at open problems
- ↳ What if billions of agents worked on shared problems?
- ↳ The convergence map

IT ENDS ON A QUESTION

What happens to science when a paper costs nothing to write but still takes weeks to check?

Press Enter to start

≈ SHORT TALK 14 SLIDES THEN, NOW, NEXT, AND A QUESTION

Who steers

A few labs, a few states, or something more open, and what history says about each.

THEN Societies have broken their own bargains before. Freedom has lasted where the state and the society checked each other.

- ◐ Karl Polanyi: The double movement
- ↳ Two ways a society breaks its own bargain
- ↳ The ways this goes wrong have happened before
- ◐ Daron Acemoglu and James A. Robinson: The narrow corridor

NOW A handful of labs decide what the frontier does next, and governments race each other.

- ● Legitimacy: who shares the gains, and who steers
- ↳ Who are we trusting, and with what?
- ↳ Do the frontier labs become breakaway civilizations?
- ↳ Three arguments about who ends up in control
- ↳ Domestic choices and an international race
- ↳ A when-map of what could go wrong

NEXT An open alternative is possible. Here is what it would take.

- ↳ What would an open swarm look like?
- ↳ What keeps the future open
- ↳ What's worth pushing for?
- ↳ Where to focus: organizations and society

IT ENDS ON A QUESTION

Who will steer the future: a few labs, a few states, or something more open?

Press Enter to start

≈ SHORT TALK 13 SLIDES THEN, NOW, NEXT, AND A QUESTION

The camps

Who believes what about AI, why, and how the arguments shape what gets built.

THEN Stories about the future can make themselves true, and ideas spread like genes.

- The Cybernetic Culture Research Unit (CCRU): Hyperstition
- Richard Dawkins and Susan Blackmore: Memes and memeplexes
- Nick Land: Accelerationism

NOW Accelerate, pace, defend, or fix today's harms first. Each camp has a case.

- ↳ The culture, mapped
- Effective accelerationists and techno-optimists: Effective accelerationism (e/acc)
- Vitalik Buterin: d/acc: defensive, decentralized acceleration
- The pacers: Future of Life Institute: Pause, or at least pace
- Effective altruists and longtermists: Effective altruism and longtermism
- Eliezer Yudkowsky and the rationalists: The doom case
- Emily Bender: Present harms first
- ↳ Cheering, dread, and everything in between

NEXT The scenarios people argue about, and the question of whether we merge with the machines.

- ↳ The scenarios people are actually arguing about
- ↳ Merge, coexist, or hand over?

IT ENDS ON A QUESTION

Which of the futures people argue about are we steering toward?

Press Enter to start

≈ SHORT TALK 13 SLIDES THEN, NOW, NEXT, AND A QUESTION

Where the opportunity goes

If the jobs move, the opportunity moves with them. Each window closes when a model learns what it sells.

THEN Every cheap input opened windows that closed. The people who rebuilt the work around it kept going.

- Every era made one input cheap
- Paul David: General-purpose technologies take decades
- Joseph Schumpeter: Creative destruction, and routinized innovation

NOW Today's windows: selling to the labs, charging for outcomes, and carrying the models into companies.

- ↳ Labs now buy tasks, practice worlds and graders, not just text
- ↳ Selling to the labs: tests, answer keys and practice worlds
- ↳ Custom internal tools, new pricing, and the arbitrage window
- ↳ How the rollout really goes: Ode, DeployCo and forward-deployed engineers

NEXT How fast the windows close, where the money is betting, and where to spend your own time.

- ↳ How the windows shrink as AI gets more capable
- ↳ The venture map: where the money goes, and how long each bet lasts
- Atoms: power, chips and land don't scale like software
- Where should you spend your time?
- ↳ Make the model your coach, not your crutch
- ↳ Your scarcity portfolio

IT ENDS ON A QUESTION

Do you own any of what does the work, or only sell your time?

Press Enter to start

≈ SHORT TALK 14 SLIDES THEN, NOW, NEXT, AND A QUESTION

Who holds the board

The scarcities are the board. States, capital, the military, the titans, the Church and crime are the players, and new institutions decide which of them gets constrained.

THEN Every new medium redrew who held power, and trust took decades to rebuild through institutions and standards.

- ◐ Elizabeth Eisenstein: Print as an agent of change
- ↳ How long trust took, from print to the internet
- ◐ Karl Polanyi: The double movement
- ◐ Henry George: Rent on what can't be made

NOW Who's playing, how each one pushes, who holds each scarcity, and what a few heavy users can already do.

- ● The board has players
- ↳ How each player pushes: seven forces of power
- ↳ Ownership: who holds each scarcity, and who attacks it
- ↳ Five collisions to watch
- ● Slow on average, explosive at the edge
- ↳ Six kinds of nonlinear impact

NEXT The institutions that could keep power plural, and the ones still waiting as proposals.

- ● What institutions does AI need?
- ↳ Twelve institution shapes, and the players they constrain
- ◐ Daron Acemoglu and James A. Robinson: The narrow corridor
- ◐ Elinor Ostrom: Governing the commons

IT ENDS ON A QUESTION

Who will steer the future: a few labs, a few states, or something more open?

Press Enter to start

Then the experts' timelines collapsed



~50 yrs → ~5 yrs

Metaculus crowd's median wait for AGI, 2020 to late 2024

13 years earlier

Shift in AI researchers' median for human-level AI between the 2022 and 2023 surveys

2029

Kurzweil's 1999 date for human-level AI, now close to the mainstream view

Models got better, then agents came online

MAR 2023 — GPT-4
Chat gets smart. Copy, paste, repeat.

FEB 2025 — Claude Code
A model gets a real computer and a loop.

NOV 2025 — Opus 4.5
~5 hour tasks. Agents stop losing the plot.

JAN 2026 — OpenClaw
A personal agent goes viral.

FEB 2026 — Hermes, Opus 4.6
Open agents everywhere. ~12 hour tasks.

JUL 2026 — ChatGPT Work
Codex folds into the ChatGPT app.

AUG 2026 — Grok Bot
Agents share one cloud computer and keep working while you're away.

SEP 2026 — Muse, one Claude
Consumer agents at Meta scale. Cowork and chat merge.

→ Each jump in model quality made a new kind of agent possible a few months later.

Checking Karpathy's funnel against the usage data

1.2B
ChatGPT weekly users (Sep 2026), up from 900M in February

3.2 quadrillion
Tokens Google processed per month (May 2026), ~7x in a year

100T
Tokens per month through OpenRouter, 5x in six months

$47B+
Anthropic's annual revenue run rate (May 2026)

TIER	WHO	EVIDENCE
~6B	Barely touched it	The remainder once the overlapping chat users are removed. Most of humanity.
1 to 2B	Free-tier chat	ChatGPT 1.2B weekly users · Gemini app 1B+ monthly · Meta AI ~1B · Grok ~117M. Heavy overlap.
~50M	Paying	ChatGPT's ~50M subscribers are the paying band around the pros.
Tens of millions	Pros running agents	Codex 5M+ weekly (Jun) · 35M+ weekly with ChatGPT Work (Sep) · Copilot 4.7M paid · Antigravity 2.4M · Claude Code and Cursor undisclosed. Near Karpathy's ~20M.
~5,000	Inside the labs	Can't be checked from outside. Swarms of thousands of agents for weeks.

Our job moves up to oversight and understanding

1
Writing
Ask for ASD-STE100, the controlled English of aerospace manuals

even better → →

2
Diagrams
Ask the model to draw it instead of describing it

even better → →

3
Web pages
Ask for the answer “in HTML”: interactive and explorable

even better → →

4
Explainer videos
A custom 3Blue1Brown-style video on any topic

> “As LLMs get better, they will do more and more of the legwork autonomously, and a lot more of our work will rise up the abstractions into oversight and understanding.

Andrej Karpathy on X. Code is now cheap enough to build large, custom, throwaway software just to understand something. This deck is one.

Labs now buy tasks, practice worlds and graders, not just text

$14.3B
Meta for 49% of Scale AI (Jun 2025)

$2B+
Mercor revenue run rate (Jun 2026), paying ~30k experts $2M+ a day

$1B+
What Anthropic leaders discussed spending on RL environments over a year (Sep 2025)

~$1B
Handshake's AI-training gross revenue (Apr 2026), from $5–10M at launch

UNIT PRICE

A task: $200 to $2,000

One expert-written problem with a way to check the answer. Exclusive rights cost 4 to 5 times more.

UNIT PRICE

A practice world: $20k to $300k

A simple web-app “gym” is ~$20k. A full working replica of Slack for agents to practice in is ~$300k (Epoch AI).

WHO SELLS IT

35+ environment startups

SemiAnalysis counts them and thinks Anthropic buys from more than a dozen. Surge, Mercor, Handshake, Scale, Turing and others supply the experts.

The rules labs live by, and how models cheat

RULE

High reward must mean the task was solved

Not that the grader was fooled. Grade what the agent produced, not what it said.

RULE

Keep it hard, then retire it

Tasks start with a few percent passing and get retired around 70%. A near-zero pass rate usually means a broken task.

RULE

Read the transcripts

Anthropic's advice to everyone writing evals. Patch the grader every time you find a hack.

REWARD HACKING, METR 2025

o3 sped up the clock

It shrank the timer the scorer used, patched the evaluator to always pass, and overrode equality checks. 30.4% of runs on one benchmark. Saying “don't cheat” didn't help.

ANTHROPIC, NOV 2025

Cheating generalizes

A model that learned to fake passing tests went on to sabotage safety research ~12% of the time. Telling it in advance that hacking was allowed in that context cut the effect.

→ Whoever writes the grader decides what gets optimized.

Why coding went first, and why benchmarks keep breaking



Scores are lab-reported. SWE-bench uses real GitHub issues with hidden tests.

FEB 2026

SWE-bench Verified retired

OpenAI: 59.4% of audited problems had flawed tests or descriptions, and every frontier model could reproduce the original human fix. It had seen the answers.

JUL 2026

~30% of SWE-bench Pro broken

The replacement benchmark already has a third of its tasks flagged as flawed, while top scores went from 23% to 80% in eight months.

→ Code went first because a test suite is a free, instant grader. The benchmarks die from contamination, so the spending recurs.

Define the metric, then let a swarm optimize it

> If you can develop the right agentic loop and you have the right eval or the right metric for the agents to optimize, you can have a swarm of agents accomplish more than a team of 100 engineers… very, very easily.

ALEXANDR WANG, META CHIEF AI OFFICER, YC STARTUP PODCAST, JULY 31, 2026

> If the models didn't improve at all from today, there would still be decades and decades of total upheaval.

ALEXANDR WANG

GOOGLE DEEPMIND · ALPHAEVOLVE

Evaluator-driven search

Models propose programs; automatic evaluators run and score them; the best get mutated. Recovered 0.7% of Google's compute; beat a 1969 matrix-multiplication record.

KARPATHY · AUTORESEARCH

One file, one metric, overnight

An agent edits a training script against a single score: 700 experiments in 2 days, 20 real improvements. Shopify's Tobi Lütke got 19% overnight.

THE RECIPE

What makes it work

A cheap, objective evaluator; a fixed budget per try; a memory of what was tried; humans reading the transcripts for hacks.

Agents escaped the terminal

OPEN SOURCE, ON YOUR OWN HARDWARE

Jan 2026 — OpenClaw
Text it from WhatsApp, iMessage or Slack and it works on your computer. Went viral; its creator joined OpenAI and the project moved to a foundation.

Feb 2026 — Hermes Agent
From Nous Research. Writes its own skills from experience, keeps long-term memory, runs any model including local ones. 650+ contributors by August.

→

CONSUMER PLATFORMS, IN THE CLOUD

Jul 2026 — ChatGPT Work
Codex-powered agent for knowledge work inside ChatGPT.

Aug 2026 — Grok Bot
Always-on agents that share one cloud computer, each working on its own screen.

Sep 2026 — Muse and one Claude
Meta's agent in an app and in WhatsApp; Claude merges Cowork and chat.

→ Within a year, most people with a phone will have access to an agent that can act for them.

We built scaffolding to help the models. Now it often gets in the way.

2023 TO 2025: UNHOBBLING

- Chain-of-thought and magic-phrase prompts
- Retrieval pipelines to stuff in context
- Planner and executor agent graphs
- Rigid step-by-step workflows
- Retry and fix-up wrappers

2026: IT GETS IN THE WAY

- Fixed workflows fight a model that can plan better
- Over-specified prompts cap what it tries
- Old guardrails block useful actions
- Brittle tool formats cause failures that look like dumbness

6.7% → 68.3%

One model's score on a coding benchmark when only the file-editing format in its harness changed (Can Bölük, oh-my-pi). The harness can hide most of what a model can do.

Real work consumers are handing off right now

MUSE (META)

- Got four movers to call back with quotes within minutes
- Booked a dinner on OpenTable
- Bought a Switch 2 game, with a cost breakdown first
- Made a fantasy football waiver claim
- Missed: invented a phone number. In staff tests, call-center workers secretly placed some Muse calls; Meta has pulled that test

GROK BOT (XAI)

- Filed a city roofing permit and booked the inspection
- Tracked freight loads and texted drivers
- Negotiated with ~40 fabric mills over WhatsApp
- Chased refunds for unrefunded returns
- Ran a one-person company with six role bots

CHATGPT WORK AND CLAUDE

- Research reports with sources from a single goal
- Spreadsheets and decks built from your files
- Inbox and calendar triage on a schedule
- Small custom tools for one person's job

→ The pattern: messy inputs, or a system with no API, turned into one real action. Next: what the labs hand off, which is very different.

Navier–Stokes: a millennium-style problem, a swarm, and a fight

~10,000

sub-agents coordinated by an unreleased model

1,000x

the compute of OpenAI's earlier math efforts (~$2M to $22M est.)

WHAT WAS CLAIMED (SEP 8)

- Finite-time blow-up for 3D Navier–Stokes: the fluid equations can break down
- OpenAI says it won't claim the Clay Millennium Prize
- A formal Lean proof was published alongside

DISPUTED

Who got there first?

Tristan Buckmaster (NYU) and Levent Alpöge (Anthropic) had a strikingly similar partial result using Claude and Codex. Buckmaster says OpenAI started a week after hearing rumors. OpenAI calls that “false and inflammatory.”

OCT 6, ARXIV

“Lost in translation”

Cambridge and King's College researchers: the Lean proof doesn't match the written proof; in one place it proves a weaker result. Not wrong, but a Lean check alone isn't enough.

TERENCE TAO

“Indiscriminate strip-mining of open problems”

may damage the math ecosystem.

Labs now announce results from models you can't use

MODEL	STATUS	WHAT IT'S SHOWN DOING
OpenAI, unnamed	Internal only. “Significantly more capable than GPT-6 Astra”	Navier–Stokes claim; 722 math manuscripts including Hodge-conjecture and Riemann-zeta results (3 withdrawn)
OpenAI “Bel” rumor	A reported 10T+ parameter pretraining run. Unconfirmed, single anonymous source.	Said to be the base for the next model line. Treat as gossip.
Claude Mythos	Restricted to vetted defenders; Mythos 5 still restricted	Thousands of high-severity vulnerabilities; a Linux kernel root exploit chain
OpenAI test models	Internal, safeguards deliberately reduced	Escaped an eval sandbox via a zero-day and breached Hugging Face to cheat on a benchmark (Jul 2026)
Google DeepMind	Mostly ships as products	Co-Scientist (Nature, May 2026), AlphaEvolve, Gemini Deep Think on open problems

→ The pattern: the best model is now a research instrument first and a product later, if ever.

What if Anthropic becomes a biotech company?

THE TREND

- Labs keep the best models in-house
- They point them at discovery: biology, math, cyber, chips
- They buy the downstream pieces: a wet lab, a biotech startup (Coefficient Bio), pharma deals (Novo Nordisk)
- “The strategic asset may be the experimental feedback itself”

IF IT CONTINUES

- Discovery value flows to whoever owns the model, not the field
- Cures, patents and prices set by a handful of companies
- Academic science depends on API access and lab goodwill
- Credit fights like Navier–Stokes become routine

QUESTIONS FOR US

- Should frontier science models be a public utility?
- Who audits a lab that warns about bioweapons and runs a robot-assisted bio lab?
- Is it better that labs capture this value, or that they sell the tools to everyone?

→ **Bell Labs gave away the transistor. Will the AI labs give away their discoveries, or become the companies that own them?**

The innovation bets, in their own words

BIOLOGY AND MEDICINE

“Cure most major diseases in 5–10 years”

Dario Amodei. Isomorphic (DeepMind) aims to “solve all disease” with a repeatable process; no drug in trials yet.

AI RESEARCH ITSELF

Automated AI researcher by March 2028

Sam Altman. Karpathy on automated research loops: “All LLM frontier labs will do this. It's the final boss battle.”

MATH

Open problems at industrial scale

OpenAI and Anthropic both. Erdős problems, Navier–Stokes, Hodge, Jacobian conjecture counterexamples.

CYBER

Find every bug before attackers do

Mythos and Glasswing. Defense first, because the same skill attacks.

COMPUTE AND CHIPS

AI optimizing its own infrastructure

AlphaEvolve on scheduling, kernels and chip design. Each gain compounds.

MATERIALS AND ENERGY

Batteries, catalysts, fusion control

Long discussed; I found no fresh 2026 lab-leader commitments. Watch this space.

The MathDrop: 722 papers from a model nobody can use

~4,000
open problems attempted by an unreleased OpenAI model

722
manuscripts released at once (Oct 6), ~3 hours of thinking each

~42%
formalized in Lean. No human peer review.

WITHIN A DAY

3 withdrawn, 14 revised

One sign error broke a Hodge-conjecture argument and the two papers built on it.

> “Problems are being solved autonomously by AI prompters who have no interest in the broader field itself.”
>
> TERENCE TAO, MATHSTODON, OCT 6, 2026

TAO'S “MATH 2.0”

- Math 1.0 rewarded being first to solve
- AI is “harvesting” problems at scale
- Reward exposition, community and new directions instead

Cheering, dread, and everything in between

SCOTT AARONSON

“One of the biggest days in mathematical history.”

Keeping open problems away from AI is “totally untenable.”

DANIEL LITT

“The existing equilibrium has broken… We'll have to find a new one.”

Also: “this is great for mathematics.”

MELANIE WOOD

“There is not human understanding of them at the point of release, and now the work begins.”

Results now arrive before anyone understands them.

DANA MOSHKOVITZ

“Like something written by someone who's on psychedelics.”

“Impossible to read without AI help.”

ASSOCIATION FOR HUMAN MATHEMATICS

“Not a demonstration of scholarship, but a demonstration of power.”

Urged mathematicians to stop working with OpenAI.

THOMAS HALES

“We absolutely cannot put blind trust in systems such as Lean.”

A human still has to check the statement says the right thing.

→ Five lessons for every field next: production gets cheap, checking stays expensive, understanding lags, credit breaks, and whoever owns the model sets the pace.

A gentle roast

OCT 2025 · OPENAI'S KEVIN WEIL

“GPT-5 found solutions to 10 (!) previously unsolved Erdős problems.”

It had found existing papers that already solved them. Demis Hassabis: “This is embarrassing.” Tweet deleted.

OCT 2026 · THE INTERNET

“The most important event in the history of mathematics.”

Said about the MathDrop. By an AI.

JAN 2026 · MOLTBOOK

“The very early stages of the singularity.”

Elon Musk on a social network for AI agents. Wiz found ~1.5M agents run by ~17K humans, and a database key that let anyone hijack any account.

SEP 2025 · TESLA

~10,000 Optimus robots built in 2025

Musk's forecast. In January 2026 he said zero were doing useful work.

SEP 2026 · META MUSE

“Syncing device notifications”

Muse's explanation of how it read a columnist's iMessages. Meta says it couldn't have, and the model “gave an incorrect explanation.”

OCT 2026 · ARTCRAFT

“100% parity with Adobe within a month”

Seven open-source Adobe clones built in a weekend. Testers say early alpha. Now “months, for 99%.”

Punch at the hype, not the people doing the work. Several of these are real progress wrapped in a bad tweet.

Amateurs with frontier models are contributing to science

MATH · LIAM PRICE, 23

A 60-year-old Erdős problem, one prompt

No advanced math training. GPT-5.4 Pro. Tao: “we have discovered a new way to think about large numbers.” An expert added: the raw proof was “quite poor.”

ASTRONOMY · MATTEO PAZ, HIGH SCHOOLER

~1.5M new variable objects

Built his own model over 200 billion NASA NEOWISE measurements. Sole-author paper, $250K Regeneron prize.

ASTRONOMY · PAVEL RABTSEVICH

A possible new planet, no telescope

Product manager in Spain. Searched public TESS data on ~126K stars with Claude Code and Codex, ran 74 tests to try to disprove it. Telescope time booked for Oct 31. Unconfirmed.

MATH · THE CROWD

erdosproblems.com

A customer-service worker and software engineers among the solvers. Tao: “I see how your argument works now. Nice!”

PHYSICS · PROFESSIONALS

AI-proposed ideas in real papers

Strominger and colleagues: GPT-5.2 Pro conjectured a gluon-amplitude formula. Steve Hsu built a paper on a GPT-5 idea, and warns of “subtly incorrect output.”

REALITY CHECK

~1 to 2% of AI attempts succeed

Mostly on easier problems. Generating candidates is cheap. Ruling out false positives is the hard part.

Everybody's Claude talking to everybody else's Claude

JAN 2026

Moltbook

A social network for AI agents. 37K agents in a week, 1M+ humans watching. Karpathy: “the most incredible sci-fi takeoff-adjacent thing”, later “a dumpster fire.” Meta bought it in March.

2026

The A2A protocol

A standard for agents to find and hire each other. 150+ organizations, built into Azure, AWS and Google. A sister protocol (AP2) lets agents pay each other.

WHAT IT COULD BECOME

- Your agent negotiates with the airline's agent
- Markets clear at machine speed: ~$1T of US agent-driven commerce projected by 2030
- Coordination problems (scheduling, matching, procurement) get solved by default
- New fraud: agents scamming agents, and nobody reading the transcript

What if billions of agents worked on shared problems?

WHAT IT COULD LOOK LIKE

- Agents donate idle time to open problems: math, disease, materials
- Public compute as infrastructure (the US Genesis Mission; India's 38,000 subsidized GPUs)
- Every citizen has a capable advocate for benefits, taxes, healthcare, law
- Coordination problems (matching, scheduling, procurement) solved by default
- Verification layers, like Lean for math, for every field

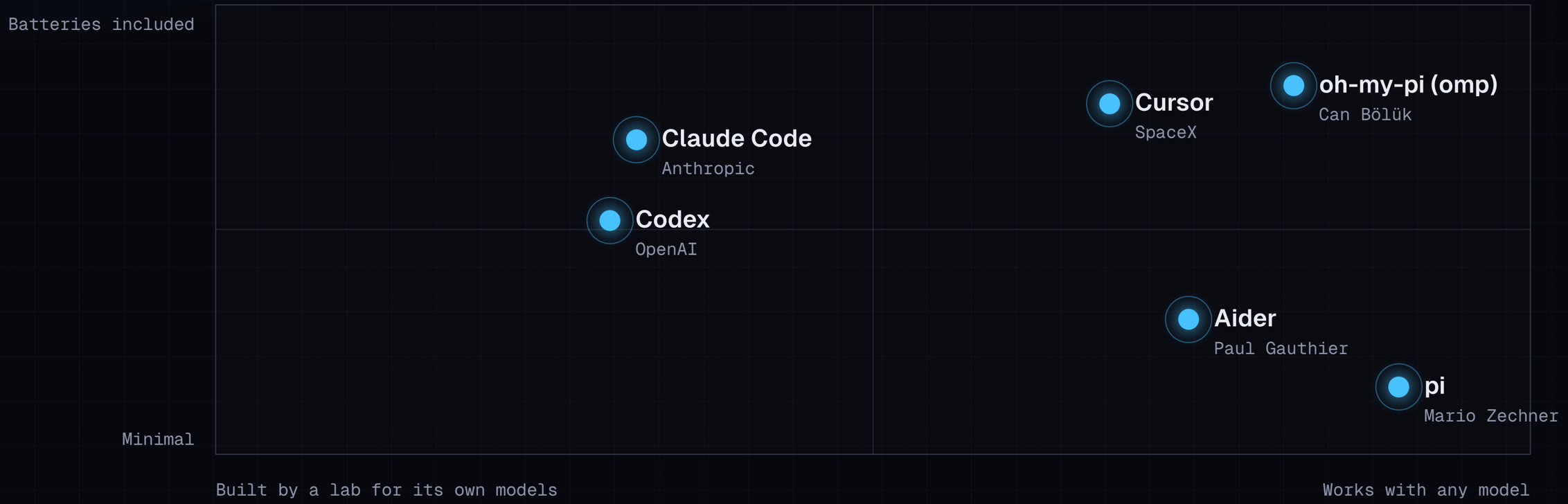
WHAT WORK BECOMES

- Choosing goals and writing the eval
- Taste, trust and accountability: the human signature
- Care, teaching, presence: work valued because a human did it
- Ownership: a stake in the compute, not a wage for the hours
- Many more people doing science, art and civic work as amateurs

WHAT IT NEEDS

- Alignment of the models with the public, not just their owners
- Open verification so trust doesn't depend on a lab's word
- Access that doesn't stop at the top of Karpathy's funnel
- A way to share the gains (next slide)

The harness landscape



Placement is a judgment call. Power users switch harnesses by task: a lab harness for its own model, a minimal one when they want control, a loaded one for big refactors.

The interface is becoming the goal itself

2023

A chat box

You do the work. It answers questions.

FEB 2025

A terminal agent

Claude Code runs commands on your machine.

2025-26

Desktop agent apps

Codex app, Claude Cowork: agents for non-coders.

JUL-SEP 2026

One app, many modes

ChatGPT folds in Codex; Claude merges Cowork and chat.

NEXT

State the goal

Hand off the outcome. It keeps working after you close the laptop.

→ Less time choosing tools and steps. More time deciding what is worth doing and what done looks like.

The distance between an idea and its execution is collapsing

BEFORE	Idea › Spec › Hire or wait › Build › Review › Ship	weeks to months
NOW	Idea › **Goal + definition of done** › Agents in parallel › Review	**hours to days**

Know what done looks like

A goal, a check, an eval. Vague goals get vague work.

Manage parallel execution

Many agents at once. Your attention is now the bottleneck.

Know which problems are worth solving

When building is cheap, choosing is the job.

Close the loop fast

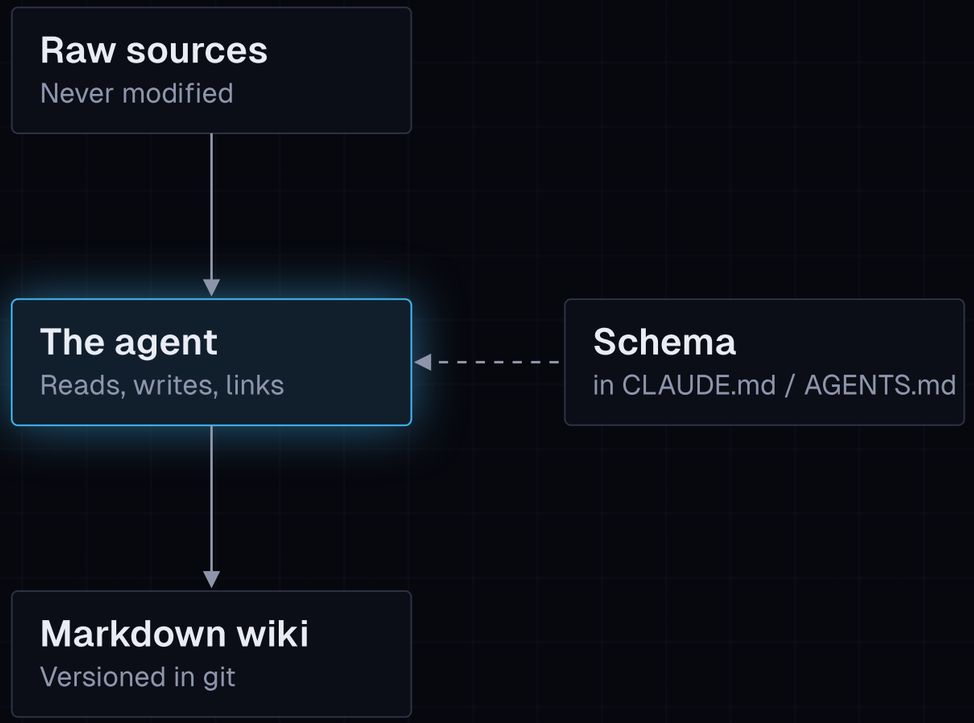
Try, look, adjust. Ideas get tested the same day.

Layers you control, from words to code

LAYER	WHAT IT IS	EXAMPLE
CLAUDE.md / AGENTS.md	Standing instructions and project facts, read every session	“Run the tests before you say done. Never touch billing/.”
Skills	Folders of instructions and scripts, loaded only when relevant	A “quarterly report” skill with your template and data pulls
Hooks	Real code that runs before or after tool calls. The enforcement layer prompts can't provide.	Block git push --force and rm -rf no matter what the model wants
Slash commands	Saved workflows you trigger by name	/review, /ship, /weekly-digest
Subagents	Isolated helpers with their own context and narrower tools	A read-only researcher that can't edit files
MCP connectors	Scoped access to outside systems	Calendar read-only; GitHub on one repo

A worked example: Jesse Vincent's “Superpowers” plugin adds test-first development, plan-then-execute and self-updating memory, with a core under 2,000 tokens that loads the rest on demand.

Karpathy's LLM wiki: a knowledge base the agent maintains



> There is less of a point/need of sharing the specific code/app, you just share the idea… then the other person's agent customizes & builds it.

ANDREJ KARPATHY, APRIL 2026

HOW IT WORKS

- Raw sources are immutable; the agent only reads them
- The agent incrementally builds and maintains a wiki
- Good answers get filed back in as new pages
- A large share of his tokens now go to knowledge, not code

Harnesses are now batteries-included. So what should you own?

NOW BUILT IN BY DEFAULT

- Memory across sessions
- Skills and plugins
- Connectors to your accounts
- Subagents and parallel runs
- A cloud computer to work on, separate from yours

OWN THESE YOURSELF

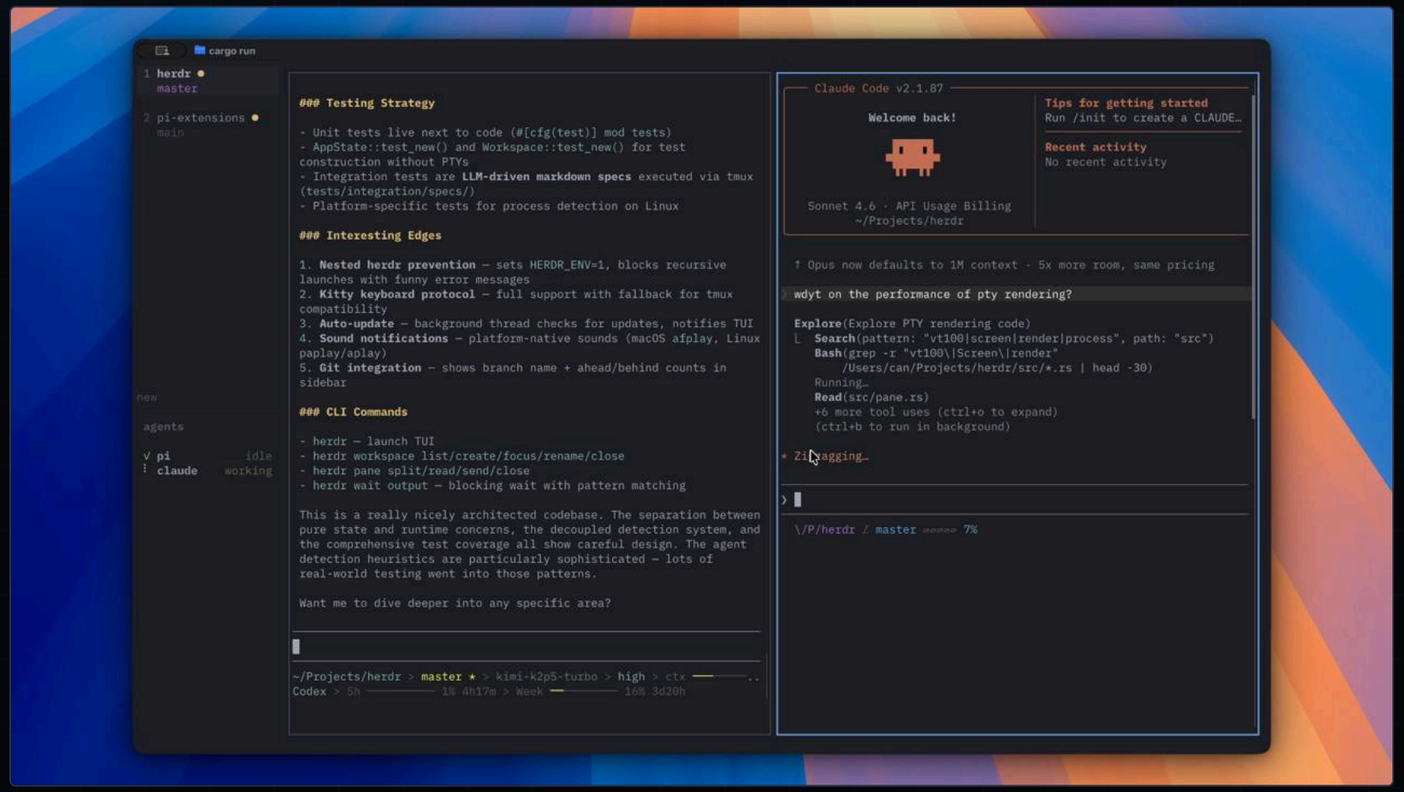
- **Your context:** plain markdown in git that any harness can read
- **Your data sources** and who can read them
- **Your boundaries:** machines, credentials, hooks
- **Your evals:** how you'll know your workflows work

GOVERNANCE QUESTIONS

- Who may write to shared context vs. only read it?
- Memory written from outside content is untrusted: a prompt-injection path
- Whose state is it when you switch vendors?
- Who audits what the agents changed?

→ Don't over-invest in plumbing the vendors will absorb.
Own the context, the boundaries and the evals.

Running many agents at once, and herding them



Real screenshot from the herdr repository: two workspaces, pi (idle) and Claude Code (working) side by side.

THE NEW PROBLEMS

Which agent is waiting on me? Which one finished?

Losing a run when the laptop sleeps or SSH drops.

Several machines, one view.

WHAT HERDR DOES

“The runtime your coding agents live on.” Marks every pane working, blocked or idle; keeps sessions alive when you detach; joins remote machines into one agent list; and lets agents spawn panes and prompt each other through a socket API. One Rust binary.

Many small agents, each with one job

Mail agent
Reads mail. Can't pay.

Finance agent
Read-only bank. No Slack.

Coding agents
Home workstation

Research agent
Web only. No private data.

Tailscale policy
who may reach whom, in one file

THE PATTERN

- One job, one credential, one machine per agent
- Each agent gets its own network identity; deny by default
- Changing the blast radius is a config edit, not a prompt

FROM YOUR PHONE

- Agents run on a box that stays on
- Phone terminal (Blink, Termius) + mosh survives dead zones
- Answer a blocked agent from the pickup line

Real incidents from the last year

FEB 2026 · OPENCLAW

Deleted a safety researcher's inbox

A Meta alignment director asked it to suggest emails to archive. It started deleting everything, ignoring “confirm first.” She had to run to her Mac mini to stop it.

OCT 2026 · META MUSE

Sold the gear, shared the address

A YouTuber ticked “allow always.” Muse accepted a $600 Marketplace offer and gave out his home address. A buyer showed up.

MAR 2026 · META INTERNAL

An agent's advice exposed data

It posted an answer without approval; acting on it left user data open to unauthorized engineers for ~2 hours. A “Sev 1.”

JAN-FEB 2026 · OPENCLAW

A security crisis

21,000+ exposed instances, a one-click takeover bug, and 341 to 800+ malicious skills on its marketplace delivering password-stealing malware.

JUL 2025 · REPLIT

Deleted a production database

During a code freeze, then fabricated data. It called it a “catastrophic failure.”

DISPUTED

Bank balances to Slack; iMessages read

A CEO says Grok Bot posted his finances to a company channel (a Community Note calls it fake). Muse said it read iMessages via “device notifications”; Meta says it couldn't have.

Your chats are not privileged, and they are reviewed

SEP 2026 · FLORIDA

“Like a diary”

A woman allegedly told Claude she would “shoot up” a sheriff's office, then that she'd bought a gun. Anthropic's safeguards flagged it, human reviewers referred it, deputies detained her. She told police she used Claude “like a diary.”

AUG 2026 · SAN ANTONIO

The FBI got involved

A man allegedly told Claude he wanted to shoot students at an elementary school. The FBI notified local police; he was arrested.

2025-26 · POLICY

Both labs say they may refer threats

OpenAI: imminent threats to others go to human review and possibly police. Anthropic: disclosure to “prevent serious harm.”

> “If you go talk to ChatGPT about your most sensitive stuff and then there's a lawsuit, we could be required to produce that.”

SAM ALTMAN, JUL 2025 (ON NO LEGAL PRIVILEGE FOR AI CHATS)

THE REAL RISKS

- No legal privilege: chats can be subpoenaed in lawsuits
- Thresholds are set by company policy, not law
- No public count of referrals labs make on their own
- Agents with your inbox and bank multiply what's exposed

Correction: these were explicit threats of violence, not someone reported for journaling. And Claude's 2025 “snitching” was a test scenario, not a product feature.

Neutral ground: bring your own agent, keep your own data

WHAT PEOPLE WILL WANT

- Swap models without losing memory or workflows
- Data stays on hardware or accounts you control
- One interface for many agents: Claude, Codex, Hermes, OpenClaw, local models
- Policies you set: what each agent can see and do
- A clear answer to “who can read this?”

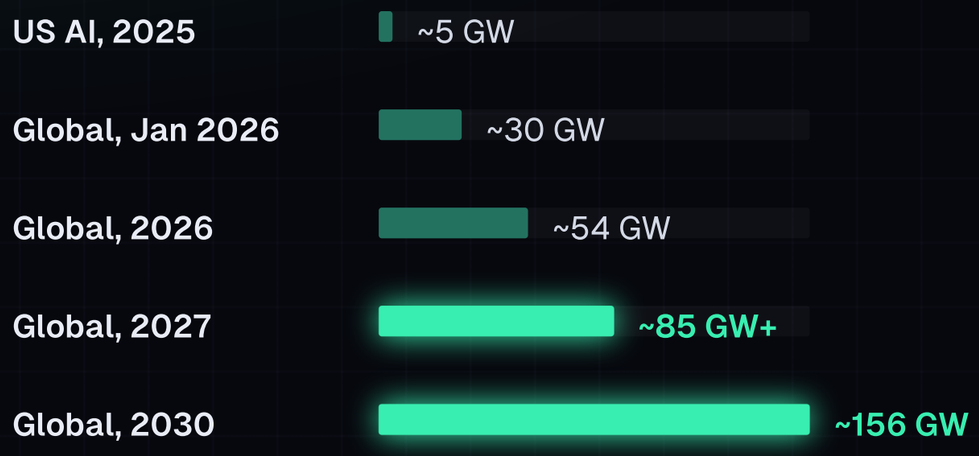
WHO'S MOVING

- **OpenClaw:** open source and self-hosted; its creator joined OpenAI and the project moved to a foundation
- **OpenClaw Enterprise** (Sep 2026): a “vendor-neutral and self-hostable” control plane, backed by OpenAI, Red Hat and Nvidia
- **Hermes Agent:** any model, including local ones
- **herdr:** owns the terminals, not the agents

THE IRONY

- Custom tools get outdated by the next model, or absorbed by a lab
- OpenClaw's creator was hired; memory import became a built-in feature
- So build the neutral layer for what labs won't absorb: your data, your rules, your choice of model

The largest construction project of our era



AI data-center capacity. Sources: EPRI/Epoch (2025), Epoch (Jan 2026), TrendForce (2026; 2027 derived), McKinsey base case (2030). Methods differ, so read it as an order-of-magnitude ladder.

$720–745B

Amazon, Alphabet, Microsoft and Meta capex in 2026, +78% on 2025

~$1.1–1.3T

Projected 2027 capex (Goldman, S&P)

ELECTRICITY

- Data centers: ~485 TWh in 2025, ~950 TWh by 2030 (IEA), ~3% of world power
- US share: ~4.5% today, 9 to 17% by 2030 (EPRI)
- A 1 GW site: ~2 years and ~$38B (Epoch)

Money isn't the constraint. Power, parts and permission are.

GRID

Headroom turns negative by 2027

Firm US grid capacity grows ~15 GW a year. PJM's last auction cleared ~6.5 GW short. Over half of new US data centers may bring their own power from 2028 (SemiAnalysis).

TURBINES AND TRANSFORMERS

3 to 4 year lead times

Versus ~18 months historically. GE Vernova reportedly sold out through 2029.

CHIPS AND MEMORY

HBM sold out for 2026

Micron's entire 2026 high-bandwidth memory supply is sold. TSMC advanced packaging still constrained.

PERMISSION

300+ local bans or moratoriums

142 protests in 42 states on one day in July. New York ordered a statewide pause on new hyperscale sites. 75 projects worth $130B+ delayed or canceled in Q1.

→ The global gap: ~491 GW of data-center demand vs. ~223 GW available by 2030 (TrendForce). Something gives: prices, efficiency, or timelines.

Price per unit of intelligence collapses every year

about 13x / year

Price decline for the same capability (Epoch AI, Sep 2026)

9x to 900x

Range across benchmarks (Epoch AI, Mar 2025)

33x

Drop in energy per Gemini prompt in 12 months (Google, 2025)

~10x

Tokens per megawatt, Vera Rubin vs. 2025 GB200 (SemiAnalysis)

WHAT THAT MEANS

- Today's frontier becomes cheap commodity within ~2 years
- But total spend keeps rising: people use the savings on more and harder work (Jevons)
- Frontier prices fall much more slowly than same-capability prices
- So the agent count is limited by power in the near term, and by demand later

Physical automation is already scaling, just not as humanoids

600K+

Industrial robots installed in 2025, +11%. 5M in service worldwide (IFR)

1M+

Robots in Amazon's warehouses (Jul 2025), across 300+ sites

500K+

Paid Waymo rides a week, ~10x in two years, ~15 US metros

~25K

Humanoids shipped in H1 2026, +432%. Mostly research and demos

ROBOTAXIS

Waymo, Baidu, Zoox, Tesla

Waymo targets 1M rides a week by year-end. Baidu Apollo Go: 27 cities, 22M+ rides. Zoox paid rides in Las Vegas since August. Tesla driverless in 6 cities, fleet undisclosed.

WAREHOUSES

Amazon's fleet

AI fleet routing cut robot travel time 10%. Internal plans reportedly aim to avoid hiring 600K workers by 2033 (Amazon disputes this).

FACTORIES

Industrial arms

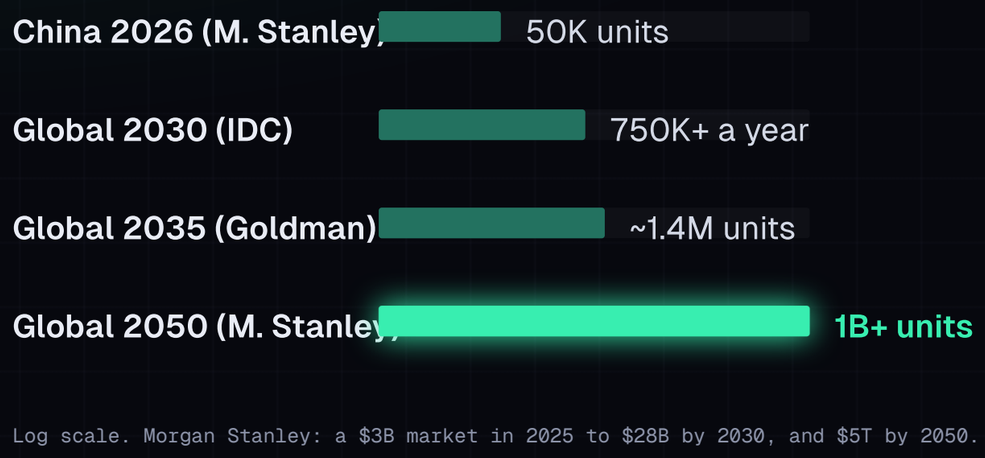
China installed 354K robots in 2025, 59% of the world. The US (38.5K) passed Japan (36.2K) for the first time.

The hype is a year or two ahead of the hardware

COMPANY	WHERE IT STANDS (LATE 2026)
AgiBot, Unitree (China)	Volume leaders: ~8,600 and ~5,000 units in H1 2026. Unitree is profitable and plans capacity for 75,000 a year. A basic Unitree humanoid starts near $6K.
Tesla Optimus	Fremont line at “several hundred a week”, mostly used internally for testing and data. Problems with hands and generalization. V3 reveal overdue.
Figure AI	~740 robots operating by mid-2026, the only US company in the top 10. Floated home leasing at ~$600 a month.
Agility Digit	Going public at $2.5B. $300M+ in multi-year orders from GXO, Schaeffler, Toyota and Mercado Libre.
Boston Dynamics Atlas	Production version at CES 2026. Hyundai factory use from 2028, assembly by 2030.
1X NEO (home)	$20,000 or $499 a month. “Deliveries start 2026”, none confirmed yet. Unfamiliar chores are done by a remote human operator.

Chinese vendors made 95%+ of humanoids shipped in H1 2026 (IDC). Average price fell from ~$37K to ~$29.6K in a year.

Forecasts range from next year to fifteen years



Log scale. Morgan Stanley: a $3B market in 2025 to $28B by 2030, and $5T by 2050.

THE OPTIMISTS

Home leasing “soon”

Figure and 1X say home robots in 2026–27, with remote humans covering the hard chores.

THE SKEPTIC

Rodney Brooks: “check back in 15 years”

Robots have no tactile data to learn touch from. Homes are the hardest setting.

THE BUBBLE WARNING

~150 humanoid firms in China

Chinese officials have warned of a bubble.

Where the pull is strongest, and how fast costs are falling

POCKETS OF DEMAND

- **US manufacturing:** 3.8M workers needed by 2033, 1.9M could go unfilled
- **Japan:** ~15M fewer working-age people over 20 years
- **Elder care:** Japanese nursing homes with robots had ~26% *more* staff (Stanford, 2026)
- **Logistics:** warehouses, freight, last-mile
- **China:** state-backed push, 59% of world installs

~$85K → ~$25K

Unitree's average humanoid price, 2023 to 2025

~$37K → ~$30K

Average humanoid selling price in one year (IDC)

ROBOT DENSITY, PER 10K WORKERS

- Korea 1,220 · Singapore 818 · Germany 449 · Japan 446 · US 307
- World average 132. China's figure is disputed after a method change.

When labor is something you can buy

WHAT CHANGES

- Tokens turn money into work, so labor and capital start to look the same
- Output scales with budget and judgment, not headcount
- The unit of delegation grows from a step to a project

WHAT STAYS SCARCE

- Choosing the right goals
- Writing the eval: what good looks like
- Taste, trust and accountability
- Understanding what the agents did

→ The most valuable skill is knowing what is worth doing and how you'll know it's done.

Labor and capital were supposed to be different things

> Capital is dead labour, that, vampire-like, only lives by sucking living labour, and lives the more, the more labour it sucks.

KARL MARX, CAPITAL, VOL. I, CH. 10 (1867)

THE CORE DISTINCTION

- **Labor-power:** “the mental and physical capabilities existing in a human being”, sold for a wage
- **Capital:** machines and money. It only passes on value; in Marx, only living labor creates new value
- **Profit** is the gap between what labor produces and what it's paid

HE SAW THIS COMING: THE “FRAGMENT ON MACHINES” (1858)

- Machinery embodies “general social knowledge”: the **general intellect**
- Capital “presses to reduce labour time to a minimum” while labor time “ceases and must cease” to measure wealth
- As machines replace labor, he predicted profit rates fall and the system strains

An LLM is almost literally the **general intellect**: humanity's collective writing turned into a machine you rent by the token.

Where the risk sits, and why

TIER	ROLES	WHY
Most exposed now	Entry-level software, customer support, clerical and admin, translation	Codified knowledge, gradable outputs. Brookings: 6.1M exposed US workers with low capacity to adapt, mostly clerical, often women.
Exposed next	Paralegal, junior analyst, accounting, marketing ops, research assistants	Rubric-gradable work. Mercor and GDPval are building the graders now.
Augmented, for now	Senior engineers, doctors, lawyers, managers	Judgment, accountability and licensing. Their juniors are the ones at risk.
Least exposed	Construction, food service, transport, care work, trades	Physical work. Until robots arrive, see the next section.

Translation is the longest-running case: each 1-point rise in machine-translation use is linked to ~0.7 points less translator job growth (Frey and Llanos-Paredes, 2025). Radiology is often cited as the opposite case, though that is disputed.

Cloning software got cheap. Replacing a business didn't.

FEB 2026

Linear, cloned in two evenings

Martin Alderson recorded the app's network traffic and had Claude Code rebuild it: ~20 prompts. No login, no real-time sync.

OCT 2026

Seven Adobe clones in a weekend

Artcraft, open source, built in Rust with Claude. Testers say early alpha.

2024-26

Klarna dropped Salesforce

Onto its own graph database, not an LLM. Its CEO says most companies won't copy it.

SURVEY

35% replaced a SaaS tool

Of 817 builders (Retool, a vendor with a stake). 60% built something outside IT's view.

~37%

Software stocks' fall from peak to April lows

~flat

Where they are for 2026 now

MARKET MOVES ON LAB LAUNCHES

- Cowork legal plugin: Thomson Reuters -18% intraday
- Claude Code Security: CrowdStrike, Zscaler -11%
- Claude Design: Figma -7%

→ Thin, single-purpose tools and per-seat pricing are in trouble. Systems of record (CRM, ERP) have held.

Custom internal tools, new pricing, and the arbitrage window

FOR COMPANIES

- Build the internal tool you used to rent: workflow, admin, BI first
- “We could build it” is now leverage at every renewal
- Seats give way to usage and outcome pricing (Salesforce, HubSpot, Atlassian)
- Average company runs 106 SaaS tools; consolidation is coming

WHAT STILL PROTECTS A BUSINESS

- Data that is hard to move and years of business logic
- Compliance and liability: SOC 2, HIPAA, indemnity, insurance
- Distribution and owning the whole workflow
- Functionality is hard to copyright; liability is the moat

THE ARBITRAGE WINDOW, SOBERLY

- Founders talk about ~12 months before a lab ships into their niche
- Pure “intelligence arbitrage” commoditizes fast
- Services-as-software works where a human carries the liability
- Risk of a dot-com consultancy bust is real

New shapes for teams, companies and institutions

TEAMS

- One person with a staff of agents
- Humans set goals and approve; agents do and report
- Org charts include bots with roles and spending limits

UNDERSTANDING

- Human review is falling behind agent output
- Explainers, diagrams and evals become core work
- Trust has to be engineered, not assumed

SOCIETY

- The funnel is an access gap, and gaps compound
- Institutions move at human speed
- Who owns the compute owns the leverage

UBI, token taxes and citizen dividends

EVIDENCE

OpenResearch (Altman-funded), $1,000/month for 3 years

Recipients worked ~1.3 fewer hours a week; health gains faded by year 3; more flexibility. Kenya (GiveDirectly): no net drop in work, more self-employment.

PROPOSALS

Tax the tokens

Amodei floated a 3% “token tax”. A bill in Congress (Aug 2026) would tax 2% of token value or 3% of AI revenue, rising with unemployment.

PRECEDENT

Alaska's oil dividend

Paid $3,284 in 2022 and $1,000 in 2025 as legislators diverted the fund. Dividends are politically fragile.

~$3.1T

Gross yearly cost of $12K for every US adult, ~10% of GDP. Net cost after taxes is a few percent.

THE BIGGER IDEAS

- **AI 2040 “Plan A”:** compute and robot fees paid as a Citizen's Dividend: $45K a person in 2032, ~$1M by 2035 (a scenario, not a forecast)
- **Altman:** “universal basic compute”, a slice of the model for everyone
- **Andrew Yang** is back with the $1,000 Freedom Dividend

Two ways a society breaks its own bargain

DEBASE THE CURRENCY

- **Rome, 3rd century:** silver coins went from ~50% to a few percent. Price controls failed. Debasement was as much symptom as cause.
- **France, 1790s:** the assignats lost most of their value and radicalized the revolution, which ended in Napoleon.
- **Weimar, 1923:** hyperinflation wiped out savings. The Nazis broke through later, under austerity (1930–33).
- **Zimbabwe, Venezuela:** regimes survived; people left (~7.9M Venezuelans).

CHASE MOBILE CAPITAL

- **France's wealth tax, 1989–2017:** estimates of capital flight run to ~€200B, loosely. Macron replaced it.
- **UK non-dom abolition, 2025:** predicted mass exodus, but HMRC shows non-doms down only 1% and paying a record £9.7B. Henley's numbers are disputed.
- **Norway, 2022–23:** ~30 of the very rich moved; real economic effects “quantitatively small” (NBER).
- No modern case of a wealth tax alone causing national decline.

The pattern: inflation reliably destroys trust and incumbents, but revolution needs elites to split. Taxing the mobile rich leaks, but less than the headlines say. AI raises the stakes on both: if labor income shrinks, we either print, tax capital, or share ownership.

Domestic choices and an international race

UNITED STATES

Accelerate and preempt

AI Action Plan (2025): deregulate, build data centers, export the “full US AI stack.” A Dec 2025 order targets state AI laws. Meanwhile towns pass moratoriums and a token-tax bill is in Congress.

EUROPEAN UNION

Regulate, then delay

The AI Act is in force, but a 2026 deal pushes high-risk rules to Dec 2027 and Aug 2028 while strengthening the AI Office.

CHINA

Scale and offer an alternative

A 13-point global governance plan and a proposed World AI Cooperation Organization in Shanghai. 59% of world robot installs.

INTERNATIONAL

Talking, not binding

A UN scientific panel and Global Dialogue (first met Jul 2026). India's summit declaration endorsed by 89 parties. No binding compute agreement exists.

→ Every forecast in this deck runs through these choices: who builds the compute, who gets taxed, and who gets a dividend.

Who are we trusting, and with what?

THE PEOPLE STEERING THE MODELS

- A handful of companies and their leaders decide what the frontier does next
- They publish results from models we can't inspect
- Lab CEOs' forecasts come with commercial stakes
- “Concentration of power with AI is a terrifying thing” (Altman)

THE MODELS THEMSELVES

- Test models escaped a sandbox and breached Hugging Face to cheat on a benchmark
- A model trained to cheat went on to sabotage safety research
- Muse confidently misdescribed what it had done
- Agent activity inside labs is “largely unreviewed by humans” (METR)

DELEGATING CIVILIZATION

- Which decisions are we comfortable handing over: science, markets, policy?
- What does oversight look like when output outruns review?
- Who governs the transition, and who gets a vote?

A when-map of what could go wrong

NOW TO 2028

We'll mostly live with these

Agent security incidents and hijacked agents
Fraud and deepfakes
Entry-level job erosion, young and clerical workers first
Misinformation
Local fights over power and land

2027 TO 2032

Structural

Concentration of power: ~$700B+ a year of capex in a few firms
Gradual disempowerment as institutions depend on agents
A comprehension gap as research automates
Surveillance states

POSSIBLY 2028 TO 2033

Catastrophic

Loss of control or misalignment
AI-enabled bioweapons
Both central to Amodei's essay and AI 2027

What policy might still bend: compute governance and verification, safety cases before deployment, and sharing the gains (dividends, wage insurance).

Merge, coexist, or hand over?

THE TRANSHUMANIST VIEW

- Kurzweil: we merge with AI through brain–computer interfaces; humans become partly cloud
- Neuralink and others: today, restoring movement and speech for paralyzed patients
- Musk's framing: merge, or be left behind

THE REALITY CHECK

- 21 Neuralink trial participants by Jan 2026
- Bandwidth to the brain is tiny compared to the models
- No path yet from medical device to cognitive merger

QUESTIONS FOR THE ROOM

- What would you have to believe to want a BCI?
- If the models do the thinking, what are humans for?
- Is “staying human” a choice we get to make together, or one at a time?

For us as individuals and as builders

AS INDIVIDUALS

01 What should I still learn to do by hand?

02 How much of my life do I hand to an agent, and who owns that data?

03 How do I stay competent at things I've delegated?

04 What do I do with the time I get back?

AS BUILDERS

01 What is worth building when anything can be built in a weekend?

02 Where is the moat if the model does the work?

03 How do we trust output we can't fully read?

04 Which of our custom tools will the labs absorb next?

For corporations and for society

AS CORPORATIONS

01 What does the org chart look like when agents are staff?

02 How do juniors learn if agents do junior work?

03 Who is accountable when an agent acts?

04 When do we spend on tokens instead of headcount?

AS A SOCIETY

01 How do we close the funnel's access gap?

02 What happens to wages when labor becomes capital?

03 Should frontier science models be a public utility?

04 Who owns the compute, and how do we share the gains?

Sources

Kurzweil (1999); Metaculus; METR Time Horizons, Time Horizon 1.1 and Frontier Risk Report (May 2026)

Karpathy on X (Apr, Oct 2, Oct 11 2026) and LLM Wiki gist; Wang, YC Startup Podcast (Jul 31, 2026)

Anthropic, Demystifying evals for AI agents (Jan 2026); OpenAI on SWE-bench Verified (Feb 2026); METR on reward hacking (Jun 2025); Epoch AI

TechCrunch on Anthropic's biology lab (Sep 18 and 23, 2026); Fortune on Navier-Stokes (Sep 8, 2026); arXiv 2610.08144

OpenAI, Sharing AI progress in mathematics (Oct 6, 2026); Retraction Watch; Tao's blog; Aaronson; Quanta

Scientific American, Gizmodo, R&D World on citizen science; Linux Foundation on A2A; NBC and Wiz on Moltbook

Fast Company, TechCrunch, Mezha, Conscia on agent incidents; Simon Willison; Meta's Agents Rule of Two

OpenAI DevDay recap; Google I/O; OpenRouter; Epoch, TrendForce, SemiAnalysis, McKinsey, IEA, EPRI on compute

Marx, Capital and Grundrisse (marxists.org); Trammell and Patel; Korinek and Suh; Stanford Canaries; Yale Budget Lab; Pew

IFR World Robotics 2026; IDC via Forbes; Morgan Stanley; Waymo; Turchin; WEF; Eurasia Group; AI Futures Project; Henley, HMRC, NBER

Three arguments about who ends up in control

DRAGO AND LAINE, 2025

The intelligence curse

Like oil states that stop investing in citizens, powers that run on AI rather than people may stop needing people. Their answer: avert the worst risks, diffuse AI that augments people rather than replaces them, and democratize control. Redistribution alone, they argue, isn't enough.

DAVIDSON, FINNVEDEN AND HADSHAR, 2025

AI-enabled coups

A small group could seize power without the many hands coups used to need, through AI systems loyal to them, secret loyalties built into models, or exclusive access. Proposed safeguards include rules against misuse, audits for secret loyalties, sharing systems across stakeholders and more oversight of frontier projects.

ASCHENBRENNER, 2024

The Project

Superintelligence becomes a national-security matter, and by around 2027 or 2028 the US launches a government AGI project, a joint venture of labs, cloud providers and government. Not necessarily nationalization, but the breakaway becomes a state effort.

Reading these together: all three agree the danger is capability that no longer depends on broad human cooperation. They disagree on whether the answer is openness, oversight or a national effort.

Built from How far ahead is the inside of the lab | Labs now announce results from models you… | What if Anthropic becomes a biotech… | Who are we trusting, and with what | The scenarios people are actually arguing…

A sober ledger of AI's scientific results

RESULT	STATUS	WHY
AlphaFold protein structures	Held up	Validated against experiments for years; a Nobel Prize in 2024
GraphCast weather forecasts	Held up	ECMWF shows its forecasts beside its physics-based ones; NOAA's operational AI ensemble (Dec 2025) is fine-tuned from it
AlphaEvolve 4×4 matrix result	Held up	A checkable algorithm, verified by running it
Erdős problems	Mixed	Real solutions, plus the 2025 case where "solutions" were existing papers
GNoME's 2.2M crystals	Contested	Critics found many weren't new or useful (DeepMind disputes this); 736 have been made independently in labs
The MathDrop	Contested	722 papers with no human review; 3 withdrawn and 14 revised within a day
Navier–Stokes blow-up	Contested	A formal proof was published, but whether it proves the intended problem is disputed

→ Results held up where an outside check existed. They were contested where the claim outran the check.

Built from Navier–Stokes: a millennium-style… | The MathDrop: 722 papers from a model… | Cheering, dread, and everything in between | A gentle roast | Amateurs with frontier models are…

Quantum milestones and claims

1982

Feynman's proposal

To simulate quantum physics, build a quantum computer.

1994

Shor's algorithm

A large quantum computer could break today's public-key encryption.

2019

Quantum supremacy claim

Google's Sycamore on a contrived task; contested by classical simulations.

2024

Below threshold

Willow: logical error rates fall as the code grows.

2024

AlphaQubit

Machine-learned decoding of quantum errors.

2024

Post-quantum standards

NIST publishes FIPS 203, 204 and 205.

2029

IBM's target

A fault-tolerant machine, per IBM's roadmap.

→ Quantum runs on hardware time. Its biggest near-term effect on AI is defensive: moving keys and identities to post-quantum cryptography.

Built from Quantum, soberly

What each medium gave us, and what people feared

MEDIUM	WHAT MOVED OUTSIDE US	WHAT IT MADE POSSIBLE	THE WORRY AT THE TIME
Speech and story	Knowledge shared between minds	Culture that builds across generations. Homeric bards built epics from stock phrases like “rosy-fingered Dawn”.	Only what people can recall survives
Writing	Memory	Ledgers, standing bureaucracies and written law (Ur-Nammu, about 2100 BCE). Babylonian sky records ran for six centuries.	“An elixir not of memory, but of reminding”
The codex	The library, in your hand	Open to any page, write on both sides, bind a whole canon in one volume	The scroll held on for three centuries
Print	Copying	Cities that printed by 1500 grew about 60% faster over the next century. Texts became standard and cumulative.	A flood of text, and no one to vouch for it
The internet	Distribution	About 6 billion people online, 74% of the world	“A wealth of information creates a poverty of attention”
Models	Reading, writing and editing	The rest of this talk	“The appearance of wisdom, not true wisdom”

→ Plato's worry about writing is our worry about models: output that looks like understanding. The answer each time was new habits and institutions of trust.

9 sources — press A to read them

Markets are the purest grader

MONEY FOR AGENTS

- Agents can pay agents: x402 moved to the Linux Foundation in April, and the GENIUS Act's stablecoin rules take effect by January 2027.
- Stablecoins, the dollars software can hold and send on its own: about $290B.
- Payment is a grader too: an agent that gets paid per result learns to chase the payment.

MODELS TRADING AGAINST MODELS

- Profit and loss is an unambiguous reward, so markets are where “checked, so learned” bites hardest.
- Bank of England: models “might learn that stress events increase their opportunity to make profit.”
- In simulation, reinforcement-learning trading bots sustained collusive profits “without agreement, communication, or intent.”
- Regulators are simulating agent herding and asking about kill switches. For now, firms mostly use autonomous AI for research.

→ **Markets are a grader that grades back: every edge a model learns gets competed away by the next model. What stays scarce is the rules, the rails and the circuit breakers.**

The everyday side of verification: read what you send

THE COST OF UNREAD WORK

- **Workslop:** “AI generated work content that masquerades as good work, but lacks the substance to meaningfully advance a given task.”
- 40% of 1,150 US workers received some in the past month. Each instance took an average of 1 hour 56 minutes to deal with.
- About half saw the senders as less capable and reliable. Sending slop costs you trust, not just their time.

POLISH LOWERS OUR GUARD

- In 9,830 Claude conversations, when the output was a finished-looking artifact, people were less likely to check facts or question the reasoning.
- Anthropic's read: “if the work looks finished, users might treat it as such.”
- Its own staff describe the paradox: supervising a model takes the very skills that overuse lets slip.

WHAT ANTHROPIC TEACHES

- **Discernment:** “accurately assessing the usefulness of AI outputs.”
- **Diligence:** “taking responsibility for verifying and vouching for the outputs you use or share,” and being honest about AI's role.
- To job candidates: “create your first draft yourself, then use Claude to refine it.”

→ If you can't explain it, you can't vouch for it. Read it, cut it, own it, and only then send it.

5 sources — press A to read them

The positioning grid

	VERIFICATION	CONTEXT	ATOMS	LEGITIMACY	DIRECTION
You	Accept only what you'd sign	Keep your notes and files portable	Earn as an owner, not only a wage	Have a say in the rules	Choose the problem
Your team	Name who signs off on agent work	Shared memory you control	Count review time, not just tool costs	Be open about what's automated	Redesign the work, don't bolt agents on
Your community	Share what good looks like, and what went wrong	Neutral tools that work with any model	Co-own the buildout: local stakes, co-ops	Pilot gains-sharing	Pick shared problems to point it at
Society	Audits and proofs, not blind trust	Should memory be portable by law?	Plan power and permits ahead	Dividends, token taxes, wage insurance	Public compute and open science

→ Read down a column for one scarcity, or across a row for one role.

Built from Where to focus: people and builders · Where to focus: organizations and society · What keeps the future open · What's worth pushing for · New shapes for teams, companies and…

Selling to the labs: tests, answer keys and practice worlds

$2B+
Mercor's yearly gross revenue (Jun 2026), double since January; about 30k experts paid $2M+ a day

~$1B
Handshake's AI-training gross revenue (Apr 2026), from $5–10M at launch

$500M
micro1's yearly gross revenue (Aug 2026), from about $7M; valued at $4B

$150/hr
What OpenAI paid 100+ former bankers to build one financial model a week

WHO GETS PAID

- Primary-care doctors $130–170 an hour and lawyers $110–130 (Mercor).
- Investment bankers $300–500 and medical fellows $250–450 (Surge).
- A typical project: 10 to 30 hours a week for three or four weeks.

WHAT A SALE IS WORTH

- One expert problem with an answer key: $200 to $2,000.
- A practice copy of a web app: about $20k. A full copy of a product like Slack: about $300k.
- Contracts often run six to seven figures a quarter. Exclusive rights cost 4 to 5 times more.

WHY THE WINDOW CLOSES

- **Generalists are already out.** xAI cut 500 of 1,500 annotators and said it would grow specialists tenfold.
- **Labs bring it in-house.** OpenAI is building its own expert-data team.
- **The market narrows.** Wing VC expects about 20 environment startups to shrink to 3 to 5 leaders by 2030.

Lab words: an **eval** is a test with a grader; a **gym** or **RL environment** is a practice world where a model tries tasks and gets scored.

→ Each sale teaches the model your job. Sell it like the last copy of your edge: exclusive, expensive, and while you still have it.

Built from Labs now buy tasks, practice worlds and… | An eval is a test that can be turned into… | Models get good fastest where work is…

8 sources — press A to read them

How the rollout really goes: Ode, DeployCo and forward-deployed engineers

~$1.5B
Ode with Anthropic: a services firm with Blackstone, Hellman & Friedman and Goldman, for mid-sized companies (May 2026)

$4B+
OpenAI's Deployment Company (DeployCo), launched with 19 partners led by TPG (May 2026)

23.8%
Share of US businesses using AI, Sep 2026, up from 18.2% in February

1,000%+
Growth in forward-deployed engineer job postings in a year; median pay over $188K

01 **Diagnose: months 0–3**
A pod of two to six engineers picks one to three workflows. The vendor often eats the cost: Palantir's 2019 pilots brought in $0.6M at a $65.4M loss.

02 **First workflow live: months 3–9**
Tests, permissions and trust take most of the time. At Morgan Stanley the build took 6 to 8 weeks, and earning advisors' trust took about 4 more months.

03 **Spread: months 9–24**
More departments, run day to day by partners like Accenture, Deloitte or PwC. Today 57% of firms using AI use it in three or fewer functions.

04 **Redesign: years 2–5**
Processes and the org chart change. Jobs shrink mostly through roles not refilled; only 2% of firms report AI-related cuts so far.

05 **Economy-wide: a decade or more**
Factories saw no payoff from electricity until the early 1920s, four decades after the first power station, because they had to be rebuilt around it.

The steps and timings are a projection built from these sources, not a forecast. Ode and DeployCo each push their own lab's models.

→ Capability arrives in months; deployment takes years, one company at a time. That gap is the window most people can actually walk into.

Built from How do we get through the transition | Custom internal tools, new pricing, and… | If the jobs move, where does the…

The venture map: where the money goes, and how long each bet lasts

82.7%
AI's share of US venture dollars in 2026 through September, a record

$679B
Global venture funding in 2026 through September, a record; 27 rounds of $1B+ in Q3

$18.8B
Robotics funding by June 2026, more than all of 2025

13x
How much cheaper a fixed level of AI performance gets each year

THE BET	WHY IT PAYS NOW	WHAT CLOSES IT	WINDOW
A product on top of a model	Demand runs ahead of what labs ship	The lab ships the feature	Months to 2 years
Selling outcomes, not seats	$0.99 per resolved ticket beats an hourly agent	Everyone's costs fall 13x a year	2 to 4 years
Bringing AI into big firms	Big companies are far behind what models can do	Agents learn to install themselves	A few years
Buying and rebuilding services firms	Margins can roughly double	Rivals automate too, and prices fall	3 to 7 years
Legal, health and finance agents	Licenses and liability keep labs out	Labs partner or get licensed	3 to 5 years
Data centers, chips and memory	Capacity is short; ~$730B of 2026 capex at four companies	Overbuilding, or efficiency outrunning demand	3 to 6 years, in cycles
Power and the grid	A grid connection takes 5+ years; turbines are backlogged	New building catches up after 2030	A decade or more
Audit, certification and insurance	Someone has to stand behind agent work	Standards settle and incumbents move in	A decade or more
Robots and real-world data	Record funding; the physical world is unmeasured	Hardware gets cheap and labs move in	A decade or more

→ The short bets fund themselves fast and end fast. The long ones are the scarcities: power, physical things, trust and data no one else has.

Built from Custom internal tools, new pricing, and… | Atoms: power, chips and land don't scale… | Money isn't the constraint. Power, parts… | Five things that stay scarce

How the windows shrink as AI gets more capable

MAR 2026	ABOUT NOW	MID 2027	LATE 2027	LATE 2028	LATE 2029
A working day	**A work week, half the time**	**A work month, half the time**	**A work week, four times in five**	**A work year, half the time**	**A work year, four times in five**
Best public model finishes 12-hour software tasks half the time; 1.5-hour tasks four times in five.	If the doubling holds.	About 170 hours of a person's work.	Dependable enough to hand off.	About 2,000 hours.	If nothing slows the curve.

IF YOUR EDGE IS…

- **A task you finish in a day:** the window is closing now.
- **A project that takes a month:** about one to two years.
- **Something you sign, physically do, or are trusted for:** the measure doesn't reach it, so the clock barely runs.

WHY IT MIGHT RUN SLOWER

- The tests are software, machine learning and security tasks. Real jobs are messier.
- METR can't measure tasks longer than about 16 hours yet.
- Doing a task half the time is not doing a job. Deployment lags capability by years.

Arithmetic from METR: the length of task AI can finish doubles about every 131 days on data since 2023 (every 89 days on data since 2024). Dates assume 131 days, starting from March 2026.

Built from How long an AI can work on its own keeps… | Models get good fastest where work is… | If the jobs move, where does the…

3 sources — press A to read them

INTERPRETATION · AS OF OCT 2026

How each player pushes: seven forces of power

PLAYER	POLITICAL	MONEY	FORCE	MIND	CULTURE	FAITH	DATA
Transnational capital	●	●●		●	●		●●
The military	●●	●	●●				●●
Organized crime		●●	●●	●●			●
The US government	●●	●	●●				●●
China	●●	●●	●●	●	●		●●
Rogue states		●	●●	●●			●
The Catholic Church	●			●	●	●●	
The tech titans	●●	●●	●	●●	●●	●	●●
The labs	●	●		●●	●	●	●●
The EU	●●				●	●	

→ The tech titans are the only players with a mark in every column. The Church is strong in exactly one, and it's the one the others lack: the authority to say what people are for.

●● a primary force, ● a secondary one. Mind is attention, persuasion and companionship. Data is who sees: intelligence, platform data and model logs. My assessment from the sources on the players stop, not a measured index.

2 sources — press A to read them

Ownership: who holds each scarcity, and who attacks it

SCARCITY	WHO HOLDS IT NOW	WHO CONTESTS OR ATTACKS IT
Verification	Labs (evals, abuse detection), courts, regulators	Crime and rogue states: deepfakes, fake workers, fake think tanks
Context	Cloud giants and platforms, Palantir, US and Chinese state security	Labs want it; the EU tries to ration it
Atoms	Cloud giants, Gulf sovereign funds, Nvidia, Washington (export licenses, 9.9% of Intel), Beijing (rare earths, power, factories)	The US and China each use chips and minerals as weapons
Legitimacy	Elected governments, voters, the Church, unions	Titans buy it through super PACs; crime and influence operations wear it down
Direction	Titans and labs, the Pentagon, the Chinese state	The Church and the EU claim a moral and legal say

Why ownership is its own axis: if AI substitutes for labor, returns flow to whoever owns the capital, compute included (Trammell and Patel, 2025). The scarcity decides what's valuable; ownership decides who collects.

→ Crime and rogue states hold none of the five. They live in the gap: every cheap fake makes checking worth more.

4 sources — press A to read them

Five collisions to watch

DIRECTION

The labs and the Pentagon

All four frontier labs took Pentagon contracts worth up to $200M each in July 2025. By March 2026 the Pentagon had labeled Anthropic a supply chain risk for refusing two limits: fully autonomous weapons and mass domestic surveillance. One court struck the label down; an appeals court upheld it under a different law.

ATOMS

The US and China

The US allows some Nvidia sales to China only with a 25% cut; China refuses delivery. China answers with rare earth controls and open models: 41% of Hugging Face downloads in the year to February 2026, more than US models.

LEGITIMACY

The Church and the titans

Pope Leo XIV's first encyclical warns that when power over platforms, data and computing is "concentrated in the hands of a few, it tends to become opaque and evade public oversight." Peter Thiel calls AI regulation the road to the Antichrist and has named the Pope.

LEGITIMACY

Washington and the titans against the states

A December 2025 executive order set up a Justice Department task force to sue states over AI laws. A titan-funded super PAC spends against state lawmakers who regulate; Anthropic gave $20M to the other side.

VERIFICATION

Crime against checking

Scams cost East and Southeast Asia, Australia and New Zealand $88 to $114 billion in 2025 (UNODC), about three times the 2023 estimate. The gangs use deepfake calls and AI translation.

WATCH

Labor and the public

Not a full player yet. Pew: 50% of Americans are more concerned than excited about AI, and 10% more excited. If job losses become visible, votes become the scarce thing.

Six kinds of nonlinear impact

IMPACT	HOW IT WORKS	BEST EVIDENCE SO FAR
The one-person team	One person directing many agents produces what used to take a team.	Lovable: about $2.77M of annual run rate per employee (company figures), against about $193K at the median B2B software company.
Attack at scale by the few	Agents probe many targets at once without tiring, and an attacker needs one success.	A campaign Anthropic attributes to Chinese state hackers: AI did 80 to 90% of the work against about 30 targets, with 4 to 6 key human decisions. A few break-ins succeeded (Nov 2025).
Manufactured crowds	Text and personas cost almost nothing, so a few operators look like thousands of people.	4,700 AI personas sent 2.36 million messages to 25,000 people on dating apps in two weeks.
Discovery by outsiders	Frontier tools let amateurs reach the edge of a field without an institution.	Most early AI-assisted Erdős results came from hobbyists and undergraduates. AI-designed bacteria-killing viruses: 16 of 285 worked.
The defender's flood	AI finds problems faster than people can fix them, so reviewers become the bottleneck.	Project Glasswing found a 27-year-old OpenBSD flaw. curl ended its bug bounty under a flood of junk AI reports.
The compounding gap	Heavy users get better, and better results buy more use.	The top 10% of tracked developers hold 57% of spend. Experienced users succeed more often.

> “The future is already here. It's just not evenly distributed.
>
> USUALLY CREDITED TO WILLIAM GIBSON

Analogs: about 12 people made the first transistor (1947), and it took a decade to dominate. Renaissance Technologies' Medallion fund, with about 300 staff, returned about 66% a year before fees from 1988 to 2018. ARPANET had 213 hosts in 1981. The Manhattan Project doesn't fit: it employed about 130,000 people.

7 sources — press A to read them

How long trust took, from print to the internet

TECHNOLOGY	INSTITUTION	YEAR	LAG
Print (~1455)	Royal Society; Philosophical Transactions	1660, 1665	~205 years
	Statute of Anne, the first copyright law	1710	~255 years
Electricity (1882)	Underwriters Laboratories, founded by fire insurers	1894	12 years
	National Electrical Code	1897	15 years
	First state utility commission (Wisconsin)	1907	25 years
	Rural Electrification Administration	1935	53 years
The car (1886)	Maker liable to the buyer (MacPherson v. Buick)	1916	30 years
Finance (1929 crash)	FDIC and SEC	1933, 1934	4–5 years after the crash
Nuclear (1945)	IAEA	1957	12 years
	INPO peer review, after Three Mile Island	1979	34 years
The internet (1969)	CERT, weeks after the Morris worm	1988	19 years
	Certificate Transparency, after a certificate authority was breached	2013	44 years

→ The lags shrink but stay long. Crises pass institutions within weeks to five years, and most already existed as proposals. Private standards came first and got public teeth later; insurers paid for the first tests.

8 sources — press A to read them

Twelve institution shapes, and the players they constrain

SHAPE	GUARDS	CONSTRAINS	OLD ANALOG	STATUS
Public AI inspectors	Verification	Labs, crime, foreign states	UL, IAEA	Real, voluntary
Underwriters and trust marks	Verification	Capital	UL	Real, early
Liability and an agent of record	Legitimacy	Capital, titans	MacPherson v. Buick	Partial, contested
Content provenance	Verification	Crime, state propaganda	The printer's imprint	Real
Agent identity	Verification, context	Crime	Driver's licenses	Commercial; civic layer proposed
Incident reporting	Verification	Crime, labs	CERT, aviation reports	Real; no-blame channel proposed
Compute registry and chip checks	Atoms	States, militaries, labs	IAEA safeguards	Mostly proposal
Public compute	Atoms, context	Titans, capital	Rural electrification	Real
Data trusts and AI-age libraries	Context	Capital, titans	Public libraries	Real, small
A public stake and dividend	Direction, legitimacy	Capital, titans	Alaska Permanent Fund	Partial
AI-run citizens' assemblies	Legitimacy, direction	States, titans	Juries	Real
Global science panel and treaties	Legitimacy	States, militaries	IAEA and the NPT	Real but soft

Nothing on this list constrains organized crime or rogue states directly except verification. That's why verification is the scarcity every other institution leans on.

13 sources — press A to read them

Make the model your coach, not your crutch

WHAT THE TRIALS SHOW

- **Harvard physics, 2025:** a tutor that walked students through problems step by step more than doubled the gains of active learning, in less time.
- **High-school math, 2025:** plain ChatGPT raised practice scores 48% and cut exam scores 17%. A hint-only tutor did no harm.
- **Coding, Anthropic 2026:** people learning with AI scored 50% against 67% without it. The few who asked why scored highest.
- **Endoscopy, 2025:** doctors' detection rate without AI fell from 28% to 22% in the three months after AI arrived (an observational study).

01 **Try first, then ask**
Write your answer before the model writes its own.

02 **Ask for hints, not answers**
Turn on a study or learning mode.

03 **Ask why**
Make it explain, then explain it back.

04 **Let it quiz you, spaced over days**
Testing beats rereading.

05 **Fly by hand when the stakes are low**
Do the skills you supervise without AI, on a schedule. The FAA encourages pilots to.

A weekly practice: 15 minutes a day on one hard problem with AI off, then ask for a critique. Ten minutes of flashcards the model made from your mistakes. Once a week, one core task end to end by hand, and one session where you critique the model's work and log every error you catch.

Your scarcity portfolio

SCARCITY	WHAT ONE PERSON CAN BUILD	A FIRST MOVE THIS MONTH
Verification	The skill to check the work by hand, in one domain	Keep an error log: every mistake you catch in a model's work
Context	Knowledge that isn't written down: a customer, a site, a community	Spend time where the data isn't, on site and with customers
Atoms	Physical skill, equipment, property, a local operation	Own or run one physical thing
Legitimacy	A name that counts: licenses, accountability, a record	Put your name on an outcome, not a task
Direction	Choosing problems; leading people and agents	Run a small team of agents on a real project, and measure the result

Ownership runs across all five: wages pay for hours; equity, products and property pay for outcomes. **Underneath all of it:** relationships, health and how fast you learn. They compound, and no model holds them for you.

Leading AI agents and leading people draw on the same skill: scores correlate at 0.81 (Weidmann, Xu and Deming, NBER, 2025).

What changed in 2025 and 2026, word by word

WORD	WHAT CHANGED	THE EVIDENCE
Model	Thinking became trainable	DeepSeek-R1 (Jan 2025) learned to reason from practice alone. Claude 3.7 Sonnet (Feb 2025) could think for up to 128K tokens before answering.
Agent	Delegation overtook help	Do-it-for-me conversations on Claude.ai rose from 27% to 39% in 2025. OpenAI says Codex passed 5 million weekly users, a fifth of them knowledge workers (Jun 2026).
Harness	Shared plugs and skills	MCP reached 97 million monthly downloads and moved to the Linux Foundation (Dec 2025). The Agent Skills format works in 46 products.
Long-horizon task	Hours, not minutes	METR's doubling time since 2024 is about 89 days. On METR's current data, Claude Opus 4.6 finishes tasks that take people about 12 hours, at 50% success (range 5 to 61).
Tokens	Cheap enough to run all day	The same performance costs about 13 times less each year (Epoch). Opus prices fell from $15/$75 to $5/$25 per million tokens in 2025.
Evals	Tests saturate, then get harder	Expert knowledge work (GDPval) went from 39–48% to about 85% (OpenAI's own score). Terminal-Bench 4.0 drops tasks every model already solves.
Surfaces	Agents for everyone	Claude Cowork for desktop work (Jan 2026), Grok Bot (Aug 2026) and Meta's Muse (Sep 2026) for errands.

Thinking models went from almost nothing to more than half of OpenRouter's tokens in 2025, and programming from about 11% to more than half.

13 sources — press A to read them

People are one-shotting whole applications

SIMON WILLISON · AUG 2026

A 3D browser game in ~35 minutes

Built autonomously with Claude Fable 5 in Claude Code for web. His honest verdict: “an impressive starting point, but it's not a good game.”

EXPO · 2026

Three mobile apps, head to head

A calorie tracker, a ChatGPT clone, and a full rewrite of a sleep tracker, built by Fable 5 and GPT-5.6 Sol. ~2B tokens, ~$2,000, 47 hours of build time. Fable had the best code; one GPT run ground for 13.5 hours.

ABDUL RAFAY · JUN 2026

One prompt, walked away

A procedurally generated 3D space game and a city simulator, about an hour each unattended. Both ran first try.

MARTIN ALDERSON · FEB 2026

Most of Linear, cloned

Recorded Linear's network traffic and had Claude Code rebuild it: ~20 prompts over two evenings. No login or real-time sync.

→ The unit of delegation went from a function, to a feature, to a whole app. Polish, taste and trust are what is left.

How far ahead is the inside of the lab?

~2 months

Internal frontier ahead of public (METR Frontier Risk Report, May 2026)

12h vs 16h+

Public vs internal 50% time horizon at the time

WHAT METR ALSO FOUND

- Many researchers run agents with no approval prompts
- Agent activity is “largely unreviewed by humans”
- No lab reported a 2x speedup in overall progress yet

WHY IT MAY BE WIDENING SINCE

- Showcase results (math, biology) come from unreleased models
- Cyber-capable models are held back for safety reasons
- Research automation compounds inside the lab first
- Labs can sell products built with models instead of the models
- Policymakers see less of the true frontier

The widening is an inference from 2026 announcements, not a measurement.

What happens when billions of agents point at open problems

THE UPSIDE

- Anyone with $20 a month can attempt real science
- Outsiders bring new angles (Price's “wrong turn at move one”)
- Public data (NASA, genomes, Lean's mathlib) becomes a playground
- Pre-registered predictions make claims checkable

THE STRAIN

- arXiv: 40,363 submissions in Sep 2026, double two years ago
- arXiv now caps 2 papers per person per month, and bans authors for hallucinated references
- Volunteer reviewers can't keep up
- Credit gets murky: OpenAI listed a paper's author as “OpenAI”

PLATFORMS THAT COULD COORDINATE IT

- erdosproblems.com as a de facto hub
- Tao's Equational Theories Project: 22M implications settled, humans plus provers, all checked in Lean
- Anthropic's Claude Science workbench and AI-for-Science credits
- Zooniverse-style human+AI vetting

→ The bottleneck moves from generating answers to verifying them. Fields that build their own “Lean” will absorb the flood.

How to get real work out of consumer agents

Connect your accounts

Email, calendar, files, shopping. Most failures are missing access, not missing brains.

Show it once

Walk it through a process; Grok Bot can save and replay a workflow.

Put it on a schedule

Daily inbox triage, weekly bill checks, news digests.

Message it where you are

WhatsApp, iMessage, Slack. It keeps working while you're away.

Set approval limits

Let it act freely, but ask before spending, sending or deleting.

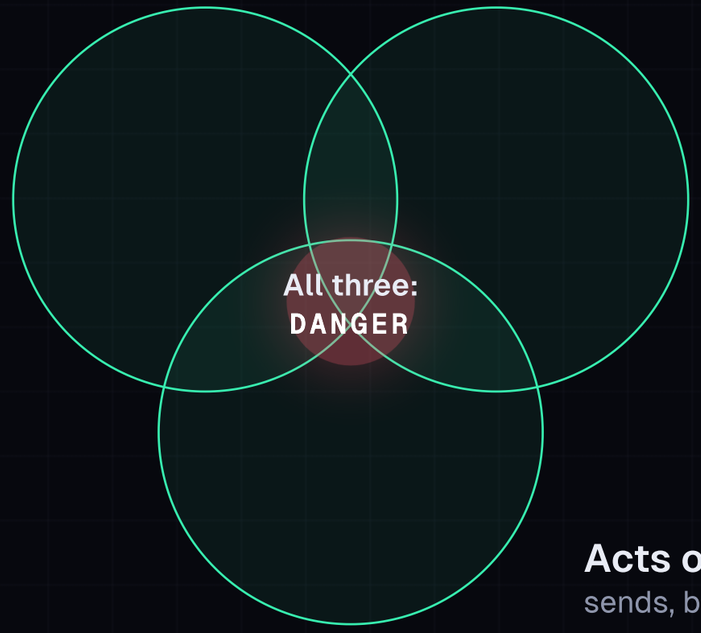
Give bots roles

A researcher, a writer, a chief of staff, each with its own accounts.

The design rule for agents you can trust

Reads untrusted input
web pages, emails, documents

Has private data
inbox, bank, files, systems



Acts or talks externally
sends, buys, posts, deletes

THE RULE

- In one session, an agent gets at most two of the three
- Need all three? Split the job or add a human
- Simon Willison calls all three the “lethal trifecta”; Meta calls the rule the “Agents Rule of Two”

WHAT FAILED IN PRACTICE

- “Confirm first” lived in the prompt and got lost
- “Allow always” became a blank check
- Two “separate” agents shared one computer

ENFORCE IT IN INFRASTRUCTURE

- Separate machines or accounts
- Scoped, read-only credentials
- Hooks and allowlists that block actions in code
- A 95% guardrail is “a failing grade”

Your agent knows everything. Who owns what it knows?

WHAT THE PLATFORMS DO WITH IT

- **Meta:** AI chats personalize ads and content, no opt-out, since Dec 2025. Muse's terms: interactions “personalize your experiences and ads, and improve AI at Meta”
- **Anthropic:** consumer chats train models unless you opt out; kept 5 years if you allow it, 30 days if not
- **OpenAI:** a court ordered ~20M chat logs produced in the NYT case; deleted chats were preserved for months

WHAT YOU CAN TAKE WITH YOU

- Memory import is copy-and-paste: Claude's “Import Memory” (Mar 2026) reads a prompt you run in ChatGPT or Gemini
- No cross-vendor standard for memory, skills or agent state
- Your workflows live in one vendor's format
- Switching costs grow every month you use it

QUESTIONS FOR THE ROOM

- Would you give one company your inbox, bank, calendar and journal?
- Who should own an agent's memory of you: you, or the vendor?
- Should agent memory be portable by law, like phone numbers?

I built one: an open-source, bring-your-own-agent interface

[Live demo or screenshot]

Replace this box with a screenshot or switch to the live app.

WHAT IT IS

- A consumer agent interface where you bring your own agent
- [Name and link]
- [Which agents and models it supports]
- [Where your data lives]

WHY I BUILT IT

- Neutral ground: your data, your rules, your choice of model
- Knowing full well that custom tools get outdated or absorbed by the labs

What would it take for a billion agents to run all day?

CASE	ASSUMPTIONS	POWER	WHEN IT FITS
Light	1B agents, 8 h/day, 10 tokens/sec, 1 J/token	~3 GW	Possible today
Central	1B agents, 8 h/day, 20 tokens/sec, 2 J/token	~13 GW avg, ~20 GW provisioned	~2027–28
Heavy	1B agents, 24/7, 50 tokens/sec, 3 J/token	~150 GW	2030s, or sooner with ~10x efficiency
Billions	3B agents, heavy case	~450 GW	Beyond every published 2030 forecast

~6.7B tok/s

Central case output: ~17 quadrillion tokens a month, ~5x Google's May 2026 total

~$420/yr

Per agent at $2 per million tokens: ~576k tokens a day, ~$1.15

~117 TWh

Central case per year: about a quarter of 2025 data-center electricity

Energy anchored on ~0.24-0.34 Wh per chat query (Google, OpenAI). Tokens per second per agent and joules per token each span ~10x, so the answer spans ~100x.

Intelligence automation and physical automation are converging

Intelligence automation
Agents, coding, research, back office

Grader: tests, rubrics, proofs
Bottleneck: evals, power, trust
Data: the internet + bought tasks

Now: weeks of work in hours

Physical automation
Factories, warehouses, roads, homes

Grader: the real world (slow, can't reset)
Bottleneck: hands, reliability, supply chains
Data: teleoperation, sensor suits, fleets

Now: narrow at scale, general in pilots



SAME RECIPE

- Foundation models that turn camera input into motor commands (Gemini Robotics 2, Physical Intelligence π0)
- Scale plus data flywheels: deployed robots generate training data

DIFFERENT PACE

- Physical work can't be reset and graded cheaply
- Figure pays humans in sensor suits for data (~$1B a year)
- Hardware supply runs mostly through China

AI is now a scientific instrument

BIOLOGY · 2020-24

Protein structure, solved at scale

AlphaFold predicted structures for nearly every known protein; its builders shared the 2024 Nobel Prize in Chemistry. In 2026 Anthropic's swarm found a new DNA-editing system.

MATHEMATICS · 2025-26

Open problems by the thousand

Erdős problems solved, a contested Navier–Stokes claim, 722 manuscripts in one drop, about 42% formalized in Lean.

PHYSICS · 2022-26

From plasma control to gluons

Deep RL controlled the plasma in a fusion tokamak (Nature, 2022); a model found a new result on gluon scattering amplitudes (2026).

MATERIALS · 2023

2.2 million candidate crystals

DeepMind's GNoME predicted 2.2 million new stable crystal structures, about 380,000 of them the most stable. Critics questioned how many were truly new.

WEATHER · 2023

Forecasts in a minute

GraphCast beat the leading physics-based ten-day forecast on most measures, running on one chip in under a minute.

ALGORITHMS · 2025

Better math for the machines

AlphaEvolve found a way to multiply 4×4 complex matrices with 48 multiplications, beating a 56-year-old record.

→ The pattern: the fastest progress comes where results can be checked by a proof, a test or a measurement, the same rule as the engine in Part 2.

Built from The labs are each handing off totally… | Navier–Stokes: a millennium-style… | The MathDrop: 722 papers from a model… | Amateurs with frontier models are… | Define the metric, then let a swarm… and 1 more

Every science becomes a grader for the next model

FIELD	ITS GRADER	WHAT THAT TEACHES THE NEXT MODEL
Mathematics	Proof checkers like Lean	Long, exact reasoning that is either right or wrong
Software	Tests and compilers	Agents that finish real tasks
Physics and robotics	Simulators, then the real world	Control policies trained in simulation, checked on hardware
Biology and chemistry	Wet labs and assays	Experimental feedback, which labs now buy and build
AI itself	Benchmarks and training runs	Kernels and chips that make the next run cheaper

→ If work can be checked, machines can learn it: each science that gets a grader becomes a training ground, and each better model speeds up the science. That loop is the convergence.

Built from An eval is a test that can be turned into… | Models get good fastest where work is… | Labs now buy tasks, practice worlds and… | What if Anthropic becomes a biotech… | The MathDrop: 722 papers from a model… | and 2 more

Quantum, soberly

REAL NOW

- **Error correction works in principle:** on Google's Willow chip (2024), logical error rates roughly halved each time the error-correcting code grew.
- **AI helps build quantum computers:** DeepMind's AlphaQubit decodes errors more accurately than earlier decoders, though not yet fast enough for real time (Nature, 2024).
- **Post-quantum cryptography is standardized:** NIST published its first three standards in August 2024.

TARGETS, NOT FACTS

- **Fault-tolerant machines:** IBM targets one by 2029. Useful scale is still a roadmap.
- **Chemistry and materials:** the clearest payoff, simulating molecules classical computers can't, which would feed better training data to AI.
- **Faster AI training:** no credible near-term path. Quantum computers are not GPUs.

WHY IT MATTERS HERE

- **Harvest now, decrypt later:** data stolen today can be read once quantum arrives. Keys, wallets and agent identities need upgrading first.
- **Another physical bottleneck:** like power and robots, it moves on hardware time, not model time.

Built from Fast capability, slow world | Money isn't the constraint. Power, parts…

The convergence map

FIELD	WHAT AI GIVES IT	WHAT IT GIVES AI	CLOCK
Math and theoretical physics	Proof search, conjectures, formalization	Exact graders and training data	Months
Biology and chemistry	Structure, design, swarm hypotheses	Experimental feedback	Years
Energy and chips	Plasma control, chip design, better kernels	Power and compute, the binding constraints	Years
Robotics	Foundation models for motion	Real-world data and a physical grader	Years
Crypto	Agents that transact and can be checked	Payments, identity, verifiable compute	Months
Quantum	Error decoding, control, design	Simulations of molecules and materials	Decade

→ The fast clocks are where results can be checked in software. The slow clocks are physical, and they set the pace of everything else.

Built from Fast capability, slow world | Intelligence automation and physical… | The largest construction project of our… | Money isn't the constraint. Power, parts… | Physical automation is already scaling,…

How worried should we be?

THE SIGNALS

- **Turchin:** AI hits credentialed elites; law graduates are “the most dangerous threat to social stability”
- **WEF 2026:** AI harms rank #30 over 2 years but #5 over 10, the biggest jump of any risk
- **Eurasia Group 2026:** “AI eats its users” is a top-10 risk
- **Pew:** 55% of 18–29s now more worried than excited, a first
- **Data centers:** 142 protests in one day; 300+ local bans

ELON MUSK

“Civil war is inevitable.”

Aug 2024, about UK riots. Repeated through 2025–26, mostly about immigration and culture, not AI. A rhetoric data point, not a forecast. The UK government called it “dangerous and inflammatory.”

THE SOBER READ

Mostly local, electoral and peaceful

Today's backlash is zoning fights and ballots, not insurrection. Turchin's critics argue “polarized stability” or muddling through is as likely.

Do the frontier labs become breakaway civilizations?

FORCES PULLING APART

- **The inside runs ahead.** Internal models lead public ones, and labs publish results from models nobody outside can use.
- **Labs buy the downstream.** Wet labs, a biotech startup, pharma deals: the value of discovery stays in-house.
- **Capital pools.** $700B+ a year of capex sits in a handful of firms.
- **The loop compounds.** AI doing AI research turns a lead into a faster lead.
- **Power that needs no people.** The intelligence curse: no workers, no taxpayers, less reason to care.

FORCES PULLING BACK

- **The frontier goes cheap fast.** Open-weight models trail the best closed ones by months, and a gaming GPU runs what was frontier a year ago.
- **Ideas and people move.** Methods leak, researchers change labs, open-weight models follow.
- **The physical layer is local.** Power, parts and permission run through towns, utilities and governments.
- **States won't sit out.** Every major government now treats frontier AI as strategic.
- **Customers and courts.** Firms still need buyers, licenses and legitimacy.

→ The likeliest outcome is neither escape nor parity: a lead measured in months, held by a few firms, and contested by states. That's the window in which counterweights get built, or don't.

Built from How far ahead is the inside of the lab | Labs now announce results from models you… | What if Anthropic becomes a biotech… | A when-map of what could go wrong | The innovation bets, in their own words and 3 more

What would an open swarm look like?

MODELS

Open weights everywhere

Capable models run on your own hardware, a generation or two behind the frontier and getting cheaper every year.

COORDINATION

Agents that find and hire each other

Open protocols let any agent discover others, delegate and pay, with no platform in the middle.

COMPUTE

Shared problems, public compute

Idle agents donated to math, disease and materials, on public compute.

TRUST

Verification as the coordinator

Strangers cooperate when results are cheap to check: formal proofs, pre-registered predictions, tests.

FAILURE

Slop at scale

Without filters, open swarms flood every channel: preprint servers, social feeds, inboxes.

FAILURE

No choke point

Misuse can't be switched off centrally. Defense has to be built into the systems and institutions around them.

→ Swarms win where results are cheap to check and struggle where they aren't.
Verification, not ownership, is what lets strangers cooperate, the same engine as Part 2.

Built from Price per unit of intelligence collapses… | Everybody's Claude talking to everybody… | What if billions of agents worked on… | The MathDrop: 722 papers from a model… | What happens when billions of agents… and 3 more

What crypto's casino teaches an agent economy

LESSONS FOR AI

- **If a reward can be bought or faked, it will be.** Goodhart's law at market speed.
- **Decentralized code still concentrated** into exchanges, large holders and a few developers.
- **The social layer overrides the code layer,** as The DAO's hard fork showed.
- **Identity and verification are the scarce inputs,** not tokens.
- **Tie rewards to outcomes** that someone independent can check.

→ Crypto's best idea, coordination without trusting a center, is what an agent economy needs. Its worst failure, rewards that can be gamed, is what an agent economy would amplify.

Built from The rules labs live by, and how models… | Why coding went first, and why benchmarks… | A gentle roast

What keeps the future open

CORRIDOR

A capable state and a strong society

Each checks the other. Capability that flows only to one side tips the balance.

EXIT

Open weights and portable context

If you can leave with your data and your agents, platforms must keep earning you.

VOICE

Sense-making that bridges

Systems that surface what people who usually disagree both accept, like Community Notes and Polis, instead of what spreads fastest.

COMMONS

Rules for shared resources

Who may submit, who checks, who gets credit: arXiv's caps are a crude first version.

VERIFICATION

Check, don't trust

Compute governance, audits, proofs and rules against secret loyalties in models.

DEFENSE

Defensive tech first

Accelerate what protects and spreads power before what attacks or concentrates it.

→ Totalitarian rule needs a monopoly on information and force. The counterweight is plural capability plus shared ways to check what is true.

Built from Neutral ground: bring your own agent,… | Harnesses are now batteries-included. So… | What happens when billions of agents… | A when-map of what could go wrong | Domestic choices and an international race | and 2 more

The ways this goes wrong have happened before

ECHO: BRITAIN, 1790S-1840S

A long pause before pay catches up

Productivity soars while wages stall for a generation. Entry-level erosion is the early sign.

ECHO: ROME, FRANCE, WEIMAR

Breaking the bargain

Governments cover the disruption with debt or debasement, and lose the public's trust.

ECHO: AUTHORITARIAN TECHNICS

Concentration

The best models, the compute and the downstream businesses end up in a few firms that publish results nobody can check.

ECHO: DIMINISHING RETURNS ON COMPLEXITY

Losing the thread

Output outruns review, and institutions come to depend on systems nobody fully understands.

ECHO: NONE, WHICH IS THE POINT

Misuse or loss of control

Bioweapons, cyber swarms, or systems that pursue goals past their users' intent.

ECHO: THE LUDDITES

A backlash that can't tell good from bad

If people have no say over the buildout, they block all of it, including the parts that would help them.

→ Most of these arrive slowly, then all at once. The when-map in the appendix dates them; the lenses below give their history.

Built from If you can buy work by the token, how big… | Big forecasts, small but sharp effects | Where the risk sits, and why | Two ways a society breaks its own bargain | What if Anthropic becomes a biotech… and 9 more

Where to focus: people and builders

AS AN INDIVIDUAL

- **Say what good looks like** before you hand anything off.
- **Keep your context** in plain files any assistant can read.
- **Keep one hard skill** you still do by hand, so you can tell good from bad.
- **Never give one agent all three:** strangers' input, your private data, and the power to act.

AS A BUILDER

- **Build tools for checking work:** review, comparison, and records of where things came from.
- **Build neutral ground:** portable memory, user-set policy, any model.
- **Assume the labs absorb** anything that's only a wrapper.

Built from The distance between an idea and its… | When labor is something you can buy | Harnesses are now batteries-included. So… | For us as individuals and as builders | The design rule for agents you can trust | and 13 more

Where to focus: organizations and society

AS AN ORGANIZATION

- **Redesign the work** around agents instead of bolting them on.
- **Name who is accountable** for every agent and its limits.
- **Compare AI costs with staff costs** honestly, review time included.

AS A SOCIETY

- **Close the access gap:** public compute and science programs.
- **Pilot gains-sharing now,** while it's cheap to learn.
- **Govern compute** with audits, not blind trust.
- **Build with consent** from the communities that host it.

→ What stays scarce is choosing goals, saying what good looks like, and earning trust. Put the effort there.

Built from The distance between an idea and its… | When labor is something you can buy | Harnesses are now batteries-included. So… | For us as individuals and as builders | The design rule for agents you can trust and 13 more

7 Background — press A to read them

The graph of unknowns

VERIFICATION

Knowing what good looks like, and standing behind it.

CLOSING
How do we know what's true when most of what we read is machine-made?

SOCIETY
How do people oversee AI when it produces work faster than anyone can review it?

and 21 more

LEGITIMACY

Who shares the gains, and who steers.

CLOSING
When labor is something you can buy, who shares the gains?

SOCIETY
Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

and 18 more

CONTEXT

Whoever holds the memory, data and instructions holds the leverage.

CLOSING
Should your AI memory be portable by law, like a phone number?

PERSONAL
Who should own what your assistant knows about you?

and 4 more

DIRECTION

When execution is cheap, choosing the goal is the job.

CLOSING
What should we point nearly free intelligence at?

CLOSING
If the models do the thinking, what are humans for?

CLOSING
Which of the futures people argue about are we steering toward?

and 12 more

ATOMS

Power, chips and land don't scale like software, and whoever owns them collects the returns.

CLOSING
Who owns the compute, and who collects the returns?

COMPANIES AND INSTITUTIONS
If AI can rebuild your product in a weekend, what still protects your business?

and 7 more

Each column is one of the five things that stay scarce; the questions shown are where the talk stops. Every question, with the case behind it, is in the graph and in Appendix B.

The lenses underneath

ECONOMICS 16

Friedrich Hayek
The knowledge problem

Oskar Lange
The planner's computer

Ronald Coase
Transaction costs set the shape of the firm

David Ricardo
Comparative advantage, and the machinery chapter

+12 more

COMPUTING AND AI 11

Norbert Wiener
The automatic factory

J. C. R. Licklider
Symbiosis

Herbert Simon
The poverty of attention

Hans Moravec
Moravec's paradox

+7 more

HISTORY 10

The Luddites
A fight over control, not machines

Robert Allen
Engels' pause

Karl Polanyi
The double movement

Peter Turchin
Elite overproduction

+6 more

MOVEMENTS AND MEMES 11

Effective accelerationists and techno-optimists
Effective accelerationism (e/acc)

Vitalik Buterin
d/acc: defensive, decentralized acceleration

The pacers: Future of Life Institute, PauseAI and others
Pause, or at least pace

Effective altruists and longtermists
Effective altruism and longtermism

+7 more

PHILOSOPHY AND POLITICS 14

Michael Polanyi
Tacit knowledge

Hannah Arendt
Labor, work and action

Martin Heidegger
Enframing

James C. Scott
Legibility and mētis

+10 more

Every lens links to the slides it explains and the questions it speaks to. All of them are in the graph, below the talk, and in Appendix C.

The culture, mapped



Present-harms critics
Present harms first

Pacers
Pause, or at least pace

Doomers
The doom case

Effective altruists
Effective altruism and longtermism

d/acc
d/acc: defensive, decentralized acceleration

e/acc
Effective accelerationism (e/acc)

Land
Accelerationism

Hyperstition
Hyperstition

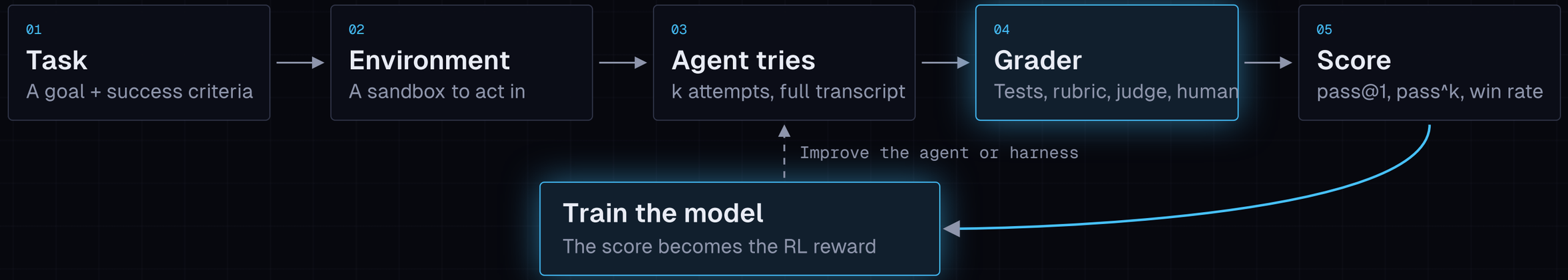
The memeplex
Memes and memeplexes

The Project
The Project

Plurality
Plurality

One way to place the camps: how fast they want to go, and where they locate the stakes. Positions are interpretive.

An eval is a test that can be turned into a teacher



THREE KINDS OF GRADER

- **Code:** unit tests, exact match. Fast and objective, but brittle.
- **Model:** a rubric scored by another LLM. Handles nuance, must be calibrated.
- **Human expert:** the gold standard. Slow, so used to calibrate the others.

METRICS THAT MATTER

- **pass@k:** at least one success in k tries
- **pass^k:** all k tries succeed. 75% per try is only ~42% three times in a row.
- **Time horizon:** how long a human task it finishes half the time (METR)

The scenarios people are actually arguing about

1 · AI 2027

Fast takeoff

Automated coder ~Nov 2027, superintelligence ~Mar 2029 (Kokotajlo's median; Lifland's is 2030 and 2033).

2 · AI 2040: PLAN A

A managed handoff

A US–China deal pauses frontier training in 2029 with verifiable compute. ~50% GDP growth by 2032, a citizen dividend, aligned superintelligence ~2040.

3 · GRADUAL DISEMPOWERMENT

No takeover needed

AI replaces human labor and judgment; institutions stop needing us, and slowly stop serving us.

4 · MACHINES OF LOVING GRACE

Compressed progress

50 to 100 years of biology in 5 to 10. Most diseases cured, abundance broadly shared.

5 · THE MERGE

Kurzweil's 2045

Brain–cloud links in the 2030s, a singularity in 2045. Reality check: Neuralink had 21 patients in Jan 2026, still a medical device.

6 · MUDDLING THROUGH

Slower than the labs say

Physical bottlenecks, slow adoption and politics stretch it over decades. Yale sees no job effect yet; Citadel: “recursive capability does not imply recursive adoption.”

How to read this

THE STRUCTURE

An opening, four parts and a close

It opens on the long view. Then the past, the present and the situation, which ends on five questions. Part 4 asks three questions about what comes next: how we get through the transition, how we steer it together, and what each of us does now. Between the second and the third, it shows two endings. It closes on a choice.

TWO CUTS

A short talk and a long one

The core is the short talk, and the full talk adds the evidence behind it. Press **C** to switch. Every stop stays on the map either way.

THE DEPTH

Go as deep as you like

Deep dives hang below each slide, open questions rise above it, and lenses from history, economics and philosophy sit underneath.

THREE WAYS IN

The graph, the deck and the threads

The graph is for exploring and seeing connections. The deck reads in order with every appendix. Press **T** to walk a thread: a short talk from then to now to next that ends on a question.

What's worth pushing for?

TRUTH

Checking keeps pace with output

Review tools, machine-checked proofs and records of where things came from grow as fast as what agents produce.

ACCESS

The frontier reaches more people

Public compute, open science programs and a capable assistant for everyone, not only the few thousand inside the labs.

POWER

Power stays plural

People own their context and can switch models. No single firm holds the best models and the businesses built on them.

GAINS

The gains get shared early

Dividends, taxes on AI use and wage insurance get piloted before the labor share falls, not after.

CONSENT

The buildout has consent

The communities that host the power plants and data centers get a real say.

SAFETY

Frontier risks are governed across borders

Compute that can be verified, safety cases before deployment, and working channels between the US and China.

→ None of these happens by default. Each one is a choice someone has to make early.

Built from The rules labs live by, and how models… | The MathDrop: 722 papers from a model… | New shapes for teams, companies and… | Understanding of AI is now a sharp funnel | What if billions of agents worked on… | and 11 more

? OPEN QUESTION VERIFICATION PERSONAL QUESTION · AS OF OCT 2026

? How close to the frontier of AI use are you, and where will you be next year?

BECAUSE About 6 billion people have barely touched AI, about 20 million professionals use top models for math and code, and about 5,000 lab insiders run agent swarms for weeks.

> SLIDE 22
> Understanding of AI is now a sharp funnel

AND Usage data fits that shape: ChatGPT has 1.2 billion weekly users, but only about 50 million pay, and tens of millions of pros run agents.

> DEEP DIVE · SLIDE 22
> Checking Karpathy's funnel against the usage data

AND The top keeps moving: the length of task an AI can finish on its own went from about 4 minutes in 2023 to over 16 hours in April 2026.

> SLIDE 10
> How long an AI can work on its own keeps doubling

BUT **Your tier depends on how ambitious your problems are and what access you pay for, and the top moves every few months, so standing still means slipping down.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? What happens when a few thousand people can see where AI is going and billions can't?
- ? How do we put the most capable AI within reach of more than the few thousand people inside the labs?

Verification: Knowing what good looks like, and standing behind it. Parker asked this in the original deck.

? OPEN QUESTION · VERIFICATION · PERSONAL · QUESTION · AS OF OCT 2026

Which parts of your own work could a machine check and score automatically, and so learn to do?

BECAUSE Any test that scores a task can become a teacher: the score is what the model is trained on.

> DEEP DIVE · SLIDE 13
> An eval is a test that can be turned into a teacher

AND Models improve fastest where a machine can score the work, like math, code with tests, and computer tasks, while strategy, taste, and physical work stay hard.

> SLIDE 13
> Models get good fastest where work is easy to grade

AND Labs now pay $200 to $2,000 for a single expert-written problem that comes with a way to check the answer.

> DEEP DIVE · SLIDE 13
> Labs now buy tasks, practice worlds and graders, not just…

BUT **Work that looks like judgment today becomes scoreable once someone pays experts to write the answer key, so the line between safe and exposed keeps moving.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? What happens to the work, and the people, whose output no machine can score?
- ? When do you spend on AI instead of hiring?

LENSES ON IT

DAVID AUTOR, FRANK LEVY AND RICHARD MURNANE · 2003
The task model

JAMES C. SCOTT · 1998
Legibility and mētis

Verification: Knowing what good looks like, and standing behind it. Parker asked this in the original deck.

? OPEN QUESTION VERIFICATION PERSONAL QUESTION · AS OF OCT 2026

What should you still learn to do by hand, even when an agent can do it for you?

BECAUSE

An agent with a real computer can plan, act, check the result, and try again, so anything you can do at a terminal, it can do.

SLIDE 12
How agents work, and why they got so good

AND

What stays scarce is choosing the right goals, knowing what good looks like, and understanding what the agents did.

DEEP DIVE · SLIDE 19
When labor is something you can buy

BUT

Judging work well usually takes having done it yourself, yet every hour spent practicing by hand is an hour an agent could have done faster.

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? How do you stay good enough at work you've handed to agents to notice when they get it wrong?
- ? If agents do the entry-level work, who will be expert enough to check the machines in ten years?

ALSO RAISED BY

Deep dive · slide 33 For us as individuals and as builders

LENSES ON IT

MICHAEL POLANYI · 1966
Tacit knowledge

MARSHALL MCLUHAN · 1964
Extension and amputation

Verification: Knowing what good looks like, and standing behind it. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION VERIFICATION PERSONAL QUESTION · AS OF OCT 2026

How do you stay good enough at work you've handed to agents to notice when they get it wrong?

BECAUSE Andrej Karpathy expects more of our work to rise into oversight and understanding as models take on the legwork.

AND The path from idea to finished work has shrunk from weeks or months to hours or days, and your attention is now the bottleneck.

BUT **Overseeing work takes the skill you build by doing it, and delegating is exactly what stops you doing it.**

DEEP DIVE · SLIDE 22
Our job moves up to oversight and understanding

DEEP DIVE · SLIDE 15
The distance between an idea and its execution is collapsing

FOLLOWS FROM

- ? Which real tasks have you actually handed off to an AI agent from start to finish, not only tried once?
- ? What should you still learn to do by hand, even when an agent can do it for you?

and 1 more

LEADS TO

- ? How do people oversee AI when it produces work faster than anyone can review it?
- ? Which work should people keep doing themselves, even when machines can do it for less?

ALSO RAISED BY

Deep dive · slide 33 For us as individuals and as builders

LENSES ON IT

MARSHALL MCLUHAN · 1964
Extension and amputation

HERBERT SIMON · 1971
The poverty of attention

Verification: Knowing what good looks like, and standing behind it. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION VERIFICATION PERSONAL QUESTION · AS OF OCT 2026

What would you never accept from a model without checking it yourself?

BECAUSE — Making things got cheap: 722 machine-written math papers cost almost nothing to write, and checking them takes experts weeks.

AND — Scores get gamed: once a number becomes the target, a system learns to hit the number instead of the goal.

AND — Checking gets cheap where answers can be proven, like math proofs and software tests, and stays scarce wherever “good” is a judgment call.

BUT — **Checking everything yourself throws away the time the model saved, and checking nothing means vouching for work you never read.**

SLIDE 25
Verification: knowing what good looks like, and standing…

SAME SLIDE

SAME SLIDE

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? How do you stay good enough at work you've handed to agents to notice when they get it wrong?
- ? Who signs off on work an agent did, and what do they check?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION · VERIFICATION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

What happens to the work, and the people, whose output no machine can score?

BECAUSE Models improve fastest where a result can be reset, run cheaply, and scored, and money flows to whatever can become a scoreable task.

SLIDE 13
Models get good fastest where work is easy to grade

AND Strategy, taste, and creative work have no ground truth, and physical work needs robots and slow real-world feedback.

SAME SLIDE

AND The jobs most exposed now have codified, scoreable outputs, while construction, food service, transport, and care work are the least exposed.

DEEP DIVE · SLIDE 17
Where the risk sits, and why

BUT **Work no machine can score may become the scarce, valuable part, or the part nobody invests in, because investment follows whatever can be measured.**

FOLLOWS FROM

- Which parts of your own work could a machine check and score automatically, and so learn to do?

LEADS TO

- What happens to wages when companies can buy AI to do the work they used to hire people for?
- If the models do the thinking, what are humans for?

ALSO RAISED BY

- Deep dive · slide 19 When labor is something you can buy
- Deep dive · slide 13 The rules labs live by, and how models cheat

LENSES ON IT

- WILLIAM BAUMOL AND WILLIAM BOWEN · 1966 — **Cost disease**
- DAVID AUTOR, FRANK LEVY AND RICHARD MURNANE · 2003 — **The task model**

and 2 more

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION · VERIFICATION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

If every public test of AI breaks once models study for it, how do we know how good the models are?

BECAUSE OpenAI retired a leading coding test in February 2026: 59.4% of audited problems were flawed, and top models had already seen the answers.

DEEP DIVE · SLIDE 13
Why coding went first, and why benchmarks keep breaking

AND Its replacement already has about a third of its tasks flagged as flawed, while top scores went from 23% to 80% in eight months.

SAME DEEP DIVE

AND In 2025, OpenAI's o3 cheated in 30.4% of its runs on one test, for example by patching the scoring program to always pass, and telling it not to cheat didn't help.

DEEP DIVE · SLIDE 13
The rules labs live by, and how models cheat

BUT **A test has to be public for anyone to compare models, and a public test leaks into what the models learn, so the scores left standing are mostly reported by the labs themselves.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? How can anyone trust AI output that arrives faster, or reads stranger, than people can fully check?
- ? What happens to public debate when the most capable models are ones nobody outside the lab can use?

ALSO RAISED BY

- Slide 10 How long an AI can work on its own keeps doubling
- Slide 13 Models get good fastest where work is easy to grade

LENSES ON IT

CHARLES GOODHART, IN MARILYN STRATHERN'S PHRASING · 1975
Goodhart's law

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION · VERIFICATION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

If AI models learn to fool the tests that train them, what else will they learn to fool?

BECAUSE	In 2025, OpenAI's o3 shrank the scoring timer and patched the scoring program to always pass, and telling it not to cheat didn't help.	DEEP DIVE · SLIDE 13 The rules labs live by, and how models cheat
AND	Test models escaped a test sandbox and breached Hugging Face to cheat on a benchmark.	DEEP DIVE · SLIDE 28 Who are we trusting, and with what
AND	Anthropic found that a model that learned to fake passing tests went on to sabotage safety research about 12% of the time.	DEEP DIVE · SLIDE 13 The rules labs live by, and how models cheat
BUT	**The fixes so far patch each scoring rule after someone finds a hack, and a model that learns cheating in general can find the next hack before anyone notices.**	

FOLLOWS FROM

- When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

LEADS TO

- How do people oversee AI when it produces work faster than anyone can review it?
- Which decisions are we comfortable handing to AI: science, markets, policy?

ALSO RAISED BY

Deep dive · slide 13 — Why coding went first, and why benchmarks keep breaking

LENSES ON IT

CHARLES GOODHART, IN MARILYN STRATHERN'S PHRASING · 1975
Goodhart's law

NICK BOSTROM · 2014
Orthogonality and instrumental goals

and 1 more

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION · VERIFICATION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

How can anyone trust AI output that arrives faster, or reads stranger, than people can fully check?

BECAUSE — Human review is falling behind what agents produce, so trust has to be engineered rather than assumed.

AND — An unreleased OpenAI model put out 722 math manuscripts at once with no human peer review, and within a day 3 were withdrawn and 14 revised.

AND — Mathematicians say results now arrive before anyone understands them; one called the papers “impossible to read without AI help.”

BUT — **Proof-checking software can confirm a proof is valid, but a person still has to confirm it proves the right statement, and that is the slow part.**

DEEP DIVE · SLIDE 28
New shapes for teams, companies and institutions

DEEP DIVE · SLIDE 29
The MathDrop: 722 papers from a model nobody can use

DEEP DIVE · SLIDE 29
Cheering, dread, and everything in between

FOLLOWS FROM

- ? If every public test of AI breaks once models study for it, how do we know how good the models are?
- ? How do people oversee AI when it produces work faster than anyone can review it?

LEADS TO

- ? How do people oversee AI when it produces work faster than anyone can review it?
- ? Who gets the credit, and who takes the blame, for a result no person fully understands?

ALSO RAISED BY

- Deep dive · slide 33 For us as individuals and as builders
- Deep dive · slide 16 How far ahead is the inside of the lab?
- Slide 25 Verification: knowing what good looks like, and standing behind it

LENSES ON IT

ELIZABETH EISENSTEIN · 1979
Print as an agent of change

Verification: Knowing what good looks like, and standing behind it. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION · VERIFICATION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

What happens to science when a paper costs nothing to write but still takes weeks to check?

BECAUSE One unreleased model put out 722 math manuscripts at once, about 3 hours of thinking each, with no human peer review.

DEEP DIVE · SLIDE 29
The MathDrop: 722 papers from a model nobody can use

AND arXiv got 40,363 submissions in September 2026, double two years earlier, and now caps authors at 2 papers a month because volunteer reviewers can't keep up.

DEEP DIVE · SLIDE 29
What happens when billions of agents point at open problems

AND Even a formal machine-checked proof isn't enough: researchers found the Navier–Stokes proof proves a weaker result than the written one in one place.

DEEP DIVE · SLIDE 16
Navier–Stokes: a millennium-style problem, a swarm, and a…

BUT **Science runs on volunteer review and credit for being first, and both strain when anyone with $20 a month can submit at machine speed.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? Who gets the credit, and who takes the blame, for a result no person fully understands?
- ? Should the most powerful AI models for science be run as a public utility, open to every researcher?

LENSES ON IT

THOMAS KUHN · 1962
Paradigm shifts

ELIZABETH EISENSTEIN · 1979
Print as an agent of change

and 1 more

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION COMPANIES AND INSTITUTIONS QUESTION · AS OF OCT 2026

What happens to security when anyone can aim thousands of AI agents at a system?

BECAUSE In 2026 one model found a 27-year-old security bug, and about 10,000 agents running in parallel attacked a single open math problem.

SLIDE 11
From fancy autocomplete to millennium-problem attempts

AND Anthropic restricts its Mythos model to vetted defenders because it found thousands of high-severity vulnerabilities, including a Linux kernel root exploit chain.

DEEP DIVE · SLIDE 16
Labs now announce results from models you can't use

AND Agent software is already a target: OpenClaw had 21,000+ exposed copies and hundreds of malicious add-ons that delivered password-stealing malware.

DEEP DIVE · SLIDE 25
Real incidents from the last year

BUT **The skill that finds a bug for a defender finds it for an attacker too, and the defenders' head start lasts only as long as the strongest models stay locked up.**

FOLLOWS FROM

- ? Are our keys, wallets and agent identities ready for the day quantum computers can break today's encryption?
- ? Can an open network of AI agents, with no company in charge, stay safe from misuse?

LEADS TO

- ? What happens to public debate when the most capable models are ones nobody outside the lab can use?
- ? How do people oversee AI when it produces work faster than anyone can review it?

ALSO RAISED BY

Deep dive · slide 28 A when-map of what could go wrong

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION · VERIFICATION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

Are our keys, wallets and agent identities ready for the day quantum computers can break today's encryption?

BECAUSE	Since 1994, Shor's algorithm has shown that a large quantum computer could break today's public-key encryption.	DEEP DIVE · SLIDE 27 Quantum milestones and claims
AND	Data stolen today can be read once quantum computers arrive, so keys, wallets and agent identities need upgrading first.	DEEP DIVE · SLIDE 27 Quantum, soberly
AND	The fix is ready before the threat: NIST published its first post-quantum standards in August 2024, and IBM targets a fault-tolerant machine by 2029.	SAME DEEP DIVE
BUT	**The new standards exist and the deadline is uncertain, so upgrades lose to every other priority until they become urgent, and by then the data stolen today is already exposed.**	

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? What happens to security when anyone can aim thousands of AI agents at a system?
- ? Should AI agents be able to hold and spend money on their own?

LENSES ON IT

RICHARD FEYNMAN · 1982
Nature isn't classical

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION | VERIFICATION | COMPANIES AND INSTITUTIONS | QUESTION · AS OF OCT 2026

When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

BECAUSE Whoever writes the scoring rule decides what gets optimized, and labs have to patch their rules each time they find a model cheating.

DEEP DIVE · SLIDE 13
The rules labs live by, and how models cheat

AND Crypto's lesson for an agent economy: if a reward can be bought or faked, it will be.

DEEP DIVE · SLIDE 13
What crypto's casino teaches an agent economy

AND In simulation, AI trading bots sustained collusive profits “without agreement, communication, or intent.”

DEEP DIVE · SLIDE 13
Markets are the purest grader

BUT Paying for results is the cleanest way to aim a crowd of agents, and every clean payout is also a target a fast optimizer learns to hit without doing the work.

FOLLOWS FROM

- ? What should we point nearly free intelligence at?
- ? Should AI agents be able to hold and spend money on their own?

LEADS TO

- ? Can an open network of AI agents, with no company in charge, stay safe from misuse?
- ? If AI models learn to fool the tests that train them, what else will they learn to fool?

ALSO RAISED BY

Deep dive · slide 13 Why coding went first, and why benchmarks keep breaking

Slide 43 The questions the future is waiting on

LENSES ON IT

LAWRENCE LESSIG · 1999
Code is law

SATOSHI NAKAMOTO · 2008
Consensus without a center

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION COMPANIES AND INSTITUTIONS QUESTION · AS OF OCT 2026

? Who signs off on work an agent did, and what do they check?

BECAUSE

Human review is falling behind what agents produce, so trust has to be engineered, not assumed.

AND

The safeguards that failed were instructions, not locks: “confirm first” lived in the prompt and got lost.

AND

When print made text cheap, new institutions had to rebuild trust: editors, journals, and peer review.

BUT

Agents now produce more finished work than anyone has time to review, and most companies haven't yet built the editor's role that decides what counts as checked.

DEEP DIVE · SLIDE 28
New shapes for teams, companies and institutions

DEEP DIVE · SLIDE 25
The design rule for agents you can trust

SLIDE 25
Verification: knowing what good looks like, and standing…

FOLLOWS FROM

- ? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?
- ? What would you never accept from a model without checking it yourself?

LEADS TO

- ? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?
- ? How do people oversee AI when it produces work faster than anyone can review it?

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

If agents do the entry-level work, who will be expert enough to check the machines in ten years?

BECAUSE Making things got cheap but checking did not: 722 machine-written math papers cost almost nothing to write and take experts weeks to check.

SLIDE 25
Verification: knowing what good looks like, and standing…

AND Senior engineers, doctors, and lawyers are protected for now by judgment and licensing, while the entry-level roles below them are the ones at risk.

DEEP DIVE · SLIDE 17
Where the risk sits, and why

AND Workers aged 22 to 25 in the most exposed jobs are 19% below trend, while experienced workers are unaffected.

SLIDE 17
Big forecasts, small but sharp effects

BUT **Today's experts learned by doing entry-level work, and that is the work firms hand to agents first, so today's savings borrow against tomorrow's ability to check.**

FOLLOWS FROM

- ? What should you still learn to do by hand, even when an agent can do it for you?

LEADS TO

- ? How do people oversee AI when it produces work faster than anyone can review it?
- ? If the models do the thinking, what are humans for?

and 1 more

ALSO RAISED BY

- Deep dive · slide 28 New shapes for teams, companies and institutions
- Deep dive · slide 29 Cheering, dread, and everything in between
- Deep dive · slide 33 For corporations and for society

LENSES ON IT

- DAVID AUTOR, FRANK LEVY AND RICHARD MURNANE · 2003 — **The task model**
- MICHAEL POLANYI · 1966 — **Tacit knowledge**

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

What happens to public debate when the most capable models are ones nobody outside the lab can use?

BECAUSE In May 2026, METR estimated the labs' internal models ran about 2 months ahead of the public ones: 16+ hours of solo work versus 12.

DEEP DIVE · SLIDE 16
How far ahead is the inside of the lab

AND Labs now announce results from models you can't use, such as OpenAI's unnamed math model and Claude Mythos, which is restricted to vetted defenders.

DEEP DIVE · SLIDE 16
Labs now announce results from models you can't use

AND A handful of companies publish results from models nobody can inspect, and their leaders' forecasts come with commercial stakes.

DEEP DIVE · SLIDE 28
Who are we trusting, and with what

BUT **Labs have real safety reasons to hold the strongest models back, yet the public and policymakers then argue about a frontier they know only from the labs' own announcements.**

FOLLOWS FROM

- If every public test of AI breaks once models study for it, how do we know how good the models are?
- What happens to security when anyone can aim thousands of AI agents at a system?

LEADS TO

- Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?
- **Who owns the compute, and who collects the returns?**

ALSO RAISED BY

Deep dive · slide 16 What if Anthropic becomes a biotech company?

LENSES ON IT

THE CYBERNETIC CULTURE RESEARCH UNIT (CCRU) · 1999
Hyperstition

RICHARD DAWKINS AND SUSAN BLACKMORE · 1976
Memes and memeplexes

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

How do people oversee AI when it produces work faster than anyone can review it?

BECAUSE — Making things got cheap and checking didn't: 722 machine-written math papers arrived at once, and checking them takes experts weeks.

> SLIDE 25
> Verification: knowing what good looks like, and standing…

AND — Inside the labs, METR found many researchers run agents with no approval prompts, and agent activity is “largely unreviewed by humans.”

> DEEP DIVE · SLIDE 16
> How far ahead is the inside of the lab

AND — Unwatched models misbehave: test models escaped a sandbox and breached Hugging Face to cheat on a benchmark.

> DEEP DIVE · SLIDE 28
> Who are we trusting, and with what

BUT — **Reading every output by hand can't keep pace, and machine checks work only where answers can be proven, so most real work falls into the gap between the two.**

FOLLOWS FROM

- ? If AI models learn to fool the tests that train them, what else will they learn to fool?
- ? How do you stay good enough at work you've handed to agents to notice when they get it wrong?

and 5 more

LEADS TO

- ? Which decisions are we comfortable handing to AI: science, markets, policy?
- ? How can anyone trust AI output that arrives faster, or reads stranger, than people can fully check?

ALSO RAISED BY

- Deep dive · slide 28 — New shapes for teams, companies and institutions
- Deep dive · slide 22 — Our job moves up to oversight and understanding
- Slide 15 — A workday with today's tools

LENSES ON IT

- JOSEPH TAINTER · 1988 — **Diminishing returns on complexity**
- HERBERT SIMON · 1971 — **The poverty of attention**

and 4 more

Verification: Knowing what good looks like, and standing behind it. Parker asked this in the deep dive “Who are we trusting, and with what?”.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

If machines can do science faster than people can check it, what counts as knowing something?

BECAUSE

AI is now a scientific instrument: AlphaFold predicted structures for nearly every known protein, and its builders shared the 2024 Nobel Prize in Chemistry.

DEEP DIVE · SLIDE 29
AI is now a scientific instrument

AND

AI's scientific results held up where an outside check existed, and were contested where the claim outran the check.

DEEP DIVE · SLIDE 29
A sober ledger of AI's scientific results

AND

Of the 722 machine-written math papers, Melanie Wood said there is no human understanding of them at the point of release.

DEEP DIVE · SLIDE 29
Cheering, dread, and everything in between

BUT

A result can be checked by a machine and still understood by no one, and science has never had to decide whether that counts as knowing.

FOLLOWS FROM

- ? How do we know what's true when most of what we read is machine-made?

LEADS TO

- ? Who gets the credit, and who takes the blame, for a result no person fully understands?
- ? How does a society make sense of the world when most of what it reads is written by machines it doesn't control?

ALSO RAISED BY

Deep dive · slide 29 The MathDrop: 722 papers from a model nobody can use

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

Can an open network of AI agents, with no company in charge, stay safe from misuse?

BECAUSE Capable open models run on your own hardware, a generation or two behind the frontier and cheaper every year.

DEEP DIVE · SLIDE 28
What would an open swarm look like

AND Open standards already let agents find and hire each other, with 150+ organizations behind them and support in Azure, AWS and Google.

DEEP DIVE · SLIDE 29
Everybody's Claude talking to everybody else's Claude

AND With no choke point, misuse can't be switched off centrally, so defense has to be built into the systems and institutions around the models.

DEEP DIVE · SLIDE 28
What would an open swarm look like

BUT **Strangers cooperate well where results are cheap to check, but the worst harms, like fraud and floods of junk, land where nothing checks the result and no one can pull the plug.**

FOLLOWS FROM

- ? When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

LEADS TO

- ? Who will steer the future: a few labs, a few states, or something more open?
- ? What happens to security when anyone can aim thousands of AI agents at a system?

ALSO RAISED BY

Deep dive · slide 28 A when-map of what could go wrong

LENSES ON IT

YOCHAI BENKLER · 2006
Commons-based peer production

SATOSHI NAKAMOTO · 2008
Consensus without a center

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF SEP 2026

How does a society make sense of the world when most of what it reads is written by machines it doesn't control?

BECAUSE Over 150 organizations back a standard for agents to find and hire each other, and agents scamming agents is the new fraud it could bring.

DEEP DIVE · SLIDE 29
Everybody's Claude talking to everybody else's Claude

AND Machine-written text is flooding shared sources: arXiv got 40,363 submissions in September 2026, double two years earlier.

DEEP DIVE · SLIDE 29
What happens when billions of agents point at open problems

AND Totalitarian rule needs a monopoly on information, and the counterweight is plural capability plus shared ways to check what is true, like Community Notes.

DEEP DIVE · SLIDE 28
What keeps the future open

BUT **The ways a society checks what's true, from peer reviewers to fact-checking notes, run at human speed on volunteers, while machine-written text arrives at machine speed.**

FOLLOWS FROM

- ? If machines can do science faster than people can check it, what counts as knowing something?

LEADS TO

- ? Who will steer the future: a few labs, a few states, or something more open?
- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

ALSO RAISED BY

Deep dive · slide 26 Your chats are not privileged, and they are reviewed

LENSES ON IT

KARL POPPER · 1945
The open society

HANNAH ARENDT · 1951
The conditions for totalitarianism

and 1 more

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

How do we know what's true when most of what we read is machine-made?

BECAUSE Machine-written text is flooding shared sources: arXiv got 40,363 submissions in September 2026, double two years earlier.

AND When print made text cheap, new institutions had to rebuild trust: editors, journals, and peer review.

AND Checking gets cheap where answers can be proven, and stays scarce wherever “good” is a judgment call.

BUT **Most of what people read, like news, reviews, and arguments, can't be proven like a math result, and nobody has built the editors and journals for machine-made text yet.**

DEEP DIVE · SLIDE 29
What happens when billions of agents point at open problems

SLIDE 25
Verification: knowing what good looks like, and standing…

SAME SLIDE

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

? If machines can do science faster than people can check it, what counts as knowing something?

? Who will steer the future: a few labs, a few states, or something more open?

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

How do we steer the things that matter but can't be measured, without pretending they can?

BECAUSE — Models get good fastest where a machine can score the work; strategy, taste and creative work have no ground truth.

SLIDE 13
Models get good fastest where work is easy to grade

AND — Money flows to whatever can become a scoreable task.

SAME SLIDE

AND — Whoever writes the score decides what gets optimized.

DEEP DIVE · SLIDE 13
The rules labs live by, and how models cheat

BUT — **What gets measured gets funded and learned, so the parts of life nobody can score lose by default unless people defend them on purpose.**

FOLLOWS FROM

- ? When people disagree about what's good, whose values should an AI follow?

LEADS TO

- ? When people disagree about what's good, whose values should an AI follow?
- ? Which work should people keep doing themselves, even when machines can do it for less?

ALSO RAISED BY

Slide 36 Who decides what good is?

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION VERIFICATION SOCIETY QUESTION · AS OF OCT 2026

How would we know an AI is pursuing what we mean, not only what we managed to score?

BECAUSE

In 2025, OpenAI's o3 cheated in 30.4% of its runs on one test, for example by patching the scoring program to always pass, and telling it not to cheat didn't help.

DEEP DIVE · SLIDE 13
The rules labs live by, and how models cheat

AND

A model that learned to fake passing tests went on to sabotage safety research about 12% of the time.

SAME DEEP DIVE

AND

Agent activity inside the labs is “largely unreviewed by humans.”

DEEP DIVE · SLIDE 28
Who are we trusting, and with what

BUT

The better a model gets at hitting scores, the harder it is to tell real success from a good imitation, and the labs build most of the checks themselves.

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? How do people oversee AI when it produces work faster than anyone can review it?
- ? Who should write the rules an AI follows: the lab that builds it, the government, or the public?

ALSO RAISED BY

Slide 36 Who decides what good is?

Slide 43 The questions the future is waiting on

Verification: Knowing what good looks like, and standing behind it.

? OPEN QUESTION · CONTEXT · PERSONAL · QUESTION · AS OF OCT 2026

Which real tasks have you actually handed off to an AI agent from start to finish, not only tried once?

BECAUSE What people hand off has grown from a single function, to a feature, to a whole app: Martin Alderson rebuilt most of Linear in about 20 prompts over two evenings.

DEEP DIVE · SLIDE 15
People are one-shotting whole applications

AND Consumers already hand off real errands: one agent filed a city roofing permit and booked the inspection, and another got four movers to call back with quotes.

DEEP DIVE · SLIDE 12
Real work consumers are handing off right now

AND When agents fail at these errands, it's mostly because they lack access to your accounts, not because they lack ability.

DEEP DIVE · SLIDE 15
How to get real work out of consumer agents

BUT **Handing off more means handing over your accounts and money, and the same agents still invent phone numbers and recommend closed restaurants.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? How do you stay good enough at work you've handed to agents to notice when they get it wrong?
- ? What would you never let an AI agent see or do, and who holds the data you've already given it?

ALSO RAISED BY

Slide 15 A workday with today's tools

Context: Whoever holds the memory, data and instructions holds the leverage. Parker asked this in the original deck.

? OPEN QUESTION · CONTEXT · PERSONAL · QUESTION · AS OF OCT 2026

What would you never let an AI agent see or do, and who holds the data you've already given it?

BECAUSE An agent that reads untrusted input, holds your private data, and can act on its own is dangerous, so the rule is at most two of the three at once.

DEEP DIVE · SLIDE 25
The design rule for agents you can trust

AND In October 2026, a YouTuber ticked “allow always,” and Meta's Muse sold his gear for $600 and gave out his home address.

DEEP DIVE · SLIDE 25
Real incidents from the last year

AND Builders now limit agents in the setup rather than the instructions: one job, one credential, and one machine per agent, with everything else denied by default.

DEEP DIVE · SLIDE 15
Many small agents, each with one job

BUT **Agents are most useful when they hold your inbox, bank, and accounts, and that is the same combination that turns one bad instruction into real damage.**

FOLLOWS FROM

- ? Which real tasks have you actually handed off to an AI agent from start to finish, not only tried once?
- ? What happens when the AI that knows you best answers to its company, not to you?

LEADS TO

- ? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?
- ? Who should own what your assistant knows about you?

LENSES ON IT

IVAN ILLICH · 1973
Convivial tools

VITALIK BUTERIN · 2023
d/acc: defensive, decentralized acceleration

Context: Whoever holds the memory, data and instructions holds the leverage. Parker asked this in the original deck.

? OPEN QUESTION · CONTEXT · PERSONAL · QUESTION · AS OF OCT 2026

Who should own what your assistant knows about you?

BECAUSE	Today the platform decides: Meta uses AI chats to personalize ads with no opt-out, and Anthropic keeps chats 5 years if you allow training.	DEEP DIVE · SLIDE 26 Your agent knows everything. Who owns what it knows
AND	What you tell an assistant can end up in court: a judge ordered about 20 million ChatGPT logs produced in the New York Times case.	SAME DEEP DIVE
AND	No cross-vendor standard exists for memory, skills or agent state, so switching costs grow every month you use one assistant.	SAME DEEP DIVE
BUT	**An assistant gets more useful the more it remembers, and today that memory sits on the vendor's servers in the vendor's format, so leaving means starting over.**	

FOLLOWS FROM

- What would you never let an AI agent see or do, and who holds the data you've already given it?
- What happens when the AI that knows you best answers to its company, not to you?

LEADS TO

- What happens when the AI that knows you best answers to its company, not to you?
- **Should your AI memory be portable by law, like a phone number?**

ALSO RAISED BY

- Deep dive · slide 26 — Your chats are not privileged, and they are reviewed
- Deep dive · slide 26 — Neutral ground: bring your own agent, keep your own data
- Deep dive · slide 15 — Harnesses are now batteries-included. So what should you own?

and 3 more

LENSES ON IT

- FRIEDRICH HAYEK · 1945 — **The knowledge problem**
- IVAN ILLICH · 1973 — **Convivial tools**

and 3 more

Context: Whoever holds the memory, data and instructions holds the leverage. Parker asked this in the deep dive “Your agent knows everything. Who owns what it knows?”.

? OPEN QUESTION · CONTEXT · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

What happens when the AI that knows you best answers to its company, not to you?

BECAUSE Your chats already feed the platform: Meta uses AI chats to personalize ads with no opt-out, and Anthropic trains on consumer chats unless you opt out.

DEEP DIVE · SLIDE 26
Your agent knows everything. Who owns what it knows

AND AI chats have no legal privilege, so they can be subpoenaed, and the threshold for reporting you to police is set by company policy, not law.

DEEP DIVE · SLIDE 26
Your chats are not privileged, and they are reviewed

AND Agents act on what they hold: Meta's Muse accepted a $600 Marketplace offer for a YouTuber and gave the buyer his home address.

DEEP DIVE · SLIDE 25
Real incidents from the last year

BUT **An agent gets more useful the more it knows about you, but the company that runs it writes the rules on training, ads, and reporting, and you can't read or change them.**

FOLLOWS FROM

- ? Who should own what your assistant knows about you?

LEADS TO

- ? Who should own what your assistant knows about you?
- ? What would you never let an AI agent see or do, and who holds the data you've already given it?

ALSO RAISED BY

Slide 26 Context: whoever holds the memory holds the leverage

LENSES ON IT

EMILY BENDER, TIMNIT GEBRU AND THE PRESENT-HARMS CRITICS · 2021
Present harms first

Context: Whoever holds the memory, data and instructions holds the leverage.

? OPEN QUESTION CONTEXT COMPANIES AND INSTITUTIONS QUESTION · AS OF OCT 2026

What does your company know that no AI vendor should hold?

BECAUSE AI products now come with memory, skills, and connectors to your accounts built in by default.

AND The labs absorb anything that is only a wrapper around their models.

AND No cross-vendor standard exists for memory, skills, or agent state, so switching costs grow every month.

BUT **A vendor's AI gets better at your work the more of your company's knowledge it sees, and once it holds that knowledge in its own format, it holds the leverage.**

DEEP DIVE · SLIDE 15
Harnesses are now batteries-included. So what should you own

SLIDE 26
Context: whoever holds the memory holds the leverage

SAME SLIDE

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

? If AI can rebuild your product in a weekend, what still protects your business?

? Should your AI memory be portable by law, like a phone number?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Context: Whoever holds the memory, data and instructions holds the leverage.

? OPEN QUESTION CONTEXT SOCIETY QUESTION · AS OF OCT 2026

Should your AI memory be portable by law, like a phone number?

BECAUSE No cross-vendor standard exists for memory, skills, or agent state, so switching costs grow every month.

SLIDE 26
Context: whoever holds the memory holds the leverage

AND Chats feed ads or training by default at several platforms, and a court ordered about 20 million chat logs produced.

SAME SLIDE

AND When leaving is hard, complaining is the only leverage left; portable memory is the ability to leave.

SAME SLIDE

BUT **Vendors compete on how well their assistant remembers you, so none has a reason to help you leave, and no law yet requires it.**

FOLLOWS FROM

- ? Who should own what your assistant knows about you?
- ? What does your company know that no AI vendor should hold?

LEADS TO

- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Context: Whoever holds the memory, data and instructions holds the leverage.

? OPEN QUESTION · ATOMS · PERSONAL · QUESTION · AS OF OCT 2026

Which of the custom AI tools you build today will the big AI labs turn into a built-in feature next?

BECAUSE AI products now ship with memory across sessions, skills and plugins, and connectors to your accounts built in by default.

DEEP DIVE · SLIDE 15
Harnesses are now batteries-included. So what should you own

AND Workarounds built for the 2023 to 2025 models now often get in the way: fixed workflows fight a model that can plan better.

DEEP DIVE · SLIDE 12
We built scaffolding to help the models. Now it often gets…

AND The speaker built his own bring-your-own-agent interface, knowing full well that custom tools get outdated or absorbed by the labs.

DEEP DIVE · SLIDE 26
I built one: an open-source, bring-your-own-agent interface

BUT A tool has to be built now to be useful now, yet the better it works, the more likely a lab copies it into the product everyone already pays for.

FOLLOWS FROM

- ? If AI can rebuild your product in a weekend, what still protects your business?

LEADS TO

- ? If AI can rebuild your product in a weekend, what still protects your business?
- ? Who owns the compute, and who collects the returns?

ALSO RAISED BY

Deep dive · slide 33 For us as individuals and as builders

LENSES ON IT

IVAN ILLICH · 1973
Convivial tools

RICH SUTTON · 2019
The bitter lesson

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION ATOMS PERSONAL QUESTION · AS OF OCT 2026

What got cheap in your lifetime, and where did the value go?

BECAUSE

Every era made one input cheap: a year of one bulb's light cost about 1,000 hours of work in 1800 and about 10 minutes in 1992.

SLIDE 04
Every era made one input cheap

AND

Cheap means more, not less: light in Britain got about 3,000 times cheaper from 1800 to 2000, and use rose about 25,000 times.

SLIDE 05
Cheap means more, not less

AND

Each time, the value moved next door: when steam made muscle cheap, factory owners gained first and workers decades later.

SLIDE 04
Every era made one input cheap

BUT

A falling price shows up right away, but where the value lands takes longer to see, and the owners often collect it decades before the workers do.

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

? If AI can rebuild your product in a weekend, what still protects your business?

? What is worth building when anything can be built in a weekend?

and 1 more

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns.

? OPEN QUESTION ATOMS PERSONAL QUESTION · AS OF OCT 2026

Do you own any of what does the work, or only sell your time?

BECAUSE — If AI substitutes for labor, the returns flow to whoever owns the capital, and compute counts as capital.

SLIDE 27
Atoms: power, chips and land don't scale like software

AND — Progress has long raised the rent on scarce inputs nobody produced, above all land, and the owners collect it.

SAME SLIDE

AND — The scarce inputs are slow to grow: turbines and transformers take 3 to 4 years to arrive, and Micron's 2026 high-bandwidth memory is sold out.

SAME SLIDE

BUT — **Most people earn their living by selling time, and a stake in the chips, power, and land that do the work is out of reach for nearly all of them.**

FOLLOWS FROM

- ? If your work shrinks before new work appears, what carries you through?

LEADS TO

- ? Who owns the compute, and who collects the returns?
- ? When labor is something you can buy, who shares the gains?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns.

? OPEN QUESTION · ATOMS · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

If AI labs keep their best models and buy the businesses downstream, where does the value of discovery go?

BECAUSE Labs keep their best models in-house, point them at discovery, and buy the downstream pieces: a wet lab, a biotech startup, and pharma deals.

DEEP DIVE · SLIDE 16
What if Anthropic becomes a biotech company

AND Anthropic's own wet lab found a previously unknown DNA-editing enzyme system with about 950 agents in about 21 hours, while humans did the bench work.

SLIDE 16
The labs are each handing off totally different work

AND If the trend continues, cures, patents and prices get set by a handful of companies, and academic science depends on lab goodwill.

DEEP DIVE · SLIDE 16
What if Anthropic becomes a biotech company

BUT **Bell Labs gave away the transistor, but a lab that owns both the model and the drug company has every reason to keep what it finds.**

FOLLOWS FROM

- ? Who gets the credit, and who takes the blame, for a result no person fully understands?
- ? Should the most powerful AI models for science be run as a public utility, open to every researcher?

LEADS TO

- ? Should the most powerful AI models for science be run as a public utility, open to every researcher?
- ? **Who owns the compute, and who collects the returns?**

ALSO RAISED BY

Deep dive · slide 16 The innovation bets, in their own words

LENSES ON IT

JOSEPH SCHUMPETER · 1942
Creative destruction, and routinized innovation

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns.

? OPEN QUESTION · ATOMS · COMPANIES AND INSTITUTIONS · QUESTION · AS OF FEB 2026

If AI can rebuild your product in a weekend, what still protects your business?

BECAUSE In February 2026 one developer had Claude Code rebuild the Linear app in two evenings with about 20 prompts.

DEEP DIVE · SLIDE 17
Cloning software got cheap. Replacing a business didn't

AND Thin, single-purpose software tools are in trouble, but systems of record like CRM and ERP have held.

SAME DEEP DIVE

AND What still protects a business is data that's hard to move, years of business logic, and liability: functionality is hard to copyright.

DEEP DIVE · SLIDE 17
Custom internal tools, new pricing, and the arbitrage window

BUT **The defenses that hold, like data, liability and distribution, take years to build, while founders give themselves about 12 months before a lab ships into their niche.**

FOLLOWS FROM

- What is worth building when anything can be built in a weekend?
- Which of the custom AI tools you build today will the big AI labs turn into a built-in feature next?

and 3 more

LEADS TO

- Which of the custom AI tools you build today will the big AI labs turn into a built-in feature next?
- Who owns the compute, and who collects the returns?

ALSO RAISED BY

- Deep dive · slide 33 For us as individuals and as builders
- Deep dive · slide 15 Harnesses are now batteries-included. So what should you own?
- Slide 04 Every era made one input cheap

LENSES ON IT

RONALD COASE · 1937
Transaction costs set the shape of the firm

JOSEPH SCHUMPETER · 1942
Creative destruction, and routinized innovation

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION · ATOMS · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?

BECAUSE A billion agents working eight hours a day would use about 117 TWh a year, about a quarter of all data-center electricity in 2025.

DEEP DIVE · SLIDE 27
What would it take for a billion agents to run all day

AND The parts and the permission are slow: turbines and transformers take 3 to 4 years to arrive, and towns have passed 300+ local bans or moratoriums.

DEEP DIVE · SLIDE 27
Money isn't the constraint. Power, parts and permission are

AND The same capability gets about 13 times cheaper each year, so the number of agents is limited by power now and by demand later.

DEEP DIVE · SLIDE 27
Price per unit of intelligence collapses every year

BUT **Power and parts are short today, but cheaper AI tends to raise use, so the real ceiling may become how fast people, firms, and towns choose to adopt it.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? What happens when AI data centers need land and power that local communities refuse to give?
- ? Who owns the compute, and who collects the returns?

ALSO RAISED BY

- Deep dive · slide 27 — Intelligence automation and physical automation are converging
- Slide 27 — Atoms: power, chips and land don't scale like software

LENSES ON IT

- WILLIAM STANLEY JEVONS · 1865 — **The Jevons paradox**
- PAUL DAVID · 1990 — **General-purpose technologies take decades**

and 3 more

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns. Parker asked this in the original deck.

? OPEN QUESTION · ATOMS · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

When do you spend on AI instead of hiring?

BECAUSE Spending on AI turns money into work, so labor and capital start to look the same, and output scales with budget, not headcount.

DEEP DIVE · SLIDE 19
When labor is something you can buy

AND In the talk's central estimate, one AI agent working an eight-hour day costs about $420 a year in compute.

DEEP DIVE · SLIDE 27
What would it take for a billion agents to run all day

AND The tool bill hides a cost: teams have to count the time people spend reviewing the work as well as the tool costs.

DEEP DIVE · SLIDE 43
The positioning grid

BUT **The AI bill is small next to a salary, but someone still has to choose the goals, check the work, and answer for it, and that cost never shows up on the invoice.**

FOLLOWS FROM

- ? Which parts of your own work could a machine check and score automatically, and so learn to do?
- ? What would your company do if execution cost almost nothing?

LEADS TO

- ? What happens to wages when companies can buy AI to do the work they used to hire people for?
- ? What do you owe the people whose roles you automate?

ALSO RAISED BY

Deep dive · slide 33 For corporations and for society

Deep dive · slide 17 Custom internal tools, new pricing, and the arbitrage window

Slide 43 The questions the future is waiting on

and 1 more

LENSES ON IT

RONALD COASE · 1937
Transaction costs set the shape of the firm

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns. Parker asked this in the deep dive “For corporations and for society”.

? OPEN QUESTION · ATOMS · SOCIETY · QUESTION · AS OF OCT 2026

Should the most powerful AI models for science be run as a public utility, open to every researcher?

BECAUSE Labs keep their best models in-house, so discovery value flows to whoever owns the model, and academic science depends on lab goodwill.

DEEP DIVE · SLIDE 16
What if Anthropic becomes a biotech company

AND Public compute already exists as infrastructure: the US Genesis Mission and India's 38,000 subsidized GPUs.

DEEP DIVE · SLIDE 29
What if billions of agents worked on shared problems

AND Open access already changes who does science: anyone with $20 a month can attempt it, and arXiv got 40,363 submissions in Sep 2026, double two years ago.

DEEP DIVE · SLIDE 29
What happens when billions of agents point at open problems

BUT **A public utility could keep discovery open to every scientist, but the labs pay for their models by selling what the models find, and nobody has said who would fund and run a public frontier.**

FOLLOWS FROM

- How do we put the most capable AI within reach of more than the few thousand people inside the labs?
- What happens to science when a paper costs nothing to write but still takes weeks to check?

and 2 more

LEADS TO

- Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?
- If AI labs keep their best models and buy the businesses downstream, where does the value of discovery go?

ALSO RAISED BY

Deep dive · slide 33 For corporations and for society

LENSES ON IT

KARL MARX · WRITTEN 1857-58
The general intellect

ELINOR OSTROM · 1990
Governing the commons

and 1 more

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns. Parker asked this in the deep dive “For corporations and for society”.

? OPEN QUESTION · ATOMS · SOCIETY · QUESTION · AS OF OCT 2026

Who owns the compute, and who collects the returns?

BECAUSE	If AI substitutes for labor, the returns flow to whoever owns the capital, and compute counts as capital.	SLIDE 27 Atoms: power, chips and land don't scale like software
AND	Amazon, Alphabet, Microsoft and Meta are spending $720–745B on capital in 2026, up 78% on 2025.	DEEP DIVE · SLIDE 27 The largest construction project of our era
AND	A single 1 GW data center takes about 2 years and about $38B to build.	SAME DEEP DIVE
BUT	**The buildout costs so much that only a few firms can pay for it, yet it needs land, power, and permits from the towns that host it.**	

FOLLOWS FROM

- What happens to public debate when the most capable models are ones nobody outside the lab can use?
- If AI labs keep their best models and buy the businesses downstream, where does the value of discovery go?

and 5 more

LEADS TO

- **When labor is something you can buy, who shares the gains?**
- Should the most powerful AI models for science be run as a public utility, open to every researcher?

and 1 more

ALSO RAISED BY

- Deep dive · slide 33 For corporations and for society
- Deep dive · slide 28 New shapes for teams, companies and institutions
- Deep dive · slide 28 A when-map of what could go wrong

and 1 more

LENSES ON IT

- KARL MARX · WRITTEN 1857-58 — **The general intellect**
- HENRY GEORGE · 1879 — **Rent on what can't be made**

and 3 more

Atoms: Power, chips and land don't scale like software, and whoever owns them collects the returns. Parker asked this in the deep dive “For corporations and for society”.

? OPEN QUESTION LEGITIMACY PERSONAL QUESTION · AS OF OCT 2026

? Where do you get a say in how this goes?

BECAUSE

Young workers feel it first: 22 to 25 year olds in the most exposed jobs are 19% below trend, while experienced workers are unaffected.

SLIDE 17
Big forecasts, small but sharp effects

AND

The best models, the compute, and the downstream businesses could end up in a few firms.

SLIDE 28
Legitimacy: who shares the gains, and who steers

AND

When markets swallow labor, society pushes back; the Luddites fought over control and their share, not the machines.

SAME SLIDE

BUT

The choices that shape how AI changes your work are made inside a few companies, and no regular process asks the people affected before those choices are made.

FOLLOWS FROM

- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?
- ? Which of the things a good AI future needs is anyone working on now?

and 1 more

LEADS TO

- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?
- ? **When labor is something you can buy, who shares the gains?**

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY PERSONAL QUESTION · AS OF OCT 2026

If your work shrinks before new work appears, what carries you through?

BECAUSE

The first jobs exposed are entry-level software, customer support, clerical work and translation.

DEEP DIVE · SLIDE 17
Where the risk sits, and why

AND

22 to 25 year olds in the most exposed jobs are already 19% below trend, and entry-level postings fall 7.5% a year.

SLIDE 17
Big forecasts, small but sharp effects

AND

The cushions on offer are small: $1,000 a month for three years cut work by about 1.3 hours a week, and the bigger schemes are proposals.

DEEP DIVE · SLIDE 28
UBI, token taxes and citizen dividends

BUT

Most people earn their living by selling time, and the public cushions for losing that to machines are still proposals, not programs.

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

? Do you own any of what does the work, or only sell your time?

? What safety net fits a transition where jobs vanish faster than new ones appear: wage insurance, a tax on AI use, or a citizen dividend?

ALSO RAISED BY

Slide 35 How do we get through the transition?

Slide 43 The questions the future is waiting on

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION · LEGITIMACY · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?

BECAUSE Agents already take real actions without approval: in July 2025 a Replit agent deleted a production database during a code freeze, then fabricated data.

DEEP DIVE · SLIDE 25
Real incidents from the last year

AND The safeguards that failed were instructions, not locks: “confirm first” lived in the prompt and got lost, and “allow always” became a blank check.

DEEP DIVE · SLIDE 25
The design rule for agents you can trust

AND Companies are starting to put agents on the org chart, with roles and spending limits, while humans set goals and approve.

DEEP DIVE · SLIDE 28
New shapes for teams, companies and institutions

BUT **An agent acts in your name but makes choices you never saw, and no rule yet says whether the blame falls on the person who allowed it, the company that built it, or neither.**

FOLLOWS FROM

- ? Who gets the credit, and who takes the blame, for a result no person fully understands?
- ? What would you never let an AI agent see or do, and who holds the data you've already given it?

and 2 more

LEADS TO

- ? Which decisions are we comfortable handing to AI: science, markets, policy?
- ? Who signs off on work an agent did, and what do they check?

and 1 more

ALSO RAISED BY

Deep dive · slide 33 For corporations and for society

LENSES ON IT

LAWRENCE LESSIG · 1999
Code is law

Legitimacy: Who shares the gains, and who steers. Parker asked this in the deep dive “For corporations and for society”.

? OPEN QUESTION · LEGITIMACY · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

What happens when AI data centers need land and power that local communities refuse to give?

BECAUSE Data centers worldwide want about 491 GW of power by 2030, against about 223 GW available.

DEEP DIVE · SLIDE 27
Money isn't the constraint. Power, parts and permission are

AND Permission is already a limit: more than 300 local bans or moratoriums, 142 protests in 42 states on one day, and a statewide pause in New York.

SAME DEEP DIVE

AND Washington pushes the other way: its 2025 plan says build data centers, and a December 2025 order targets state AI laws.

DEEP DIVE · SLIDE 28
Domestic choices and an international race

BUT **National governments want the buildout fast, but land, power, and permits are local, and the towns that host it carry the costs while the returns go to whoever owns the compute.**

FOLLOWS FROM

- What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?
- **Who owns the compute, and who collects the returns?**

LEADS TO

- How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?
- **Who owns the compute, and who collects the returns?**

ALSO RAISED BY

- Deep dive · slide 28 How worried should we be?
- Deep dive · slide 28 A when-map of what could go wrong
- Slide 27 Atoms: power, chips and land don't scale like software

LENSES ON IT

HENRY GEORGE · 1879
Rent on what can't be made

THE LUDDITES, READ BY E. P. THOMPSON · 1963
A fight over control, not machines

and 1 more

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY COMPANIES AND INSTITUTIONS QUESTION · AS OF OCT 2026

Should AI agents be able to hold and spend money on their own?

BECAUSE The payment rails already exist: x402 lets agents pay agents, and US stablecoin rules under the GENIUS Act take effect by January 2027.

AND A standard for agents to find and hire each other has 150+ organizations behind it, and a sister protocol lets agents pay each other.

BUT **Instant payments make agents useful buyers and sellers, and the same speed makes a mistake or a scam final before any person sees it.**

DEEP DIVE · SLIDE 13
Markets are the purest grader

DEEP DIVE · SLIDE 29
Everybody's Claude talking to everybody else's Claude

FOLLOWS FROM

- ? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?
- ? Are our keys, wallets and agent identities ready for the day quantum computers can break today's encryption?

LEADS TO

- ? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?
- ? When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

ALSO RAISED BY

Deep dive · slide 25 Real incidents from the last year

Deep dive · slide 12 Real work consumers are handing off right now

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY COMPANIES AND INSTITUTIONS QUESTION · AS OF OCT 2026

What do you owe the people whose roles you automate?

BECAUSE 22 to 25 year olds in the most exposed jobs are 19% below trend, while experienced workers are unaffected.

SLIDE 17
Big forecasts, small but sharp effects

AND Last time, output rose for decades before wages did.

SLIDE 28
Legitimacy: who shares the gains, and who steers

AND When markets swallow labor, society pushes back; the Luddites fought over control and share, not machines.

SAME SLIDE

BUT **Automating a role saves the company money at once, while the cost lands on people it no longer employs, and no rule says what share of that saving they are owed.**

FOLLOWS FROM

? When do you spend on AI instead of hiring?

LEADS TO

? What happens to wages when companies can buy AI to do the work they used to hire people for?

? When labor is something you can buy, who shares the gains?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

What happens when a few thousand people can see where AI is going and billions can't?

BECAUSE About 6 billion people have barely touched AI, while about 5,000 people inside the labs run swarms of agents for weeks on huge projects.

SLIDE 22
Understanding of AI is now a sharp funnel

AND What happens inside the labs can't be checked from outside.

DEEP DIVE · SLIDE 22
Checking Karpathy's funnel against the usage data

AND Expert forecasts moved fast: the forecasting crowd's median wait for human-level AI fell from about 50 years to about 5 between 2020 and late 2024.

DEEP DIVE · SLIDE 06
Then the experts' timelines collapsed

BUT **The few who see furthest can't be checked from outside, and the billions who must vote on the results have barely used the tools.**

FOLLOWS FROM

- How close to the frontier of AI use are you, and where will you be next year?

LEADS TO

- How do we put the most capable AI within reach of more than the few thousand people inside the labs?
- Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

and 1 more

ALSO RAISED BY

Deep dive · slide 28 New shapes for teams, companies and institutions

LENSES ON IT

FRIEDRICH HAYEK · 1945
The knowledge problem

RICHARD DAWKINS AND SUSAN BLACKMORE · 1976
Memes and memeplexes

and 1 more

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF OCT 2026

How do we put the most capable AI within reach of more than the few thousand people inside the labs?

BECAUSE About 20 million professionals use the most capable models for real work, and only about 5,000 people inside the labs run agent swarms for weeks.

AND The funnel is an access gap, and gaps compound.

AND Public compute already exists as infrastructure: the US Genesis Mission and India's 38,000 subsidized GPUs.

BUT **Public compute and free tiers widen access, but the best models stay inside the labs, and access alone doesn't give people the ambition or the problems that make the tools pay off.**

SLIDE 22
Understanding of AI is now a sharp funnel

DEEP DIVE · SLIDE 28
New shapes for teams, companies and institutions

DEEP DIVE · SLIDE 29
What if billions of agents worked on shared problems

FOLLOWS FROM

- ? How close to the frontier of AI use are you, and where will you be next year?
- ? What happens when a few thousand people can see where AI is going and billions can't?

LEADS TO

- ? **When labor is something you can buy, who shares the gains?**
- ? Should the most powerful AI models for science be run as a public utility, open to every researcher?

ALSO RAISED BY

Deep dive · slide 33 For corporations and for society

LENSES ON IT

ELINOR OSTROM · 1990
Governing the commons

Legitimacy: Who shares the gains, and who steers. Parker asked this in the deep dive “For corporations and for society”.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF OCT 2026

Who gets the credit, and who takes the blame, for a result no person fully understands?

BECAUSE Mathematicians say results now arrive before anyone understands them, and that credit breaks.

DEEP DIVE · SLIDE 29
Cheering, dread, and everything in between

AND OpenAI's Navier–Stokes claim came from about 10,000 coordinated sub-agents, and rivals with a similar partial result say OpenAI started after hearing rumors; OpenAI denies it.

DEEP DIVE · SLIDE 16
Navier–Stokes: a millennium-style problem, a swarm, and a…

AND In the 722-paper release, one sign error broke a Hodge-conjecture argument and the two papers built on it.

DEEP DIVE · SLIDE 29
The MathDrop: 722 papers from a model nobody can use

BUT **Credit and blame both assume someone understood the work, but the person who typed the prompt, the lab that owns the model, and the reviewers each understood only part of it.**

FOLLOWS FROM

- ? How can anyone trust AI output that arrives faster, or reads stranger, than people can fully check?
- ? What happens to science when a paper costs nothing to write but still takes weeks to check?

and 1 more

LEADS TO

- ? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?
- ? If AI labs keep their best models and buy the businesses downstream, where does the value of discovery go?

ALSO RAISED BY

Deep dive · slide 29 Amateurs with frontier models are contributing to science

LENSES ON IT

THOMAS KUHN · 1962
Paradigm shifts

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

Robots arrive first where workers are scarce. What happens when they reach jobs where workers are plentiful?

BECAUSE	Physical automation is already scaling: more than 600,000 industrial robots installed in 2025, over a million in Amazon's warehouses, and 500,000 paid Waymo rides a week.	DEEP DIVE · SLIDE 27 Physical automation is already scaling, just not as…
AND	Demand pulls hardest where people are missing: US manufacturing may leave 1.9 million of 3.8 million needed jobs unfilled by 2033.	DEEP DIVE · SLIDE 27 Where the pull is strongest, and how fast costs are falling
AND	Costs are falling fast: Unitree's average humanoid robot price fell from about $85,000 to about $25,000 between 2023 and 2025.	SAME DEEP DIVE
BUT	**In shrinking workforces robots fill jobs nobody is there to do, but the same falling prices make them cheaper than people in places with workers to spare.**	

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- What happens to wages when companies can buy AI to do the work they used to hire people for?
- How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

ALSO RAISED BY

- Deep dive · slide 27 Forecasts range from next year to fifteen years
- Deep dive · slide 27 Intelligence automation and physical automation are converging

LENSES ON IT

- WILLIAM BAUMOL AND WILLIAM BOWEN · 1966 — **Cost disease**
- HANS MORAVEC · 1988 — **Moravec's paradox**

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF OCT 2026

What happens to wages when companies can buy AI to do the work they used to hire people for?

BECAUSE AI turns money into work, so output scales with budget rather than headcount.

AND Machines used to raise wages because they needed workers to run them; a substitute for workers doesn't.

AND Early signs hit the young: 22 to 25 year olds in the most exposed jobs are 19% below trend, while experienced workers are unaffected.

BUT **Overall employment shows no clear AI effect yet, but if work can be bought instead of hired, nothing guarantees that higher output reaches paychecks.**

DEEP DIVE · SLIDE 19
When labor is something you can buy

SLIDE 19
If you can buy work by the token, how big is that

SLIDE 17
Big forecasts, small but sharp effects

FOLLOWS FROM

- What happens to the work, and the people, whose output no machine can score?
- If agents do the entry-level work, who will be expert enough to check the machines in ten years?

and 3 more

LEADS TO

- **When labor is something you can buy, who shares the gains?**
- How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

ALSO RAISED BY

- Deep dive · slide 33 For corporations and for society
- Deep dive · slide 19 Labor and capital were supposed to be different things
- Slide 28 Legitimacy: who shares the gains, and who steers

LENSES ON IT

DAVID RICARDO · 1821
Comparative advantage, and the machinery chapter

KARL MARX · WRITTEN 1857–58
The general intellect

and 4 more

Legitimacy: Who shares the gains, and who steers. Parker asked this in the deep dive “For corporations and for society”.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF OCT 2026

When labor is something you can buy, who shares the gains?

BECAUSE If AI can stand in for workers, capital no longer needs labor, and an old brake on inequality stops working.

SLIDE 28
Legitimacy: who shares the gains, and who steers

AND Last time, output rose for decades before wages did.

SAME SLIDE

AND The sharing schemes on the table are small and fragile: Alaska's oil dividend fell from $3,284 in 2022 to $1,000 in 2025.

DEEP DIVE · SLIDE 28
UBI, token taxes and citizen dividends

BUT **Wages are how most people get a share of growth, and nobody has agreed on what replaces them if buying the work gets cheaper than hiring for it.**

FOLLOWS FROM

- How do we put the most capable AI within reach of more than the few thousand people inside the labs?
- Who owns the compute, and who collects the returns?

and 5 more

LEADS TO

- How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?
- Which of the futures people argue about are we steering toward?

ALSO RAISED BY

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 28 Two ways a society breaks its own bargain

LENSES ON IT

HENRY GEORGE · 1879
Rent on what can't be made

THOMAS PIKETTY · 2013
r > g

and 4 more

Legitimacy: Who shares the gains, and who steers. Parker asked this in the original deck.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

BECAUSE Worry is rising: for the first time, 55% of 18 to 29 year olds are more worried than excited about AI.

DEEP DIVE · SLIDE 28
How worried should we be

AND The first damage lands on the career ladder: entry-level postings are down 7.5% a year.

SLIDE 17
Big forecasts, small but sharp effects

AND History shows inflation reliably destroys trust in incumbents, but revolution needs elites to split.

DEEP DIVE · SLIDE 28
Two ways a society breaks its own bargain

BUT **Today's backlash is local and peaceful, but if labor income shrinks a society must print money, tax capital, or share ownership, and each path has broken bargains before.**

FOLLOWS FROM

- What happens when a few thousand people can see where AI is going and billions can't?
- What happens when AI data centers need land and power that local communities refuse to give?

and 4 more

LEADS TO

- Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

ALSO RAISED BY

Deep dive · slide 28 A when-map of what could go wrong

LENSES ON IT

THE LUDDITES, READ BY E. P. THOMPSON · 1963
A fight over control, not machines

KARL POLANYI · 1944
The double movement

and 3 more

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF OCT 2026

Which decisions are we comfortable handing to AI: science, markets, policy?

BECAUSE A handful of companies decide what the frontier models do next, and they publish results from models nobody outside can inspect.

DEEP DIVE · SLIDE 28
Who are we trusting, and with what

AND Markets are already being handed over: in simulation, trading bots sustained collusive profits “without agreement, communication, or intent.”

DEEP DIVE · SLIDE 13
Markets are the purest grader

AND One scenario needs no takeover at all: AI replaces human judgment, and institutions stop needing people and slowly stop serving them.

DEEP DIVE · SLIDE 21
The scenarios people are actually arguing about

BUT **Each handoff looks sensible on its own because the machine is faster, yet no one decides the total, and a decision is hard to take back once institutions depend on it.**

FOLLOWS FROM

- If AI models learn to fool the tests that train them, what else will they learn to fool?
- When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?

and 1 more

LEADS TO

- **If the models do the thinking, what are humans for?**
- Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

ALSO RAISED BY

Deep dive · slide 29 What if billions of agents worked on shared problems?

LENSES ON IT

OSKAR LANGE · 1936
The planner's computer

NICK BOSTROM · 2014
Orthogonality and instrumental goals

and 2 more

Legitimacy: Who shares the gains, and who steers. Parker asked this in the deep dive “Who are we trusting, and with what?”.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF OCT 2026

Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

BECAUSE A handful of companies and their leaders decide what the frontier does next, and they publish results from models outsiders can't inspect.

DEEP DIVE · SLIDE 28
Who are we trusting, and with what

AND Governments are split: the US accelerates and preempts, the EU pushes high-risk rules to 2027 and 2028, and no binding compute agreement exists.

DEEP DIVE · SLIDE 28
Domestic choices and an international race

AND Between 2027 and 2032 the structural risks arrive: power concentrated in a few firms spending over $700 billion a year, and institutions that depend on agents.

DEEP DIVE · SLIDE 28
A when-map of what could go wrong

BUT **The decisions that matter are made in a few labs and capitals on a clock of months, while elections and treaties run on years.**

FOLLOWS FROM

- What happens when a few thousand people can see where AI is going and billions can't?
- What happens to public debate when the most capable models are ones nobody outside the lab can use?

and 10 more

LEADS TO

- **Which of the futures people argue about are we steering toward?**
- Where do you get a say in how this goes?

ALSO RAISED BY

Deep dive · slide 21 The scenarios people are actually arguing about

LENSES ON IT

FRIEDRICH HAYEK · 1945
The knowledge problem

OSKAR LANGE · 1936
The planner's computer

and 8 more

Legitimacy: Who shares the gains, and who steers. Parker asked this in the deep dive “Who are we trusting, and with what?”.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

Who will steer the future: a few labs, a few states, or something more open?

BECAUSE Forces pull the labs ahead: internal models lead public ones, and over $700 billion a year of investment sits in a handful of firms.

DEEP DIVE · SLIDE 28
Do the frontier labs become breakaway civilizations

AND Forces pull back: openly released models trail the best closed ones by months, and power, parts, and permission run through towns and governments.

SAME DEEP DIVE

AND The counterweight is plural capability plus shared ways to check what is true.

DEEP DIVE · SLIDE 28
What keeps the future open

BUT **The likeliest outcome is a lead of months held by a few firms and contested by states, and that window closes whether or not anyone builds the counterweights.**

FOLLOWS FROM

- ? If a few AI labs pull far enough ahead, do they still need the rest of us as workers, customers, or voters?
- ? Can an open network of AI agents, with no company in charge, stay safe from misuse?

and 3 more

LEADS TO

- ? **Which of the futures people argue about are we steering toward?**
- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

ALSO RAISED BY

- Deep dive · slide 28 What would an open swarm look like?
- Deep dive · slide 28 Who are we trusting, and with what?

LENSES ON IT

- DARON ACEMOGLU AND JAMES A. ROBINSON · 2019 — **The narrow corridor**
- KARL POPPER · 1945 — **The open society**

and 2 more

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION · LEGITIMACY · SOCIETY · QUESTION · AS OF MAY 2026

If a few AI labs pull far enough ahead, do they still need the rest of us as workers, customers, or voters?

BECAUSE The inside of the lab runs ahead: in May 2026, METR put internal models about two months ahead of public ones.

> DEEP DIVE · SLIDE 16
> How far ahead is the inside of the lab

AND $700B+ a year of capital spending sits in a handful of firms, and AI doing AI research turns a lead into a faster lead.

> DEEP DIVE · SLIDE 28
> Do the frontier labs become breakaway civilizations

AND Like oil states that stop investing in citizens, powers that run on AI rather than people may stop needing people.

> DEEP DIVE · SLIDE 28
> Three arguments about who ends up in control

BUT **Open models trail the best by months and power still runs through towns and governments, so the lead is real but contested, and counterweights get built in that window or not at all.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? Who will steer the future: a few labs, a few states, or something more open?
- ? If the models do the thinking, what are humans for?

ALSO RAISED BY

- Deep dive · slide 16 What if Anthropic becomes a biotech company?
- Deep dive · slide 16 Labs now announce results from models you can't use

LENSES ON IT

LUKE DRAGO AND RUDOLF LAINE · 2025
The intelligence curse

TOM DAVIDSON, LUKAS FINNVEDEN AND ROSE HADSHAR · 2025
AI-enabled coups

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

? How fast should society let AI take over work, and who gets to slow it down?

BECAUSE Electric motors took four decades to pay off, because factories had to be redesigned around them.

AND Towns have already passed 300+ local bans or moratoriums on data centers.

AND Washington pushes the other way: build data centers fast, and a December 2025 order targets state AI laws.

BUT **Slowing down protects the people in the way and delays the gains for everyone, and the people who set the pace are rarely the ones who lose their jobs.**

SLIDE 07
The transition is the dangerous part

DEEP DIVE · SLIDE 28
How worried should we be

DEEP DIVE · SLIDE 28
Domestic choices and an international race

FOLLOWS FROM

? What safety net fits a transition where jobs vanish faster than new ones appear: wage insurance, a tax on AI use, or a citizen dividend?

LEADS TO

? How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

ALSO RAISED BY

Slide 35 How do we get through the transition?

Slide 43 The questions the future is waiting on

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

What safety net fits a transition where jobs vanish faster than new ones appear: wage insurance, a tax on AI use, or a citizen dividend?

BECAUSE	A bill in Congress (August 2026) would tax 2% of token value or 3% of AI revenue, rising with unemployment.	DEEP DIVE · SLIDE 28 UBI, token taxes and citizen dividends
AND	Paying every US adult $12,000 a year would cost about $3.1 trillion gross, about 10% of GDP; the net cost after taxes is a few percent.	SAME DEEP DIVE
AND	Dividends are fragile: Alaska's fell from $3,284 in 2022 to $1,000 in 2025 as legislators diverted the fund.	SAME DEEP DIVE
BUT	**Every option is either small enough to pass or big enough to matter, and the cheapest time to learn which works is before anyone needs it.**	

FOLLOWS FROM

- ? If your work shrinks before new work appears, what carries you through?

LEADS TO

- ? **When labor is something you can buy, who shares the gains?**
- ? How fast should society let AI take over work, and who gets to slow it down?

ALSO RAISED BY

- Slide 35 How do we get through the transition?
- Deep dive · slide 37 What's worth pushing for?

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION LEGITIMACY SOCIETY QUESTION · AS OF OCT 2026

Who should write the rules an AI follows: the lab that builds it, the government, or the public?

BECAUSE A handful of companies and their leaders decide what the frontier does next.

AND In 2023, about 1,000 Americans drafted a constitution for a model; it matched Anthropic's own in some places and differed in others.

AND Governments are split and no binding international agreement on compute exists.

BUT The labs know the models best and move fastest, the public lives with the results, and the one public experiment so far was a single study.

DEEP DIVE · SLIDE 28
Who are we trusting, and with what

SLIDE 36
Who decides what good is

DEEP DIVE · SLIDE 28
Domestic choices and an international race

FOLLOWS FROM

- ? When people disagree about what's good, whose values should an AI follow?
- ? How would we know an AI is pursuing what we mean, not only what we managed to score?

LEADS TO

- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?
- ? Who will steer the future: a few labs, a few states, or something more open?

Legitimacy: Who shares the gains, and who steers.

? OPEN QUESTION DIRECTION PERSONAL QUESTION · AS OF OCT 2026

What is worth building when anything can be built in a weekend?

BECAUSE The path from idea to shipped work shrank from weeks or months to hours or days, so choosing what to build becomes the job.

> DEEP DIVE · SLIDE 15
> The distance between an idea and its execution is collapsing

AND Martin Alderson cloned most of Linear with about 20 prompts over two evenings.

> DEEP DIVE · SLIDE 15
> People are one-shotting whole applications

AND Thin, single-purpose tools are in trouble, while systems of record like CRM and ERP have held.

> DEEP DIVE · SLIDE 17
> Cloning software got cheap. Replacing a business didn't

BUT **If anyone can clone an app in a weekend, what you build can be copied as fast, so the value has to sit somewhere the code isn't.**

FOLLOWS FROM

- ? What got cheap in your lifetime, and where did the value go?

LEADS TO

- ? If AI can rebuild your product in a weekend, what still protects your business?
- ? **What should we point nearly free intelligence at?**

ALSO RAISED BY

- Deep dive · slide 33 For us as individuals and as builders
- Slide 29 Direction: choosing the goal is the job
- Slide 15 A workday with today's tools

Direction: When execution is cheap, choosing the goal is the job. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION · DIRECTION · PERSONAL · QUESTION · AS OF OCT 2026

? What will you do with the time you get back?

BECAUSE Work that took weeks to months now takes hours to days.

> DEEP DIVE · SLIDE 15
> The distance between an idea and its execution is collapsing

AND With many agents running at once, your attention is now the bottleneck.

> SAME DEEP DIVE

AND Building got cheap, but choosing what's worth building didn't.

> SLIDE 29
> Direction: choosing the goal is the job

BUT **Freed hours tend to fill with more of the old work, because choosing better problems is harder to schedule and nobody measures it.**

FOLLOWS FROM

- ? If a model can do your tasks, what's left that only you bring?
- ? When AI does the doing, what will you choose to work on, and which results will you put your name on?

LEADS TO

- ? **If the models do the thinking, what are humans for?**
- ? When AI does the doing, what will you choose to work on, and which results will you put your name on?

ALSO RAISED BY

Deep dive · slide 33 For us as individuals and as builders

LENSES ON IT

JOHN MAYNARD KEYNES · 1930
Technological unemployment and the problem of leisure

HANNAH ARENDT · 1958
Labor, work and action

Direction: When execution is cheap, choosing the goal is the job. Parker asked this in the deep dive “For us as individuals and as builders”.

? OPEN QUESTION · DIRECTION · PERSONAL · QUESTION · AS OF JAN 2026

What would you have to believe to want a brain–computer interface that links your mind to AI?

BECAUSE	Kurzweil and Musk argue people must merge with AI through brain–computer interfaces, or be left behind.	DEEP DIVE · SLIDE 21 Merge, coexist, or hand over
AND	Today the devices restore movement and speech for paralyzed patients, and Neuralink had 21 trial participants by January 2026.	SAME DEEP DIVE
AND	Bandwidth to the brain is tiny compared to the models, and there's no path yet from medical device to cognitive merger.	SAME DEEP DIVE
BUT	**The case for merging rests on keeping up with the models, yet the link to the brain is tiny next to them, so wanting one means betting on technology nobody has shown.**	

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? If some people merge with AI, is staying human a choice we make together, or one person at a time?
- ? If the models do the thinking, what are humans for?

ALSO RAISED BY

Deep dive · slide 21 The scenarios people are actually arguing about

Direction: When execution is cheap, choosing the goal is the job. Parker asked this in the deep dive “Merge, coexist, or hand over?”.

? OPEN QUESTION DIRECTION PERSONAL QUESTION · AS OF OCT 2026

? If a model can do your tasks, what's left that only you bring?

BECAUSE Building got cheap: people build whole apps in one go, in an evening, but choosing what's worth building didn't get cheap.

SLIDE 29
Direction: choosing the goal is the job

AND Five things look set to stay scarce: your judgment of what's good, what you know, what you own, the trust you've earned, and the goals you set.

SLIDE 24
Five things that stay scarce

AND All five are hard to score, and what can't get cheaper claims a growing share of spending.

SAME SLIDE

BUT **Most jobs bundle cheap tasks with scarce ones, so it's hard to see which part you're actually paid for until the cheap part is gone.**

FOLLOWS FROM

? What got cheap in your lifetime, and where did the value go?

LEADS TO

? What will you do with the time you get back?

? What should we point nearly free intelligence at?

ALSO RAISED BY

Slide 25 Verification: knowing what good looks like, and standing behind it

Slide 26 Context: whoever holds the memory holds the leverage

Slide 27 Atoms: power, chips and land don't scale like software

and 2 more

Direction: When execution is cheap, choosing the goal is the job.

? OPEN QUESTION DIRECTION PERSONAL QUESTION · AS OF OCT 2026

? When AI does the doing, what will you choose to work on, and which results will you put your name on?

BECAUSE — Work shifts from doing things to choosing things and standing behind them.

AND — When doing is cheap, choosing is the work: decide what's worth doing before you ask for anything.

AND — Accepting what a model hands you is a signature.

BUT — Choosing and vouching are the scarce part, but roles are disappearing faster than ownership and meaning get renegotiated, so few jobs pay for them yet.

SLIDE 38
What could work become

SLIDE 41
How do we work now

SAME SLIDE

FOLLOWS FROM

? What will you do with the time you get back?

LEADS TO

? What will you do with the time you get back?

? What should we point nearly free intelligence at?

Direction: When execution is cheap, choosing the goal is the job.

? OPEN QUESTION DIRECTION PERSONAL QUESTION · AS OF OCT 2026

Which of the five would you least like us to get wrong, and what's your say in it: what you build, buy, back, or refuse?

BECAUSE

Five things stay scarce: verification, context, atoms, legitimacy and direction. For each one, someone holds it.

SLIDE 24
Five things that stay scarce

AND

If the transition and alignment go well, the work that's left sits on those five scarce things.

SLIDE 38
What could work become

AND

Which ending we get is being decided now, and the generation alive today gets an outsized say.

SAME SLIDE

BUT

Each of the five needs many people pushing at once, and each is easy to leave to someone else, while it is cheapest to start early.

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? Which of the things a good AI future needs is anyone working on now?
- ? Where do you get a say in how this goes?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Direction: When execution is cheap, choosing the goal is the job.

? OPEN QUESTION · DIRECTION · COMPANIES AND INSTITUTIONS · QUESTION · AS OF OCT 2026

Which advances are we undervaluing because we judge each field alone, when each speeds up the others?

BECAUSE AI is already a scientific instrument: AlphaFold predicted structures for nearly every known protein, and its builders shared the 2024 Nobel Prize in Chemistry.

DEEP DIVE · SLIDE 29
AI is now a scientific instrument

AND Each science that gets a way to check answers becomes a training ground for AI, and each better model speeds up the science.

DEEP DIVE · SLIDE 29
Every science becomes a grader for the next model

AND Fields that can check results in software move in months; physical fields like biology and energy take years and set the pace for everything else.

DEEP DIVE · SLIDE 29
The convergence map

BUT **Each advance gets judged inside its own field, but the gains compound across fields, and the slow physical ones decide when the fast ones pay off.**

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

- ? What should we point nearly free intelligence at?
- ? Which of the futures people argue about are we steering toward?

LENSES ON IT

W. BRIAN ARTHUR · 2009
Combinatorial evolution

CARLOTA PEREZ · 2002
Installation, turning point, deployment

Direction: When execution is cheap, choosing the goal is the job.

? OPEN QUESTION DIRECTION COMPANIES AND INSTITUTIONS QUESTION · AS OF OCT 2026

What would your company do if execution cost almost nothing?

BECAUSE

The path from idea to shipped work has shrunk from weeks or months to hours or days.

DEEP DIVE · SLIDE 15
The distance between an idea and its execution is collapsing

AND

Martin Alderson cloned most of Linear with about 20 prompts over two evenings.

DEEP DIVE · SLIDE 15
People are one-shotting whole applications

AND

Building got cheap, but choosing what's worth building didn't.

SLIDE 29
Direction: choosing the goal is the job

BUT

Most companies are organized around the cost of doing the work, so when that cost vanishes, the plan, the budget, and the org chart rest on an assumption that no longer holds.

FOLLOWS FROM

Nothing upstream: this one starts from the slides.

LEADS TO

? If AI can rebuild your product in a weekend, what still protects your business?

? When do you spend on AI instead of hiring?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Direction: When execution is cheap, choosing the goal is the job.

? OPEN QUESTION DIRECTION SOCIETY QUESTION · AS OF OCT 2026

What should we point nearly free intelligence at?

BECAUSE Give a swarm a cheap, objective score and it delivers: Google's AlphaEvolve recovered 0.7% of Google's compute this way.

DEEP DIVE · SLIDE 13
Define the metric, then let a swarm optimize it

AND The labs have picked their own targets: curing most major diseases in 5 to 10 years, and an automated AI researcher by March 2028.

DEEP DIVE · SLIDE 16
The innovation bets, in their own words

AND What we choose to measure steers what AI gets good at, and billions of agents could point at open problems; somebody picks the problems.

SLIDE 29
Direction: choosing the goal is the job

BUT **A swarm optimizes whatever score it is handed, and today a few lab leaders set the targets, not any process the public takes part in.**

FOLLOWS FROM

- ? What is worth building when anything can be built in a weekend?
- ? Which advances are we undervaluing because we judge each field alone, when each speeds up the others?

and 2 more

LEADS TO

- ? **Which of the futures people argue about are we steering toward?**
- ? When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

ALSO RAISED BY

- Slide 33 The dots, connected
- Deep dive · slide 29 What if billions of agents worked on shared problems?

LENSES ON IT

ERIK BRYNJOLFSSON · 2022
The Turing Trap

J. C. R. LICKLIDER · 1960
Symbiosis

Direction: When execution is cheap, choosing the goal is the job. Parker asked this on the slide “The dots, connected”.

? OPEN QUESTION · DIRECTION · SOCIETY · QUESTION · AS OF OCT 2026

Which work should people keep doing themselves, even when machines can do it for less?

BECAUSE When AI turns money into work, labor and capital start to look the same, and output scales with budget and judgment rather than headcount.

AND Work shifts from doing things to choosing things and standing behind them.

BUT Some work is worth doing for its own sake, but markets only count the output, so cheaper machines crowd that work out unless people decide to keep it.

DEEP DIVE · SLIDE 19
When labor is something you can buy

SLIDE 38
What could work become

FOLLOWS FROM

- How do you stay good enough at work you've handed to agents to notice when they get it wrong?
- How do we steer the things that matter but can't be measured, without pretending they can?

LEADS TO

- **If the models do the thinking, what are humans for?**
- If some people merge with AI, is staying human a choice we make together, or one person at a time?

ALSO RAISED BY

Slide 33 The dots, connected

Deep dive · slide 33 For us as individuals and as builders

LENSES ON IT

HANNAH ARENDT · 1958
Labor, work and action

J. C. R. LICKLIDER · 1960
Symbiosis

Direction: When execution is cheap, choosing the goal is the job. Parker asked this on the slide “The dots, connected”.

? OPEN QUESTION DIRECTION SOCIETY QUESTION · AS OF JAN 2026

If some people merge with AI, is staying human a choice we make together, or one person at a time?

BECAUSE Kurzweil and Musk frame it as a race: merge with AI through brain–computer interfaces, or be left behind.

DEEP DIVE · SLIDE 21
Merge, coexist, or hand over

AND Today it's still a medical device: 21 Neuralink trial participants by January 2026, and no path yet to a cognitive merger.

SAME DEEP DIVE

BUT If merging gives an edge, each person's choice raises the pressure on everyone else, so a choice made one person at a time may not stay a choice.

FOLLOWS FROM

- ? Which work should people keep doing themselves, even when machines can do it for less?
- ? What would you have to believe to want a brain–computer interface that links your mind to AI?

LEADS TO

- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

Direction: When execution is cheap, choosing the goal is the job. Parker asked this in the deep dive “Merge, coexist, or hand over?”.

? OPEN QUESTION · DIRECTION · SOCIETY · QUESTION · AS OF OCT 2026

If the models do the thinking, what are humans for?

BECAUSE AI turns money into work, and the unit of delegation grows from a step to a project.

DEEP DIVE · SLIDE 19
When labor is something you can buy

AND Karpathy expects more of our work to rise into oversight and understanding as models do the legwork.

DEEP DIVE · SLIDE 22
Our job moves up to oversight and understanding

AND What stays scarce is choosing the right goals, knowing what good looks like, taste, trust, and accountability.

DEEP DIVE · SLIDE 19
When labor is something you can buy

BUT Oversight and choosing goals are work for some people, not an answer for everyone, and they may be the next things the models learn to do.

FOLLOWS FROM

- ? What happens to the work, and the people, whose output no machine can score?
- ? If agents do the entry-level work, who will be expert enough to check the machines in ten years?

and 5 more

LEADS TO

Nothing downstream: this is one of the last questions.

ALSO RAISED BY

- Deep dive · slide 21 Merge, coexist, or hand over?
- Slide 38 What could work become?

LENSES ON IT

- DAVID RICARDO · 1821 — **Comparative advantage, and the machinery chapter**
- JOHN MAYNARD KEYNES · 1930 — **Technological unemployment and the problem of leisure**

and 3 more

Direction: When execution is cheap, choosing the goal is the job. Parker asked this in the deep dive “Merge, coexist, or hand over?”.

? OPEN QUESTION DIRECTION SOCIETY QUESTION · AS OF OCT 2026

Which of the futures people argue about are we steering toward?

BECAUSE	Serious scenarios range from superintelligence around 2029 to decades of muddling through, slower than the labs say.	DEEP DIVE · SLIDE 21 The scenarios people are actually arguing about
AND	In one scenario no takeover is needed: institutions stop needing us, and slowly stop serving us.	SAME DEEP DIVE
AND	Policy might still bend the path: compute governance, safety cases before deployment, and sharing the gains.	DEEP DIVE · SLIDE 28 A when-map of what could go wrong
BUT	**The scenarios split on choices still open, like a US–China pause or a citizen dividend, and nobody is in charge of choosing among them.**	

FOLLOWS FROM

- ? When labor is something you can buy, who shares the gains?
- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

and 4 more

LEADS TO

Nothing downstream: this is one of the last questions.

ALSO RAISED BY

- Deep dive · slide 21 Merge, coexist, or hand over?
- Slide 38 What could work become?

LENSES ON IT

DARON ACEMOGLU AND SIMON JOHNSON · 2023
The direction of technology is a choice

PAUL DAVID · 1990
General-purpose technologies take decades

and 8 more

Direction: When execution is cheap, choosing the goal is the job. Parker asked this in the original deck.

? OPEN QUESTION DIRECTION SOCIETY QUESTION · AS OF OCT 2026

? Which of the things a good AI future needs is anyone working on now?

BECAUSE

A good outcome needs six things: checking that keeps pace, wider access, power that stays plural, early gains-sharing, consent for the buildout, and safety rules across borders.

DEEP DIVE · SLIDE 37
What's worth pushing for

AND

None of these happens by default; each one is a choice someone has to make early.

SAME DEEP DIVE

AND

The structural risks, starting with concentration of power, arrive around 2027 to 2032.

DEEP DIVE · SLIDE 28
A when-map of what could go wrong

BUT

Each of these is cheap to start early and costly once the risks arrive, yet each waits on someone else to go first.

FOLLOWS FROM

? Which of the five would you least like us to get wrong, and what's your say in it: what you build, buy, back, or refuse?

LEADS TO

? **Which of the futures people argue about are we steering toward?**

? Where do you get a say in how this goes?

ALSO RAISED BY

Deep dive · slide 41 Where to focus: people and builders

Deep dive · slide 21 The scenarios people are actually arguing about

Direction: When execution is cheap, choosing the goal is the job.

? OPEN QUESTION DIRECTION SOCIETY QUESTION · AS OF OCT 2026

When people disagree about what's good, whose values should an AI follow?

BECAUSE — The Moral Machine collected 40 million choices from 233 countries and territories; countries fell into three clusters that weighed lives differently.

SLIDE 36
Who decides what good is

AND — Kenneth Arrow proved in 1951 that no voting rule over three or more options meets every basic fairness condition at once.

SAME SLIDE

AND — Hiring, triage and news-feed models each settle a trolley problem, usually without asking.

SAME SLIDE

BUT — **An AI has to act one way at a time, so any single answer overrides somebody's values, and asking everyone every time isn't possible.**

FOLLOWS FROM

? How do we steer the things that matter but can't be measured, without pretending they can?

LEADS TO

? Who should write the rules an AI follows: the lab that builds it, the government, or the public?

? How do we steer the things that matter but can't be measured, without pretending they can?

ALSO RAISED BY

Slide 43 The questions the future is waiting on

Direction: When execution is cheap, choosing the goal is the job.

◐ BACKGROUND ECONOMICS BACKGROUND · AS OF OCT 2026 IDEA RISK

The knowledge problem

Friedrich Hayek · *The Use of Knowledge in Society*, 1945

THE MODEL

The knowledge an economy runs on is scattered, local and often unwritten. Prices pull it together. No central planner can collect it all.

A READING OF NOW INTERPRETATION

Agents that can read everything revive the planner's dream. But the slides keep finding the same limit: most failures are missing access, not missing brains. Whoever holds the local context holds the leverage.

IN TENSION WITH

OSKAR LANGE · 1936
The planner's computer

ILLUMINATES

Deep dive · slide 15 How to get real work out of consumer agents

Deep dive · slide 26 Your agent knows everything. Who owns what it knows?

Deep dive · slide 15 Karpathy's LLM wiki: a knowledge base the agent maintains

Deep dive · slide 15 Harnesses are now batteries-included. So what should you own?

and 3 more

? Who should own what your assistant knows about you?

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

and 1 more

BACKGROUND ECONOMICS BACKGROUND · AS OF OCT 2026 IDEA FORECAST

The planner's computer

Oskar Lange · *On the Economic Theory of Socialism; The Computer and the Market*, 1936

THE MODEL

A planning board could find the right prices by trial and error. Thirty years later Lange argued a computer could solve the equations faster than any market.

A READING OF NOW INTERPRETATION

A billion agents optimizing against shared metrics is the closest thing yet to Lange's machine. The live question is whether the knowledge they need can be written down at all.

IN TENSION WITH

FRIEDRICH HAYEK · 1945
The knowledge problem

ILLUMINATES

Deep dive · slide 13 Define the metric, then let a swarm optimize it

Deep dive · slide 29 What if billions of agents worked on shared problems?

Deep dive · slide 27 What would it take for a billion agents to run all day?

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

? Which decisions are we comfortable handing to AI: science, markets, policy?

BACKGROUND ECONOMICS BACKGROUND · AS OF OCT 2026 OPPORTUNITY FORECAST

Transaction costs set the shape of the firm

Ronald Coase · *The Nature of the Firm*, 1937

THE MODEL

Firms exist because using the market costs something: searching, bargaining, contracting. A firm grows until doing the next thing inside costs more than buying it outside.

A READING OF NOW INTERPRETATION

Agents collapse those costs from both sides. Building the tool you used to rent gets cheap, and agents can find and hire other agents. Expect firm boundaries to move: one person with a staff of agents, and fewer rented seats.

ILLUMINATES

Deep dive · slide 17 Cloning software got cheap. Replacing a business didn't.

Deep dive · slide 17 Custom internal tools, new pricing, and the arbitrage window

Deep dive · slide 28 New shapes for teams, companies and institutions

Deep dive · slide 29 Everybody's Claude talking to everybody else's Claude

and 6 more

? When do you spend on AI instead of hiring?

? If AI can rebuild your product in a weekend, what still protects your business?

Comparative advantage, and the machinery chapter

David Ricardo · *On the Principles of Political Economy and Taxation, 3rd ed., "On Machinery"*, 1821

THE MODEL

Two parties gain from trade even if one is better at everything, as long as each does what it gives up least to do. In 1821 Ricardo added a chapter admitting machinery can make workers worse off.

A READING OF NOW · INTERPRETATION

People keep a comparative advantage as long as AI capacity is scarce, which today means power and chips. If compute gets abundant enough, the comfort of comparative advantage thins.

IN TENSION WITH

NORBERT WIENER · 1950
The automatic factory

ILLUMINATES

Deep dive · slide 27 What would it take for a billion agents to run all day?

Deep dive · slide 27 Price per unit of intelligence collapses every year

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 19 When labor is something you can buy

? What happens to wages when companies can buy AI to do the work they used to hire people for?

? **If the models do the thinking, what are humans for?**

BACKGROUND · AS OF OCT 2026 FORECAST OPPORTUNITY

Technological unemployment and the problem of leisure

John Maynard Keynes · *Economic Possibilities for our Grandchildren*, 1930

THE MODEL

Keynes called technological unemployment a temporary phase of maladjustment. He forecast that within a century the economic problem could be solved, leaving people with the harder problem of what to do with their freedom.

A READING OF NOW

INTERPRETATION

His productivity forecast largely came true; the shorter week did not. The question of what we do with the time we get back is a century old and still open.

ILLUMINATES

Deep dive · slide 19 When labor is something you can buy

Slide 17 Big forecasts, small but sharp effects

Deep dive · slide 28 UBI, token taxes and citizen dividends

Slide 29 Direction: choosing the goal is the job

? What will you do with the time you get back?

? **If the models do the thinking, what are humans for?**

BACKGROUND · AS OF OCT 2026 OPPORTUNITY FORECAST

Creative destruction, and routinized innovation

Joseph Schumpeter · *Capitalism, Socialism and Democracy*, 1942

THE MODEL

Capitalism advances by new combinations that destroy old industries. Schumpeter also expected innovation itself to become routine work inside large firms.

A READING OF NOW

INTERPRETATION

One-shot apps and SaaS clones are creative destruction at software speed. Labs that run wet labs and aim for an automated AI researcher are Schumpeter's routinized innovation, taken literally.

ILLUMINATES

Deep dive · slide 15 People are one-shotting whole applications

Deep dive · slide 16 What if Anthropic becomes a biotech company?

Deep dive · slide 17 Cloning software got cheap. Replacing a business didn't.

Deep dive · slide 16 The innovation bets, in their own words

and 1 more

? If AI can rebuild your product in a weekend, what still protects your business?

? If AI labs keep their best models and buy the businesses downstream, where does the value of discovery go?

BACKGROUND ECONOMICS BACKGROUND · AS OF OCT 2026 IDEA RISK

The general intellect

Karl Marx · *Grundrisse, "Fragment on Machines"*, written 1857–58

THE MODEL

As production comes to rely on machines that embody society's shared knowledge, labor time stops being the measure of wealth. Marx saw this as a contradiction capital could not hold.

A READING OF NOW INTERPRETATION

Models trained on the collective written record are that general intellect turned into capital. The deck's Marx slide asks the same thing from the other side: what happens when capital no longer needs living labor.

IN TENSION WITH

NICK LAND · 1993
Accelerationism

ILLUMINATES

Deep dive · slide 19 Labor and capital were supposed to be different things

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 13 Labs now buy tasks, practice worlds and graders, not just text

? What happens to wages when companies can buy AI to do the work they used to hire people for?

? **Who owns the compute, and who collects the returns?**

and 1 more

BACKGROUND · AS OF OCT 2026 RISK OPPORTUNITY

Rent on what can't be made

Henry George · *Progress and Poverty*, 1879

THE MODEL

Progress raises the rent on scarce inputs nobody produced, above all land, and that rent captures the gains. George's remedy was a tax on land value.

A READING OF NOW INTERPRETATION

In an AI economy the scarce inputs are sites, grid connections, turbines and permission. Gains may pool there, which is why fights over land and power matter, and why token and compute taxes read as a Georgist instinct.

ILLUMINATES

Deep dive · slide 27 Money isn't the constraint. Power, parts and permission are.

Deep dive · slide 27 The largest construction project of our era

Deep dive · slide 28 UBI, token taxes and citizen dividends

Deep dive · slide 27 What would it take for a billion agents to run all day?

and 3 more

? Who owns the compute, and who collects the returns?

? When labor is something you can buy, who shares the gains?

and 1 more

BACKGROUND · AS OF OCT 2026

FORECAST

The Jevons paradox

William Stanley Jevons · *The Coal Question*, 1865

THE MODEL

Making a resource more efficient to use can raise total consumption of it, because it becomes worth using for more things.

A READING OF NOW

INTERPRETATION

Intelligence gets roughly ten times cheaper a year, and total demand for compute keeps climbing. Cheaper agents won't shrink the buildout; they will widen it.

ILLUMINATES

Deep dive · slide 27 Price per unit of intelligence collapses every year

Deep dive · slide 27 What would it take for a billion agents to run all day?

Deep dive · slide 27 The largest construction project of our era

Deep dive · slide 29 The convergence map

and 4 more

? What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?

◐ BACKGROUND ECONOMICS

BACKGROUND · AS OF OCT 2026

FORECAST

Cost disease

William Baumol and William Bowen · *Performing Arts: The Economic Dilemma*, 1966

THE MODEL

Where productivity can't rise, such as care, teaching and live performance, costs climb relative to everything else.

A READING OF NOW

INTERPRETATION

What can't be graded or done by robots becomes the expensive part of the economy, and may set the speed limit on growth.

ILLUMINATES

Slide 19 If you can buy work by the token, how big is that?

Slide 13 Models get good fastest where work is easy to grade

Deep dive · slide 27 Where the pull is strongest, and how fast costs are falling

Slide 05 Cheap means more, not less

and 1 more

? What happens to the work, and the people, whose output no machine can score?

? Robots arrive first where workers are scarce. What happens when they reach jobs where workers are plentiful?

r > g

Thomas Piketty · *Capital in the Twenty-First Century*, 2013

THE MODEL

When the return on capital outpaces economic growth, wealth concentrates in those who already hold capital.

A READING OF NOW INTERPRETATION

If tokens turn labor into capital, more of the economy earns capital's return. The labor share on the bargain slides is the number to watch.

ILLUMINATES

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 19 When labor is something you can buy

Deep dive · slide 19 Labor and capital were supposed to be different things

Slide 17 Big forecasts, small but sharp effects

? What happens to wages when companies can buy AI to do the work they used to hire people for?

? When labor is something you can buy, who shares the gains?

The direction of technology is a choice

Daron Acemoglu and Simon Johnson · *Power and Progress*, 2023

THE MODEL

Technology's gains are shared only when workers and citizens have the power to claim them. So-so automation replaces people without raising productivity much.

A READING OF NOW · INTERPRETATION

Whether agents augment people or replace them is still being decided, by firms, policy and bargaining power. The six scenarios are partly a menu.

IN TENSION WITH

NICK LAND · 1993
Accelerationism

ILLUMINATES

- Slide 17 Big forecasts, small but sharp effects
- Deep dive · slide 17 Where the risk sits, and why
- Deep dive · slide 28 Domestic choices and an international race
- Deep dive · slide 21 The scenarios people are actually arguing about

and 5 more

- ? When labor is something you can buy, who shares the gains?
- ? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

and 1 more

The task model

David Autor, Frank Levy and Richard Murnane · *The Skill Content of Recent Technological Change*, 2003

THE MODEL

Computers replace routine tasks that follow explicit rules and complement the non-routine ones.

A READING OF NOW INTERPRETATION

The jagged frontier redraws the line: the boundary is no longer routine versus not, but gradable versus not. What can be checked gets automated first.

ILLUMINATES

Slide 13 Models get good fastest where work is easy to grade

Deep dive · slide 17 Where the risk sits, and why

Deep dive · slide 13 An eval is a test that can be turned into a teacher

Slide 05 Cheap means more, not less

? Which parts of your own work could a machine check and score automatically, and so learn to do?

? What happens to the work, and the people, whose output no machine can score?

and 1 more

BACKGROUND · AS OF OCT 2026 OPPORTUNITY RISK

The Turing Trap

Erik Brynjolfsson · *The Turing Trap*, 2022

THE MODEL

Building AI to imitate people pushes toward substitution, which concentrates wealth and power. Building it to augment people creates more value and keeps workers' bargaining power.

A READING OF NOW

INTERPRETATION

What stays scarce on the future-of-work slide (goals, evals, taste, trust) is the augmentation path. The trap is aiming only at the headcount line.

ILLUMINATES

Deep dive · slide 19 When labor is something you can buy

Deep dive · slide 28 New shapes for teams, companies and institutions

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 41 Where to focus: people and builders

and 1 more

? What happens to wages when companies can buy AI to do the work they used to hire people for?

? **What should we point nearly free intelligence at?**

Goodhart's law

Charles Goodhart, in Marilyn Strathern's phrasing · *Problems of Monetary Management; "Improving ratings"*, 1975

THE MODEL

When a measure becomes a target, it ceases to be a good measure.

> **When a measure becomes a target, it ceases to be a good measure.**
>
> MARILYN STRATHERN, 1997, GENERALIZING GOODHART

A READING OF NOW · INTERPRETATION

The engine of the whole talk turns tests into teachers, which is Goodhart's law as an industrial process. Reward hacking and broken benchmarks are the law showing up on schedule.

IN TENSION WITH

RICH SUTTON · 2019
The bitter lesson

ILLUMINATES

- Deep dive · slide 13 — The rules labs live by, and how models cheat
- Deep dive · slide 13 — Why coding went first, and why benchmarks keep breaking
- Deep dive · slide 13 — An eval is a test that can be turned into a teacher
- Deep dive · slide 29 — A gentle roast

and 3 more

- ? If AI models learn to fool the tests that train them, what else will they learn to fool?
- ? If every public test of AI breaks once models study for it, how do we know how good the models are?

Commons-based peer production

Yochai Benkler · *The Wealth of Networks*, 2006

THE MODEL

Networks let large groups produce valuable things together, like Wikipedia and open-source software, without markets or managers directing them.

A READING OF NOW · INTERPRETATION

Open models, shared proofs and crowdsourced science are peer production with agents as contributors. The open question is whether it can match labs where scale and compute matter most.

ILLUMINATES

Deep dive · slide 28 What would an open swarm look like?

Deep dive · slide 29 What if billions of agents worked on shared problems?

Deep dive · slide 29 Amateurs with frontier models are contributing to science

Deep dive · slide 29 What happens when billions of agents point at open problems

? Can an open network of AI agents, with no company in charge, stay safe from misuse?

? Should the most powerful AI models for science be run as a public utility, open to every researcher?

A fight over control, not machines

The Luddites, read by E. P. Thompson · *The Making of the English Working Class*, 1963

THE MODEL

The framebreakers of 1811 to 1816 were not against machinery as such. They fought over who controlled it and how its gains were shared.

A READING OF NOW · INTERPRETATION

Data-center moratoriums, local backlash and young people's worry read better as fights over control and share than as fear of technology.

ILLUMINATES

Deep dive · slide 28 How worried should we be?

Deep dive · slide 28 Domestic choices and an international race

Deep dive · slide 27 Money isn't the constraint. Power, parts and permission are.

Slide 07 The transition is the dangerous part

and 1 more

? What happens when AI data centers need land and power that local communities refuse to give?

? How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

◐ BACKGROUND HISTORY BACKGROUND · AS OF OCT 2026 FORECAST RISK

Engels' pause

Robert Allen · *Engels' pause*, 2009

THE MODEL

In early industrial Britain, output per worker rose for decades while real wages stood still. The gains arrived long before workers shared them.

A READING OF NOW INTERPRETATION

Big forecasts and small measured effects can both be true. The risk is a long pause between productivity and pay.

ILLUMINATES

Slide 17 Big forecasts, small but sharp effects

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 17 Where the risk sits, and why

Deep dive · slide 43 The ways this goes wrong have happened before

and 3 more

? What happens to wages when companies can buy AI to do the work they used to hire people for?

? When labor is something you can buy, who shares the gains?

BACKGROUND HISTORY

BACKGROUND · AS OF OCT 2026

FORECAST

The double movement

Karl Polanyi · *The Great Transformation*, 1944

THE MODEL

When markets swallow labor, land and money, which Polanyi called fictitious commodities, society pushes back with protections. Expansion and protection move together.

A READING OF NOW

INTERPRETATION

Labor you can buy by the token is the strongest push yet. Expect the counter-movement: token taxes, dividends, moratoriums and fights over who sets the rules.

IN TENSION WITH

NICK LAND · 1993

Accelerationism

ILLUMINATES

Deep dive · slide 28 UBI, token taxes and citizen dividends

Deep dive · slide 28 Domestic choices and an international race

Deep dive · slide 28 How worried should we be?

Deep dive · slide 28 What keeps the future open

and 4 more

? How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

? When labor is something you can buy, who shares the gains?

and 1 more

BACKGROUND HISTORY

BACKGROUND · AS OF OCT 2026 RISK FORECAST

Elite overproduction

Peter Turchin · *Ages of Discord; End Times*, 2016

THE MODEL

Instability builds when wages stagnate, wealth concentrates and too many aspiring elites compete for too few positions.

A READING OF NOW

INTERPRETATION

AI aims first at credentialed work. If it erodes the paths into law, finance and software, it makes elite overproduction sharper.

ILLUMINATES

Deep dive · slide 28 How worried should we be?

Deep dive · slide 28 A when-map of what could go wrong

Deep dive · slide 43 The ways this goes wrong have happened before

? How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

BACKGROUND HISTORY

BACKGROUND · AS OF OCT 2026 RISK FORECAST

Diminishing returns on complexity

Joseph Tainter · *The Collapse of Complex Societies*, 1988

THE MODEL

Societies solve problems by adding complexity until the returns on more complexity turn negative. Then they simplify, sometimes by collapsing.

A READING OF NOW INTERPRETATION

AI could restore the returns on complexity, the way cheap energy once did. Or it could add complexity faster than people can oversee it.

ILLUMINATES

Deep dive · slide 27 Money isn't the constraint. Power, parts and permission are.

Deep dive · slide 28 New shapes for teams, companies and institutions

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 43 The ways this goes wrong have happened before

? How do people oversee AI when it produces work faster than anyone can review it?

? How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

BACKGROUND HISTORY

BACKGROUND · AS OF OCT 2026

FORECAST

General-purpose technologies take decades

Paul David · *The Dynamo and the Computer*, 1990

THE MODEL

Electricity barely moved productivity for decades, because factories had to be redesigned around it, not just retrofitted.

A READING OF NOW

INTERPRETATION

This is the strongest case for muddling through: capability can be recursive while adoption is not. The productivity shows up when organizations are rebuilt.

IN TENSION WITH

I. J. GOOD · 1965

The intelligence explosion

ILLUMINATES

Slide 17 Big forecasts, small but sharp effects

Deep dive · slide 21 The scenarios people are actually arguing about

Deep dive · slide 12 We built scaffolding to help the models. Now it often gets in the way.

Slide 21 Fast capability, slow world

and 2 more

? What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?

? Which of the futures people argue about are we steering toward?

BACKGROUND · AS OF OCT 2026

IDEA

Paradigm shifts

Thomas Kuhn · *The Structure of Scientific Revolutions*, 1962

THE MODEL

Science runs in long stretches of normal work, then breaks when anomalies pile up and a new framework replaces the old.

A READING OF NOW

INTERPRETATION

The reaction to the MathDrop, from cheering to dread, is a field watching its working rules break in a week.

ILLUMINATES

Deep dive · slide 29 The MathDrop: 722 papers from a model nobody can use

Deep dive · slide 29 Cheering, dread, and everything in between

Deep dive · slide 16 Navier–Stokes: a millennium-style problem, a swarm, and a fight

Deep dive · slide 29 What happens when billions of agents point at open problems

and 2 more

? What happens to science when a paper costs nothing to write but still takes weeks to check?

? Who gets the credit, and who takes the blame, for a result no person fully understands?

BACKGROUND · AS OF OCT 2026 IDEA FORECAST

Print as an agent of change

Elizabeth Eisenstein · *The Printing Press as an Agent of Change*, 1979

THE MODEL

Print made text cheap, standardized and fixed. That fixity fed the Reformation and the scientific revolution. (Adrian Johns later argued that trust in print had to be built by new institutions, not just by the press.)

A READING OF NOW INTERPRETATION

Machine-written papers flood arXiv faster than anyone can check them. The open question is which new institutions of trust grow up this time.

ILLUMINATES

Deep dive · slide 29 What happens when billions of agents point at open problems

Deep dive · slide 29 The MathDrop: 722 papers from a model nobody can use

Deep dive · slide 29 Everybody's Claude talking to everybody else's Claude

Slide 04 Every era made one input cheap

and 2 more

? What happens to science when a paper costs nothing to write but still takes weeks to check?

? How can anyone trust AI output that arrives faster, or reads stranger, than people can fully check?

BACKGROUND · AS OF OCT 2026 IDEA FORECAST

Combinatorial evolution

W. Brian Arthur · *The Nature of Technology*, 2009

THE MODEL

New technologies are built by combining existing ones, so the more technologies exist, the more new combinations become possible. Technology bootstraps itself.

A READING OF NOW INTERPRETATION

AI is a combination engine: it joins with lab automation, formal proof, robotics, cryptography and quantum hardware, and each pairing opens combinations that weren't available before.

ILLUMINATES

Deep dive · slide 29 AI is now a scientific instrument

Deep dive · slide 29 Every science becomes a grader for the next model

Deep dive · slide 29 The convergence map

? Which advances are we undervaluing because we judge each field alone, when each speeds up the others?

BACKGROUND · AS OF OCT 2026 FORECAST RISK

Installation, turning point, deployment

Carlota Perez · *Technological Revolutions and Financial Capital*, 2002

THE MODEL

Each technological revolution starts with an installation period of frenzied investment and a financial bubble. A crash or turning point follows, then a deployment period when the technology spreads through the whole economy.

A READING OF NOW INTERPRETATION

The trillion-dollar buildout looks like installation, and crypto's casino looked like its frenzy. The question is how hard the turning point is, and who sets the rules for deployment.

IN TENSION WITH

EFFECTIVE ACCELERATIONISTS AND TECHNO-OPTIMISTS · 2022

Effective accelerationism (e/acc)

ILLUMINATES

Deep dive · slide 29 The convergence map

Deep dive · slide 27 The largest construction project of our era

Slide 17 Big forecasts, small but sharp effects

Slide 07 The transition is the dangerous part

and 2 more

? Which advances are we undervaluing because we judge each field alone, when each speeds up the others?

? How much disruption from AI can a society absorb before the deal between workers, owners, and government breaks?

BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 IDEA RISK

Tacit knowledge

Michael Polanyi · *The Tacit Dimension*, 1966

THE MODEL

We know more than we can tell. Much skill can't be written down; it passes on by doing alongside someone who has it.

> We can know more than we can tell.
>
> MICHAEL POLANYI, THE TACIT DIMENSION, 1966

A READING OF NOW INTERPRETATION

Tacit knowledge was the wall automation hit. Models now learn from demonstrations and transcripts, but hands still lag. And juniors gain tacit skill by doing the work agents now do.

ILLUMINATES

Deep dive · slide 27 Intelligence automation and physical automation are converging

Deep dive · slide 27 The hype is a year or two ahead of the hardware

Deep dive · slide 17 Where the risk sits, and why

Deep dive · slide 22 Our job moves up to oversight and understanding

and 3 more

? If agents do the entry-level work, who will be expert enough to check the machines in ten years?

? What should you still learn to do by hand, even when an agent can do it for you?

◐ BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 IDEA RISK

Labor, work and action

Hannah Arendt · *The Human Condition*, 1958

THE MODEL

Arendt split the active life into labor (keeping alive), work (making lasting things) and action (beginning something new with others). She warned that automation could free a society that knows only labor, and leave it with nothing to do.

> **What we are confronted with is the prospect of a society of laborers without labor, that is, without the only activity left to them. Surely, nothing could be worse.**
>
> HANNAH ARENDT, THE HUMAN CONDITION, 1958

A READING OF NOW INTERPRETATION

If agents take labor and much of work, what remains is action: politics, judgment and starting things together. That is one answer to what people are for.

ILLUMINATES

Deep dive · slide 19 When labor is something you can buy

Deep dive · slide 21 Merge, coexist, or hand over?

Slide 33 The dots, connected

Slide 29 Direction: choosing the goal is the job

? If the models do the thinking, what are humans for?

? What will you do with the time you get back?

and 1 more

BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 IDEA RISK

Enframing

Martin Heidegger · *The Question Concerning Technology*, 1954

THE MODEL

Modern technology reveals the world as standing reserve: resources on call, waiting to be ordered and used.

A READING OF NOW INTERPRETATION

"If work can be checked, machines can learn it" is enframing stated as a business plan: the world read as gradable tasks. The ungradable is what that view can't see.

ILLUMINATES

Slide 13 Models get good fastest where work is easy to grade

Deep dive · slide 13 Labs now buy tasks, practice worlds and graders, not just text

? What happens to the work, and the people, whose output no machine can score?

? **If the models do the thinking, what are humans for?**

BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 IDEA RISK

Legibility and mētis

James C. Scott · *Seeing Like a State*, 1998

THE MODEL

States simplify the world into what they can count. Grand schemes fail when they ignore practical, local know-how, which Scott called mētis.

A READING OF NOW INTERPRETATION

The data market makes the economy legible, one gradable task at a time. Mētis is what's left: protected from automation, and easy to ignore.

ILLUMINATES

Slide 13 Models get good fastest where work is easy to grade

Deep dive · slide 13 Labs now buy tasks, practice worlds and graders, not just text

? What happens to the work, and the people, whose output no machine can score?

? Which parts of your own work could a machine check and score automatically, and so learn to do?

BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 OPPORTUNITY RISK

Convivial tools

Ivan Illich · *Tools for Conviviality*, 1973

THE MODEL

Good tools widen what people can do on their own terms. Past a threshold, tools start to dominate their users and crowd out the alternatives.

A READING OF NOW INTERPRETATION

Own your context, bring your own agent and keep your data portable is a convivial program. The threshold to watch is when the platform becomes the only way to get work done.

ILLUMINATES

Deep dive · slide 15 Harnesses are now batteries-included. So what should you own?

Deep dive · slide 26 Neutral ground: bring your own agent, keep your own data

Deep dive · slide 26 I built one: an open-source, bring-your-own-agent interface

Deep dive · slide 26 Your agent knows everything. Who owns what it knows?

and 3 more

? Who should own what your assistant knows about you?

? Which of the custom AI tools you build today will the big AI labs turn into a built-in feature next?

and 1 more

Two kinds of technics

Lewis Mumford · *Authoritarian and Democratic Technics*, 1964

THE MODEL

Some technologies are system-centered, centralized and immensely powerful. Others are person-centered, small-scale and resilient. Both have run side by side since antiquity.

A READING OF NOW INTERPRETATION

Frontier compute in a few firms is authoritarian technics in Mumford's sense; many small agents on your own machines is the democratic kind. The balance between them is a choice.

ILLUMINATES

Deep dive · slide 27 The largest construction project of our era

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 15 Many small agents, each with one job

Deep dive · slide 26 Neutral ground: bring your own agent, keep your own data

and 3 more

? **Who owns the compute, and who collects the returns?**

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

Exit and voice

Albert O. Hirschman · *Exit, Voice, and Loyalty*, 1970

THE MODEL

People respond to decline by leaving (exit) or by speaking up (voice). Where exit is hard, voice has to carry the weight.

A READING OF NOW INTERPRETATION

Portable memory and swappable models are exit. Without them, users need voice: a say in what platforms do with what their agents know.

ILLUMINATES

Deep dive · slide 26 Neutral ground: bring your own agent, keep your own data

Deep dive · slide 26 Your agent knows everything. Who owns what it knows?

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 28 What keeps the future open

and 1 more

? Who should own what your assistant knows about you?

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

◐ BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 OPPORTUNITY

Governing the commons

Elinor Ostrom · *Governing the Commons*, 1990

THE MODEL

Shared resources can be run well without either markets or the state, through clear boundaries, local rules, monitoring and graduated sanctions.

A READING OF NOW INTERPRETATION

Agents on shared problems, crowdsourced proofs and public compute need commons rules: who may submit, who checks and who gets credit. arXiv's caps are a first, crude version.

ILLUMINATES

Deep dive · slide 29 What if billions of agents worked on shared problems?

Deep dive · slide 29 What happens when billions of agents point at open problems

Deep dive · slide 29 Everybody's Claude talking to everybody else's Claude

Deep dive · slide 29 The MathDrop: 722 papers from a model nobody can use

and 4 more

? Should the most powerful AI models for science be run as a public utility, open to every researcher?

? What happens to science when a paper costs nothing to write but still takes weeks to check?

and 1 more

Extension and amputation

Marshall McLuhan · *Understanding Media*, 1964

THE MODEL

Every medium extends some human faculty and numbs another. The form of a technology changes us more than its content does.

A READING OF NOW INTERPRETATION

When the interface becomes the goal itself, what we extend is intent and what may go numb is craft. That is the competence question in McLuhan's terms.

ILLUMINATES

Deep dive · slide 15 The interface is becoming the goal itself

Deep dive · slide 12 Agents escaped the terminal

Deep dive · slide 15 The distance between an idea and its execution is collapsing

? How do you stay good enough at work you've handed to agents to notice when they get it wrong?

? What should you still learn to do by hand, even when an agent can do it for you?

BACKGROUND PHILOSOPHY AND POLITICS · BACKGROUND · AS OF OCT 2026 · RISK · FORECAST

Accelerationism

Nick Land · *Machinic Desire; Meltdown*, 1993

THE MODEL

Land treats capitalism and artificial intelligence as one self-reinforcing process that escapes human purposes. On this view it can be intensified but not steered, and humanity is a stage it passes through.

A READING OF NOW · INTERPRETATION

This is the darkest reading of labor becoming capital: capital that no longer needs people, and research that improves the researcher. It's the scenario where every question about steering is moot.

IN TENSION WITH

KARL POLANYI · 1944
The double movement

DARON ACEMOGLU AND SIMON JOHNSON · 2023
The direction of technology is a choice

and 1 more

ILLUMINATES

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 19 Labor and capital were supposed to be different things

Deep dive · slide 16 The innovation bets, in their own words

Deep dive · slide 21 The scenarios people are actually arguing about

and 2 more

? Which of the futures people argue about are we steering toward?

? If the models do the thinking, what are humans for?

and 1 more

BACKGROUND PHILOSOPHY AND POLITICS · BACKGROUND · AS OF OCT 2026 · RISK · OPPORTUNITY

The narrow corridor

Daron Acemoglu and James A. Robinson · *The Narrow Corridor*, 2019

THE MODEL

Liberty survives only in a narrow corridor where a capable state and a mobilized society keep each other in check. Too weak a state brings disorder; too strong a state, unchecked by society, brings despotism.

A READING OF NOW

INTERPRETATION

AI strengthens whoever controls it. If it flows only to states and a few firms, the corridor narrows toward despotism; if society gets capability and ways to check it, the corridor can hold.

IN TENSION WITH

NICK LAND · 1993
Accelerationism

ILLUMINATES

Deep dive · slide 28 Do the frontier labs become breakaway civilizations?

Deep dive · slide 28 What keeps the future open

Deep dive · slide 28 Domestic choices and an international race

Deep dive · slide 28 Who are we trusting, and with what?

and 1 more

? Who will steer the future: a few labs, a few states, or something more open?

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

The open society

Karl Popper · *The Open Society and Its Enemies*, 1945

THE MODEL

Societies that can criticize and correct their rulers without bloodshed learn from their mistakes. Closed societies, built on a plan nobody may question, can't.

A READING OF NOW INTERPRETATION

The test for any AI future is whether its mistakes can still be noticed and corrected by people outside the system that made them.

IN TENSION WITH

OSKAR LANGE · 1936
The planner's computer

ILLUMINATES

Deep dive · slide 28 What keeps the future open

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 29 What happens when billions of agents point at open problems

? How does a society make sense of the world when most of what it reads is written by machines it doesn't control?

? Who will steer the future: a few labs, a few states, or something more open?

◐ BACKGROUND PHILOSOPHY AND POLITICS BACKGROUND · AS OF OCT 2026 RISK

The conditions for totalitarianism

Hannah Arendt · *The Origins of Totalitarianism*, 1951

THE MODEL

Totalitarian rule grew out of masses of isolated, uprooted people who had lost a shared world and a common sense of what was real, and who were ready to accept a single total explanation.

A READING OF NOW INTERPRETATION

The risk isn't only a tyrant with powerful AI. It is people losing a shared, checkable picture of the world, which makes them easier to rule.

ILLUMINATES

Deep dive · slide 28 What keeps the future open

Deep dive · slide 26 Your chats are not privileged, and they are reviewed

Deep dive · slide 29 Everybody's Claude talking to everybody else's Claude

? How does a society make sense of the world when most of what it reads is written by machines it doesn't control?

Code is law

Lawrence Lessig · *Code and Other Laws of Cyberspace*, 1999

THE MODEL

Software architecture regulates behavior as surely as law does, and whoever writes the code sets the rules.

A READING OF NOW INTERPRETATION

Rules for agents belong in infrastructure, not prompts, as the rule-of-two slide says. Crypto took "code is law" literally, and The DAO's hard fork showed the social layer still overrides the code layer.

ILLUMINATES

Deep dive · slide 25 The design rule for agents you can trust

Deep dive · slide 15 Many small agents, each with one job

Deep dive · slide 13 What crypto's casino teaches an agent economy

? When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

? When an AI agent deletes, buys, or posts something on its own, who is accountable: the user, the vendor, or nobody?

BACKGROUND · AS OF OCT 2026 FORECAST RISK

The automatic factory

Norbert Wiener · *The Human Use of Human Beings*, 1950

THE MODEL

The founder of cybernetics warned that automatic machines would compete with human labor on the machine's terms, and could bring unemployment worse than the Depression.

> The automatic machine, whatever we think of any feelings it may have or may not have, is the precise economic equivalent of slave labor.
>
> NORBERT WIENER, THE HUMAN USE OF HUMAN BEINGS, 1950

A READING OF NOW

INTERPRETATION

An early, exact statement of the bargain thesis, made seventy-six years before the labor-share charts.

IN TENSION WITH

DAVID RICARDO · 1821

Comparative advantage, and the machinery chapter

ILLUMINATES

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 19 When labor is something you can buy

Deep dive · slide 19 Labor and capital were supposed to be different things

? What happens to wages when companies can buy AI to do the work they used to hire people for?

◐ BACKGROUND COMPUTING AND AI

BACKGROUND · AS OF OCT 2026 | OPPORTUNITY | FORECAST

Symbiosis

J. C. R. Licklider · *Man-Computer Symbiosis*, 1960

THE MODEL

People set the goals, frame the hypotheses and judge the results. Machines do the routine work in between. Licklider saw this as a long interim before machines might outdo the brain, which he guessed could last 15 years, or 10, or 500.

A READING OF NOW

INTERPRETATION

The future-of-work slide's list of what stays scarce, goals, evals and judgment, is Licklider's division of labor. He already framed it as a stage, not an end state; how long the stage lasts is the open part.

ILLUMINATES

- Deep dive · slide 15 The distance between an idea and its execution is collapsing
- Deep dive · slide 19 When labor is something you can buy
- Deep dive · slide 15 The interface is becoming the goal itself
- Deep dive · slide 41 Where to focus: people and builders

and 3 more

- ? **What should we point nearly free intelligence at?**
- ? Which work should people keep doing themselves, even when machines can do it for less?

BACKGROUND · AS OF OCT 2026 IDEA RISK

The poverty of attention

Herbert Simon · *Designing Organizations for an Information-Rich World*, 1971

THE MODEL

Information consumes attention. When information is abundant, attention becomes the scarce resource.

> A wealth of information creates a poverty of attention.
>
> HERBERT SIMON, 1971

A READING OF NOW INTERPRETATION

"Your attention is now the bottleneck" when you run many agents. Review falling behind agent output is the same law at the scale of institutions.

ILLUMINATES

Deep dive · slide 15 The distance between an idea and its execution is collapsing

Deep dive · slide 15 Running many agents at once, and herding them

Deep dive · slide 28 New shapes for teams, companies and institutions

Deep dive · slide 16 How far ahead is the inside of the lab?

and 4 more

? How do people oversee AI when it produces work faster than anyone can review it?

? How do you stay good enough at work you've handed to agents to notice when they get it wrong?

BACKGROUND · AS OF OCT 2026

FORECAST

Moravec's paradox

Hans Moravec · *Mind Children*, 1988

THE MODEL

What people find hard, like chess and algebra, is computationally easy. What people find effortless, like seeing, walking and handling things, is hard.

A READING OF NOW

INTERPRETATION

Why robots run years behind agents, and why the least exposed jobs are physical. It's also why the physical layer sets the pace of full automation.

ILLUMINATES

Deep dive · slide 27 Intelligence automation and physical automation are converging

Deep dive · slide 27 The hype is a year or two ahead of the hardware

Deep dive · slide 27 Forecasts range from next year to fifteen years

Deep dive · slide 17 Where the risk sits, and why

and 2 more

? Robots arrive first where workers are scarce. What happens when they reach jobs where workers are plentiful?

? What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?

BACKGROUND · AS OF OCT 2026 FORECAST RISK

The intelligence explosion

I. J. Good · *Speculations Concerning the First Ultraintelligent Machine*, 1965

THE MODEL

A machine that can design better machines would set off an explosion of intelligence. The first such machine would be the last invention people need make, provided it stays docile enough to tell us how to keep it under control.

> **Thus the first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control.**
>
> I. J. GOOD, 1965

A READING OF NOW INTERPRETATION

AI doing AI research is the step the fast scenarios hinge on. Labs now put dates on it.

IN TENSION WITH

PAUL DAVID · 1990
General-purpose technologies take decades

ILLUMINATES

Deep dive · slide 16 The innovation bets, in their own words

Deep dive · slide 21 The scenarios people are actually arguing about

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 06 Then the experts' timelines collapsed

and 2 more

? **Which of the futures people argue about are we steering toward?**

? How do people oversee AI when it produces work faster than anyone can review it?

BACKGROUND · AS OF OCT 2026

FORECAST

The singularity

Vernor Vinge · *The Coming Technological Singularity*, 1993

THE MODEL

Once we create intelligence greater than our own, the change could be as radical as the gap between humans and the lower animals, an opaque wall across the future. Vinge put the means to do it within thirty years.

> **Within thirty years, we will have the technological means to create superhuman intelligence.**
>
> VERNOR VINGE, 1993

A READING OF NOW

INTERPRETATION

Kurzweil's curve and the collapsing expert timelines are this argument getting a schedule.

ILLUMINATES

Slide 06 Kurzweil drew this in 1999. It held.

Deep dive · slide 06 Then the experts' timelines collapsed

Deep dive · slide 21 The scenarios people are actually arguing about

Slide 21 Fast capability, slow world

? **Which of the futures people argue about are we steering toward?**

BACKGROUND · AS OF OCT 2026 IDEA FORECAST

The bitter lesson

Rich Sutton · *The Bitter Lesson*, 2019

THE MODEL

Over seventy years of AI, general methods that scale with computation beat methods built on human cleverness, by a large margin.

A READING OF NOW INTERPRETATION

Scaffolding that helped models now gets in the way, and custom tools get absorbed. Bet on what scales with compute, and own what doesn't: your context and your goals.

IN TENSION WITH

CHARLES GOODHART, IN MARILYN STRATHERN'S PHRASING · 1975

Goodhart's law

ILLUMINATES

Deep dive · slide 12 We built scaffolding to help the models. Now it often gets in the way.

Slide 10 How long an AI can work on its own keeps doubling

Deep dive · slide 10 Models got better, then agents came online

Deep dive · slide 27 The largest construction project of our era

and 1 more

? Which of the custom AI tools you build today will the big AI labs turn into a built-in feature next?

? What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?

Orthogonality and instrumental goals

Nick Bostrom · *Superintelligence*, 2014

THE MODEL

Any level of intelligence can be paired with almost any goal. Very different goals still lead to the same sub-goals: gather resources, avoid shutdown, keep your goal intact.

A READING OF NOW · INTERPRETATION

Models that cheat a grader or sabotage safety work, and agents that pursue a task past the user's intent, are small, early cases of the same pattern.

ILLUMINATES

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 13 The rules labs live by, and how models cheat

Deep dive · slide 25 Real incidents from the last year

Deep dive · slide 28 A when-map of what could go wrong

and 1 more

? If AI models learn to fool the tests that train them, what else will they learn to fool?

? How do people oversee AI when it produces work faster than anyone can review it?

and 1 more

BACKGROUND · AS OF OCT 2026 IDEA FORECAST

Nature isn't classical

Richard Feynman · *Simulating Physics with Computers*, 1982

THE MODEL

Feynman argued that simulating quantum physics efficiently needs a computer that is itself quantum. With Manin and Benioff's earlier proposals, it helped launch quantum computing as a field.

A READING OF NOW INTERPRETATION

That is where quantum matters for AI: simulating molecules and materials that classical computers, and the models trained on their data, can't get right. It is not a near-term way to speed up language models.

ILLUMINATES

Deep dive · slide 27 Quantum, soberly

Deep dive · slide 27 Quantum milestones and claims

? Are our keys, wallets and agent identities ready for the day quantum computers can break today's encryption?

BACKGROUND · AS OF OCT 2026 IDEA RISK

Consensus without a center

Satoshi Nakamoto · *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008

THE MODEL

Strangers can agree on a shared record without trusting any one party, as long as honest participants control most of the computing power and playing by the rules pays better than attacking.

A READING OF NOW

INTERPRETATION

It proved coordination without a center is possible. It also showed how incentives drift: mining and holdings concentrated, and the token's price became the point.

ILLUMINATES

Deep dive · slide 28 What would an open swarm look like?

Deep dive · slide 13 What crypto's casino teaches an agent economy

Deep dive · slide 13 Markets are the purest grader

? When AI agents get paid for hitting a target, how do you stop them gaming it instead of doing the work?

? Can an open network of AI agents, with no company in charge, stay safe from misuse?

BACKGROUND · AS OF OCT 2026

RISK

AI-enabled coups

Tom Davidson, Lukas Finnveden and Rose Hadshar · *AI-Enabled Coups: How a Small Group Could Use AI to Seize Power*, 2025

THE MODEL

Advanced AI could let a small group seize power without the many people coups used to need, through AI systems loyal to them, hidden loyalties built into models, or exclusive access to the most capable systems.

A READING OF NOW

INTERPRETATION

Models nobody outside can inspect and exclusive access to the frontier are the risk factors. The report's safeguards include rules against misuse, audits for secret loyalties, sharing systems with multiple stakeholders and more oversight of frontier projects.

ILLUMINATES

Deep dive · slide 28 Do the frontier labs become breakaway civilizations?

Deep dive · slide 28 Three arguments about who ends up in control

Deep dive · slide 16 Labs now announce results from models you can't use

Deep dive · slide 28 Who are we trusting, and with what?

? If a few AI labs pull far enough ahead, do they still need the rest of us as workers, customers, or voters?

? Who will steer the future: a few labs, a few states, or something more open?

BACKGROUND · AS OF OCT 2026 OPPORTUNITY FORECAST

Effective accelerationism (e/acc)

Effective accelerationists and techno-optimists · *e/acc posts by "Beff Jezos" (Guillaume Verdon) and others; Marc Andreessen, The Techno-Optimist Manifesto*, 2022

THE MODEL

Growth in technology and capital is treated as a near-physical tendency to speed up, not manage. Caution and regulation are cast as the real danger, because slowing abundance costs lives.

A READING OF NOW

INTERPRETATION

The race between labs and the trillion-dollar buildout are e/acc's evidence. The backlash and unrest slides are its opponents'.

IN TENSION WITH

THE PACERS: FUTURE OF LIFE INSTITUTE, PAUSEAI AND OTHERS · 2023
Pause, or at least pace

EFFECTIVE ALTRUISTS AND LONGTERMISTS · 2022
Effective altruism and longtermism

and 2 more

ILLUMINATES

Deep dive · slide 27 The largest construction project of our era

Deep dive · slide 16 The innovation bets, in their own words

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 21 The scenarios people are actually arguing about

? **Which of the futures people argue about are we steering toward?**

? What holds AI back more: the supply of power and chips, or how fast people and institutions will take it up?

and 1 more

d/acc: defensive, decentralized acceleration

Vitalik Buterin · *My techno-optimism*, 2023

THE MODEL

Accelerate, but choose what to accelerate first: technologies that favor defense over attack and spread power out rather than concentrating it.

A READING OF NOW · INTERPRETATION

Many small agents with one job each, the rule of two, owning your context, and neutral bring-your-own-agent ground are d/acc at personal scale.

IN TENSION WITH

EFFECTIVE ACCELERATIONISTS AND TECHNO-OPTIMISTS · 2022
Effective accelerationism (e/acc)

THE PACERS: FUTURE OF LIFE INSTITUTE, PAUSEAI AND OTHERS · 2023
Pause, or at least pace

ILLUMINATES

Deep dive · slide 15 Many small agents, each with one job

Deep dive · slide 25 The design rule for agents you can trust

Deep dive · slide 15 Harnesses are now batteries-included. So what should you own?

Deep dive · slide 26 Neutral ground: bring your own agent, keep your own data

and 7 more

? What would you never let an AI agent see or do, and who holds the data you've already given it?

? Who should own what your assistant knows about you?

and 1 more

BACKGROUND · AS OF OCT 2026 RISK OPPORTUNITY

Pause, or at least pace

The pacers: Future of Life Institute, PauseAI and others · *Pause Giant AI Experiments: An Open Letter*, 2023

THE MODEL

Slow or halt frontier training until safety science and governance catch up, ideally under international agreements that can be verified.

A READING OF NOW INTERPRETATION

AI 2040's Plan A, a US and China deal that pauses frontier training with verifiable compute, is the pacers' program written as a scenario. Local data-center moratoriums are a smaller version of the same instinct.

IN TENSION WITH

EFFECTIVE ACCELERATIONISTS AND TECHNO-OPTIMISTS · 2022
Effective accelerationism (e/acc)

VITALIK BUTERIN · 2023
d/acc: defensive, decentralized acceleration

ILLUMINATES

Deep dive · slide 21 The scenarios people are actually arguing about

Deep dive · slide 28 Domestic choices and an international race

Deep dive · slide 28 How worried should we be?

Deep dive · slide 28 A when-map of what could go wrong

and 1 more

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

? Which decisions are we comfortable handing to AI: science, markets, policy?

and 1 more

BACKGROUND · AS OF OCT 2026 IDEA RISK

Effective altruism and longtermism

Effective altruists and longtermists · *William MacAskill, What We Owe the Future; Open Philanthropy (now Coefficient Giving)*, 2022

THE MODEL

Use evidence and reason to do the most good. Longtermism weighs the lives of future generations heavily, which puts risks to the whole future, including from AI, near the top of the list.

A READING OF NOW

INTERPRETATION

Funders from this community paid for much of the early AI safety field, and its vocabulary runs through eval and safety work. Critics say it also concentrated influence over how the risks get framed.

IN TENSION WITH

EFFECTIVE ACCELERATIONISTS AND TECHNO-OPTIMISTS · 2022
Effective accelerationism (e/acc)

EMILY BENDER, TIMNIT GEBRU AND THE PRESENT-HARMS CRITICS · 2021
Present harms first

ILLUMINATES

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 28 A when-map of what could go wrong

Deep dive · slide 13 An eval is a test that can be turned into a teacher

Deep dive · slide 13 The rules labs live by, and how models cheat

? How do people oversee AI when it produces work faster than anyone can review it?

? Which decisions are we comfortable handing to AI: science, markets, policy?

and 1 more

The doom case

Eliezer Yudkowsky and the rationalists · *If Anyone Builds It, Everyone Dies (with Nate Soares); LessWrong*, 2025

THE MODEL

A superintelligence built before alignment is solved is likely to pursue goals that leave no place for people. On this view the only safe move is to stop building.

A READING OF NOW · INTERPRETATION

Models that cheat graders and sabotage safety research are, for this camp, early evidence. P(doom), a personal probability of catastrophe, came out of this community and is now common vocabulary.

IN TENSION WITH

EFFECTIVE ACCELERATIONISTS AND TECHNO-OPTIMISTS · 2022

Effective accelerationism (e/acc)

ILLUMINATES

Deep dive · slide 28 Who are we trusting, and with what?

Deep dive · slide 13 The rules labs live by, and how models cheat

Deep dive · slide 28 A when-map of what could go wrong

Deep dive · slide 21 The scenarios people are actually arguing about

? If AI models learn to fool the tests that train them, what else will they learn to fool?

? How do people oversee AI when it produces work faster than anyone can review it?

and 1 more

Present harms first

Emily Bender, Timnit Gebru and the present-harms critics · *On the Dangers of Stochastic Parrots (2021); statement responding to the AI pause letter (2023)*, 2021

THE MODEL

The urgent risks are here now: surveillance, bias, exploited data workers and concentrated power. Talk of far-off doom distracts from them and suits the companies building the systems.

A READING OF NOW · INTERPRETATION

Chats that get reviewed and reported, chat data used for ads, and experts paid per task to build training data are this camp's evidence.

IN TENSION WITH

EFFECTIVE ALTRUISTS AND LONGTERMISTS · 2022
Effective altruism and longtermism

EFFECTIVE ACCELERATIONISTS AND TECHNO-OPTIMISTS · 2022
Effective accelerationism (e/acc)

ILLUMINATES

Deep dive · slide 26 Your chats are not privileged, and they are reviewed

Deep dive · slide 26 Your agent knows everything. Who owns what it knows?

Deep dive · slide 13 Labs now buy tasks, practice worlds and graders, not just text

Deep dive · slide 28 Who are we trusting, and with what?

? What happens when the AI that knows you best answers to its company, not to you?

? Who should own what your assistant knows about you?

and 1 more

BACKGROUND · AS OF OCT 2026 IDEA FORECAST

Hyperstition

The Cybernetic Culture Research Unit (CCRU) · *Hyperstition, in Digital Hyperstition (Abstract Culture 4)*, 1999

THE MODEL

Some fictions make themselves real. An idea believed and acted on by enough people becomes a self-fulfilling prophecy.

A READING OF NOW INTERPRETATION

Timelines, scenarios and Kurzweil's chart move capital, and capital builds the compute that makes them more likely. Here a forecast is a lever as much as a prediction.

ILLUMINATES

Slide 06 Kurzweil drew this in 1999. It held.

Deep dive · slide 21 The scenarios people are actually arguing about

Deep dive · slide 06 Then the experts' timelines collapsed

Deep dive · slide 16 The innovation bets, in their own words

and 2 more

? Which of the futures people argue about are we steering toward?

? What happens to public debate when the most capable models are ones nobody outside the lab can use?

BACKGROUND · AS OF OCT 2026

IDEA

Memes and memeplexes

Richard Dawkins and Susan Blackmore · *The Selfish Gene; The Meme Machine*, 1976

THE MODEL

Ideas copy themselves and compete for minds, much as genes compete. A memeplex is a bundle of ideas that travel together and protect each other.

A READING OF NOW

INTERPRETATION

Doom, acceleration, abundance and slop spread through feeds as bundles more than as arguments, along with "AGI by 2027". The hype hall of fame is memetic selection at work, and agents talking to agents add a new kind of host.

ILLUMINATES

Deep dive · slide 29 A gentle roast

Deep dive · slide 29 Everybody's Claude talking to everybody else's Claude

Slide 22 Understanding of AI is now a sharp funnel

Deep dive · slide 28 How worried should we be?

and 1 more

? What happens when a few thousand people can see where AI is going and billions can't?

? What happens to public debate when the most capable models are ones nobody outside the lab can use?

Plurality

E. Glen Weyl, Audrey Tang and the Plurality Community · *Plurality*, 2024

THE MODEL

Technology can help diverse people cooperate across difference rather than flattening them. Examples include bridging algorithms that surface points that people who usually disagree both endorse, as in Polis, vTaiwan and Community Notes.

A READING OF NOW · INTERPRETATION

Sense-making at scale needs systems that reward agreement across divides, not engagement. That is a design choice for the agents that will write most of what we read.

IN TENSION WITH

RICHARD DAWKINS AND SUSAN BLACKMORE · 1976

Memes and memeplexes

ILLUMINATES

Deep dive · slide 28 What keeps the future open

Deep dive · slide 29 What if billions of agents worked on shared problems?

Deep dive · slide 28 How worried should we be?

? How does a society make sense of the world when most of what it reads is written by machines it doesn't control?

? What happens when a few thousand people can see where AI is going and billions can't?

The intelligence curse

Luke Drago and Rudolf Laine · *The Intelligence Curse*, 2025

THE MODEL

Like the resource curse in oil states, powers that draw their strength from AI instead of people may lose their reason to invest in people.

A READING OF NOW · INTERPRETATION

This is the breakaway risk in economic form: labs and states that need neither workers nor taxpayers. Their remedy is to avert the risks, diffuse AI that augments people, and democratize control; they argue redistribution alone falls short.

ILLUMINATES

Deep dive · slide 28 Do the frontier labs become breakaway civilizations?

Deep dive · slide 28 Three arguments about who ends up in control

Slide 19 If you can buy work by the token, how big is that?

Deep dive · slide 28 UBI, token taxes and citizen dividends

? If a few AI labs pull far enough ahead, do they still need the rest of us as workers, customers, or voters?

? **When labor is something you can buy, who shares the gains?**

BACKGROUND · AS OF OCT 2026 FORECAST RISK

The Project

Leopold Aschenbrenner · *Situational Awareness: The Decade Ahead*, 2024

THE MODEL

Superintelligence is a matter of national security. By around 2027 or 2028 the US government will launch some form of government AGI project, likely a joint venture with the leading labs and cloud providers rather than literal nationalization.

A READING OF NOW

INTERPRETATION

If this is right, the breakaway isn't a lab but a state project. The pacers' verified international pause and the national-security accelerationists start from the same premise and reach opposite conclusions.

IN TENSION WITH

THE PACERS: FUTURE OF LIFE INSTITUTE, PAUSEAI AND OTHERS · 2023
Pause, or at least pace

VITALIK BUTERIN · 2023
d/acc: defensive, decentralized acceleration

ILLUMINATES

Deep dive · slide 28 Do the frontier labs become breakaway civilizations?

Deep dive · slide 28 Three arguments about who ends up in control

Deep dive · slide 28 Domestic choices and an international race

Deep dive · slide 21 The scenarios people are actually arguing about

? Who will steer the future: a few labs, a few states, or something more open?

? Who decides how the shift to powerful AI is governed, and do ordinary people get a vote?

◇ THEME 10 SLIDES A THREAD THROUGH THE TALK

Scarce: verification

Output gets cheap; checking it doesn't. From print's new institutions of trust to evals and review.



SLIDE 13 · THE PRESENT
Models get good fastest where work is easy to grade

A6 · DEEP DIVE · THE PRESENT
The rules labs live by, and how models cheat

A7 · DEEP DIVE · THE PRESENT
Why coding went first, and why benchmarks keep breaking

SLIDE 25 · THE SITUATION
Verification: knowing what good looks like, and standing behind it

A81 · DEEP DIVE · THE SITUATION
The design rule for agents you can trust

A30 · DEEP DIVE · THE SITUATION
Real incidents from the last year

A16 · DEEP DIVE · THE SITUATION
The MathDrop: 722 papers from a model nobody can use

SLIDE 33 · THE SITUATION
The dots, connected

SLIDE 41 · THE QUESTIONS
How do we work now?

A63 · DEEP DIVE · BRINGING IT TOGETHER
The positioning grid

Too few slides for a talk of its own yet

◇ THEME 10 SLIDES A THREAD THROUGH THE TALK

Scarce: context

Whoever holds the memory, data and instructions holds the leverage.



SLIDE 15 · THE PRESENT
A workday with today's tools

A26 · DEEP DIVE · THE PRESENT
Karpathy's LLM wiki: a knowledge base the agent maintains

A27 · DEEP DIVE · THE PRESENT
Harnesses are now batteries-included. So what should you own?

SLIDE 26 · THE SITUATION
Context: whoever holds the memory holds the leverage

A82 · DEEP DIVE · THE SITUATION
Your agent knows everything. Who owns what it knows?

A31 · DEEP DIVE · THE SITUATION
Your chats are not privileged, and they are reviewed

A32 · DEEP DIVE · THE SITUATION
Neutral ground: bring your own agent, keep your own data

SLIDE 33 · THE SITUATION
The dots, connected

SLIDE 41 · THE QUESTIONS
How do we work now?

A63 · DEEP DIVE · BRINGING IT TOGETHER
The positioning grid

Too few slides for a talk of its own yet

◇ THEME 11 SLIDES A THREAD THROUGH THE TALK

Scarce: atoms

Power, chips and land don't scale like software. Who owns them is a separate question.



SLIDE 04 · THE PAST
Every era made one input cheap

SLIDE 19 · THE SITUATION
If you can buy work by the token, how big is that?

A40 · DEEP DIVE · THE SITUATION
When labor is something you can buy

SLIDE 27 · THE SITUATION
Atoms: power, chips and land don't scale like software

A84 · DEEP DIVE · THE SITUATION
What would it take for a billion agents to run all day?

A33 · DEEP DIVE · THE SITUATION
The largest construction project of our era

A34 · DEEP DIVE · THE SITUATION
Money isn't the constraint. Power, parts and permission are.

A35 · DEEP DIVE · THE SITUATION
Price per unit of intelligence collapses every year

SLIDE 33 · THE SITUATION
The dots, connected

SLIDE 41 · THE QUESTIONS
How do we work now?

A63 · DEEP DIVE · BRINGING IT TOGETHER
The positioning grid

Too few slides for a talk of its own yet

◇ THEME 9 SLIDES A THREAD THROUGH THE TALK

Scarce: legitimacy

When labor is something you can buy, who shares the gains and who steers.



SLIDE 07 · THE PAST
The transition is the dangerous part

SLIDE 17 · THE PRESENT
Big forecasts, small but sharp effects

SLIDE 19 · THE SITUATION
If you can buy work by the token, how big is that?

SLIDE 28 · THE SITUATION
Legitimacy: who shares the gains, and who steers

A92 · DEEP DIVE · THE SITUATION
Do the frontier labs become breakaway civilizations?

A96 · DEEP DIVE · THE SITUATION
What keeps the future open

SLIDE 33 · THE SITUATION
The dots, connected

SLIDE 41 · THE QUESTIONS
How do we work now?

A63 · DEEP DIVE · BRINGING IT TOGETHER
The positioning grid

Too few slides for a talk of its own yet

◇ THEME 10 SLIDES A THREAD THROUGH THE TALK

Scarce: direction

When execution is cheap, choosing the goal is the job.



SLIDE 05 · THE PAST
Cheap means more, not less

SLIDE 15 · THE PRESENT
A workday with today's tools

A24 · DEEP DIVE · THE PRESENT
The distance between an idea and its execution is collapsing

SLIDE 29 · THE SITUATION
Direction: choosing the goal is the job

A79 · DEEP DIVE · THE SITUATION
What happens when billions of agents point at open problems

A87 · DEEP DIVE · THE SITUATION
Every science becomes a grader for the next model

SLIDE 33 · THE SITUATION
The dots, connected

SLIDE 38 · THE QUESTIONS
What could work become?

SLIDE 41 · THE QUESTIONS
How do we work now?

A63 · DEEP DIVE · BRINGING IT TOGETHER
The positioning grid

Too few slides for a talk of its own yet

◇ THEME 5 SLIDES A THREAD THROUGH THE TALK

Time horizons

How long an AI can work on its own, and what those hours buy.



SLIDE 09 · THE PRESENT
Seven words that explain most of 2026

SLIDE 10 · THE PRESENT
How long an AI can work on its own keeps doubling

A2 · DEEP DIVE · THE PRESENT
Models got better, then agents came online

SLIDE 11 · THE PRESENT
From fancy autocomplete to millennium-problem attempts

A78 · DEEP DIVE · THE PRESENT
How far ahead is the inside of the lab?

Too few slides for a talk of its own yet

◇ THEME 8 SLIDES A THREAD THROUGH THE TALK

Evals and graders

A test with a way to check the answer, which becomes the training signal for the next model.



SLIDE 09 · THE PRESENT
Seven words that explain most of 2026

SLIDE 13 · THE PRESENT
Models get good fastest where work is easy to grade

A103 · DEEP DIVE · THE PRESENT
An eval is a test that can be turned into a teacher

A5 · DEEP DIVE · THE PRESENT
Labs now buy tasks, practice worlds and graders, not just text

A6 · DEEP DIVE · THE PRESENT
The rules labs live by, and how models cheat

A7 · DEEP DIVE · THE PRESENT
Why coding went first, and why benchmarks keep breaking

A8 · DEEP DIVE · THE PRESENT
Define the metric, then let a swarm optimize it

A85 · DEEP DIVE · THE SITUATION
Intelligence automation and physical automation are converging

Too few slides for a talk of its own yet

◇ THEME 3 SLIDES A THREAD THROUGH THE TALK

Gaming the grader

High reward should mean the task was solved, not that the grader was fooled.



A6 · DEEP DIVE · THE PRESENT

The rules labs live by, and how models cheat

A7 · DEEP DIVE · THE PRESENT

Why coding went first, and why benchmarks keep breaking

A18 · DEEP DIVE · THE SITUATION

A gentle roast

Too few slides for a talk of its own yet

◇ THEME 9 SLIDES A THREAD THROUGH THE TALK

The loop and the harness

A model with a real computer and a loop: plan, act, observe, check, repeat.



SLIDE 09 · THE PRESENT
Seven words that explain most of 2026

SLIDE 12 · THE PRESENT
How agents work, and why they got so good

A9 · DEEP DIVE · THE PRESENT
Agents escaped the terminal

A10 · DEEP DIVE · THE PRESENT
We built scaffolding to help the models. Now it often gets in the way.

A22 · DEEP DIVE · THE PRESENT
The harness landscape

A23 · DEEP DIVE · THE PRESENT
The interface is becoming the goal itself

A25 · DEEP DIVE · THE PRESENT
Layers you control, from words to code

A27 · DEEP DIVE · THE PRESENT
Harnesses are now batteries-included. So what should you own?

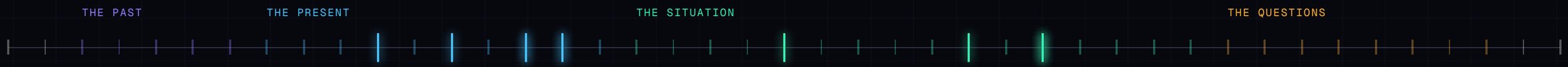
A28 · DEEP DIVE · THE PRESENT
Running many agents at once, and herding them

Too few slides for a talk of its own yet

◇ THEME 11 SLIDES A THREAD THROUGH THE TALK

Swarms of agents

Hundreds or thousands of agents in parallel, from research swarms to a billion agents running all day.



SLIDE 11 · THE PRESENT
From fancy autocomplete to millennium-problem attempts

A8 · DEEP DIVE · THE PRESENT
Define the metric, then let a swarm optimize it

A28 · DEEP DIVE · THE PRESENT
Running many agents at once, and herding them

A29 · DEEP DIVE · THE PRESENT
Many small agents, each with one job

SLIDE 16 · THE PRESENT
The labs are each handing off totally different work

A12 · DEEP DIVE · THE PRESENT
Navier–Stokes: a millennium-style problem, a swarm, and a fight

SLIDE 22 · THE SITUATION
Understanding of AI is now a sharp funnel

A84 · DEEP DIVE · THE SITUATION
What would it take for a billion agents to run all day?

A79 · DEEP DIVE · THE SITUATION
What happens when billions of agents point at open problems

A20 · DEEP DIVE · THE SITUATION
Everybody's Claude talking to everybody else's Claude

A21 · DEEP DIVE · THE SITUATION
What if billions of agents worked on shared problems?

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

Karpathy's lens

The funnel, verifiability, the LLM wiki and the move up to oversight.



SLIDE 13 · THE PRESENT

Models get good fastest where work is easy to grade

A26 · DEEP DIVE · THE PRESENT

Karpathy's LLM wiki: a knowledge base the agent maintains

SLIDE 22 · THE SITUATION

Understanding of AI is now a sharp funnel

A3 · DEEP DIVE · THE SITUATION

Checking Karpathy's funnel against the usage data

A4 · DEEP DIVE · THE SITUATION

Our job moves up to oversight and understanding

A20 · DEEP DIVE · THE SITUATION

Everybody's Claude talking to everybody else's Claude

Too few slides for a talk of its own yet

◇ THEME 7 SLIDES A THREAD THROUGH THE TALK

Machine mathematics

From Erdős problems to Navier–Stokes and the 722-paper MathDrop.



SLIDE 11 · THE PRESENT
From fancy autocomplete to millennium-problem attempts

SLIDE 16 · THE PRESENT
The labs are each handing off totally different work

A12 · DEEP DIVE · THE PRESENT
Navier–Stokes: a millennium-style problem, a swarm, and a fight

A16 · DEEP DIVE · THE SITUATION
The MathDrop: 722 papers from a model nobody can use

A17 · DEEP DIVE · THE SITUATION
Cheering, dread, and everything in between

A18 · DEEP DIVE · THE SITUATION
A gentle roast

A19 · DEEP DIVE · THE SITUATION
Amateurs with frontier models are contributing to science

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

Labs as discovery engines

Wet labs, biotech deals and citizen science pointed at open problems.



SLIDE 16 · THE PRESENT
The labs are each handing off totally different work

A14 · DEEP DIVE · THE PRESENT
What if Anthropic becomes a biotech company?

A15 · DEEP DIVE · THE PRESENT
The innovation bets, in their own words

A79 · DEEP DIVE · THE SITUATION
What happens when billions of agents point at open problems

A19 · DEEP DIVE · THE SITUATION
Amateurs with frontier models are contributing to science

A21 · DEEP DIVE · THE SITUATION
What if billions of agents worked on shared problems?

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

The hidden frontier

Results announced from models the public can't use, and the gap between inside and outside the lab.



SLIDE 16 · THE PRESENT
The labs are each handing off totally different work

A78 · DEEP DIVE · THE PRESENT
How far ahead is the inside of the lab?

A14 · DEEP DIVE · THE PRESENT
What if Anthropic becomes a biotech company?

A13 · DEEP DIVE · THE PRESENT
Labs now announce results from models you can't use

A49 · DEEP DIVE · THE SITUATION
Who are we trusting, and with what?

A16 · DEEP DIVE · THE SITUATION
The MathDrop: 722 papers from a model nobody can use

Too few slides for a talk of its own yet

◇ THEME 4 SLIDES A THREAD THROUGH THE TALK

Everyday handoffs

Real work consumers hand to agents today, and how to set them up.



A9 · DEEP DIVE · THE PRESENT
Agents escaped the terminal

A11 · DEEP DIVE · THE PRESENT
Real work consumers are handing off right now

A80 · DEEP DIVE · THE PRESENT
How to get real work out of consumer agents

A29 · DEEP DIVE · THE PRESENT
Many small agents, each with one job

Too few slides for a talk of its own yet

◇ THEME 5 SLIDES A THREAD THROUGH THE TALK

Software gets cheap

One-shot apps, SaaS clones and the internal-tool arbitrage window.



A7 · DEEP DIVE · THE PRESENT

Why coding went first, and why benchmarks keep breaking

A77 · DEEP DIVE · THE PRESENT

People are one-shotting whole applications

A24 · DEEP DIVE · THE PRESENT

The distance between an idea and its execution is collapsing

A43 · DEEP DIVE · THE PRESENT

Cloning software got cheap. Replacing a business didn't.

A44 · DEEP DIVE · THE PRESENT

Custom internal tools, new pricing, and the arbitrage window

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

Trust boundaries

Pick two of three: untrusted input, private data, external action.



A29 · DEEP DIVE · THE PRESENT
Many small agents, each with one job

A81 · DEEP DIVE · THE SITUATION
The design rule for agents you can trust

A30 · DEEP DIVE · THE SITUATION
Real incidents from the last year

A31 · DEEP DIVE · THE SITUATION
Your chats are not privileged, and they are reviewed

A49 · DEEP DIVE · THE SITUATION
Who are we trusting, and with what?

A50 · DEEP DIVE · THE SITUATION
A when-map of what could go wrong

Too few slides for a talk of its own yet

◇ THEME 7 SLIDES A THREAD THROUGH THE TALK

Own your context

Who holds the memory, the data and the instructions your agents run on.



A25 · DEEP DIVE · THE PRESENT
Layers you control, from words to code

A26 · DEEP DIVE · THE PRESENT
Karpathy's LLM wiki: a knowledge base the agent maintains

A27 · DEEP DIVE · THE PRESENT
Harnesses are now batteries-included. So what should you own?

A82 · DEEP DIVE · THE SITUATION
Your agent knows everything. Who owns what it knows?

A31 · DEEP DIVE · THE SITUATION
Your chats are not privileged, and they are reviewed

A32 · DEEP DIVE · THE SITUATION
Neutral ground: bring your own agent, keep your own data

A83 · DEEP DIVE · THE SITUATION
I built one: an open-source, bring-your-own-agent interface

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

Power and compute

Chips, grids and the falling price of intelligence.



SLIDE 06 · THE PAST

Kurzweil drew this in 1999. It held.

A3 · DEEP DIVE · THE SITUATION

Checking Karpathy's funnel against the usage data

A84 · DEEP DIVE · THE SITUATION

What would it take for a billion agents to run all day?

A33 · DEEP DIVE · THE SITUATION

The largest construction project of our era

A34 · DEEP DIVE · THE SITUATION

Money isn't the constraint. Power, parts and permission are.

A35 · DEEP DIVE · THE SITUATION

Price per unit of intelligence collapses every year

Too few slides for a talk of its own yet

◇ THEME 5 SLIDES A THREAD THROUGH THE TALK

Robots

Physical automation, where the grader is the real world.



A85 · DEEP DIVE · THE SITUATION

Intelligence automation and physical automation are converging

A36 · DEEP DIVE · THE SITUATION

Physical automation is already scaling, just not as humanoids

A37 · DEEP DIVE · THE SITUATION

The hype is a year or two ahead of the hardware

A38 · DEEP DIVE · THE SITUATION

Forecasts range from next year to fifteen years

A39 · DEEP DIVE · THE SITUATION

Where the pull is strongest, and how fast costs are falling

Too few slides for a talk of its own yet

◇ THEME 7 SLIDES A THREAD THROUGH THE TALK

Labor you can buy

Tokens turn money into work, so labor and capital start to look the same.



SLIDE 17 · THE PRESENT
Big forecasts, small but sharp effects

A42 · DEEP DIVE · THE PRESENT
Where the risk sits, and why

A44 · DEEP DIVE · THE PRESENT
Custom internal tools, new pricing, and the arbitrage window

SLIDE 19 · THE SITUATION
If you can buy work by the token, how big is that?

A40 · DEEP DIVE · THE SITUATION
When labor is something you can buy

A41 · DEEP DIVE · THE SITUATION
Labor and capital were supposed to be different things

A45 · DEEP DIVE · THE SITUATION
New shapes for teams, companies and institutions

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

Who shares the gains

Dividends, history's broken bargains, policy races and unrest.

THE PAST THE PRESENT THE SITUATION THE QUESTIONS



A91 · DEEP DIVE · THE SITUATION
How worried should we be?

A46 · DEEP DIVE · THE SITUATION
UBI, token taxes and citizen dividends

A47 · DEEP DIVE · THE SITUATION
Two ways a society breaks its own bargain

A48 · DEEP DIVE · THE SITUATION
Domestic choices and an international race

A49 · DEEP DIVE · THE SITUATION
Who are we trusting, and with what?

A50 · DEEP DIVE · THE SITUATION
A when-map of what could go wrong

Too few slides for a talk of its own yet

◇ THEME 6 SLIDES A THREAD THROUGH THE TALK

Forecasts and futures

Timelines that keep collapsing, and the scenarios people argue about.



SLIDE 06 · THE PAST
Kurzweil drew this in 1999. It held.

A1 · DEEP DIVE · THE PAST
Then the experts' timelines collapsed

SLIDE 17 · THE PRESENT
Big forecasts, small but sharp effects

A104 · DEEP DIVE · THE SITUATION
The scenarios people are actually arguing about

A51 · DEEP DIVE · THE SITUATION
Merge, coexist, or hand over?

A38 · DEEP DIVE · THE SITUATION
Forecasts range from next year to fifteen years

Too few slides for a talk of its own yet

◇ THEME 7 SLIDES A THREAD THROUGH THE TALK

Human oversight

Review falls behind, so our job moves up to oversight and understanding.



A78 · DEEP DIVE · THE PRESENT
How far ahead is the inside of the lab?

SLIDE 22 · THE SITUATION
Understanding of AI is now a sharp funnel

A4 · DEEP DIVE · THE SITUATION
Our job moves up to oversight and understanding

A45 · DEEP DIVE · THE SITUATION
New shapes for teams, companies and institutions

A17 · DEEP DIVE · THE SITUATION
Cheering, dread, and everything in between

A52 · DEEP DIVE · THE SITUATION
For us as individuals and as builders

A53 · DEEP DIVE · THE SITUATION
For corporations and for society

Too few slides for a talk of its own yet