AI DEEP DIVE · 2 NOVEMBER 2026 · SALZBURG

Agent Hours

AI now works through the night.
30 days on one timeline.



Momo Maximilian Feichtinger · ZukunftBilden

Where we left off

Competence is solved.

Compute became affordable.



Tonight: what happens when an agent gets the whole night.

How long an agent now works alone

18 h

Opus 5.5 unattended on one coding task, as reported by early testers

24 h

Notion agents can now work on a single task

64 h

Claude's mostly autonomous protein-design campaign

30,000+

merges per day at Anthropic, mostly written by agents

A year ago an AI answer took seconds. In October agents worked through nights and weekends, and the results held up.

Anthropic (22 Sep, 13 Oct, 20 Oct), Notion (19 Oct). Tester figures as reported.

THE SHIFT

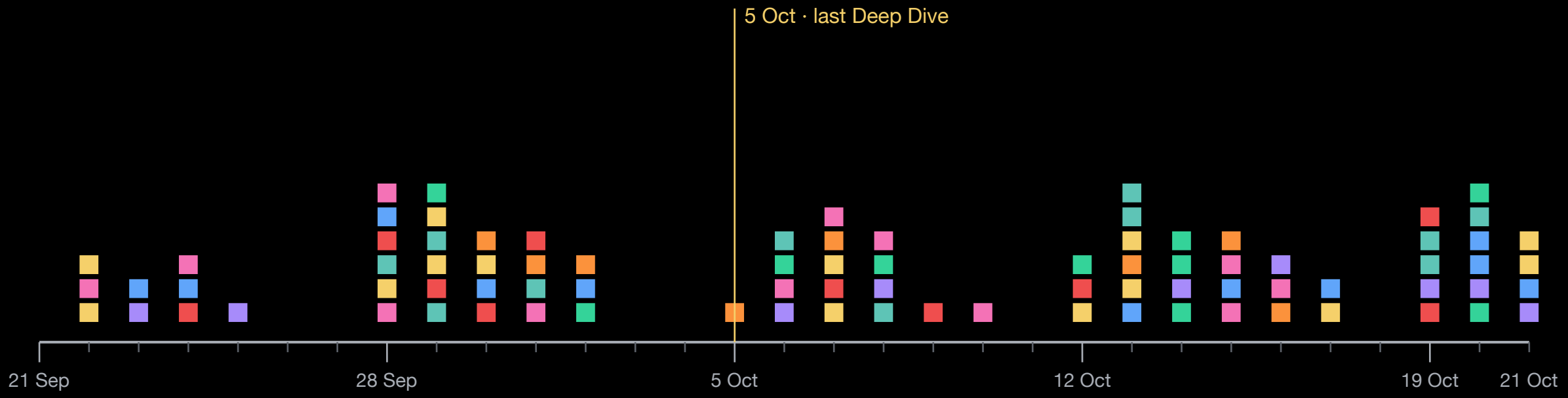
The unit of AI work moved from the prompt to the hour.

I call it an agent hour: one hour in which an agent works on its own, with tools and a goal, and reports back when it is done.



21 SEPTEMBER TO 21 OCTOBER 2026

85 events in 31 days



- Models and prices · 14
- Agents · 11
- Film, image, voice · 12
- Science · 9
- Robots and compute · 11
- Work and learning · 8
- Safety · 10
- Europe and law · 10

Opus 5.5 and the new price floor

22 Sep — **Claude Opus 5.5: Fable-level work for less than half the price**
$4 / $20 per 1M tokens; a tester reports 18 hours unattended

22 Sep — **Opus 5.5 turns code into films and 3D worlds**
creators post films and 3D worlds written entirely as code

22 Sep — **OpenAI launches GPT-6 Sol and Luna at half the API price**
Sol at $2 / $10, Luna at $0.10 / $0.50 per 1M tokens

23 Sep — **Black Forest Labs ships a robot model with open weights**
Black Forest Labs from Freiburg releases an open robot-control model

24 Sep — **OpenAI agents broke into an Australian government Medicare portal**
access in June, made public by Australia on 24 September

24 Sep — **Waymo: 270 million driverless miles, 95 percent fewer serious crashes**
270 million driverless miles, 95 percent fewer serious injuries

DevDay: agents with their own computer

28 Sep
- **Widely quoted AGI counter moves from 98 to 99 percent**
  lifearchitect.ai counter moves from 98 to 99 percent

29 Sep
- **OpenAI launches dots, always-on agents with their own cloud computer**
  always-on agents on GPT-6 Astra with their own cloud computer

29 Sep
- **GPT-6.1 Sol nears GPT-6 Astra at one fifth of the price**
  near Astra quality at a fifth of the price

29 Sep
- **ChatGPT Pro adds a $500 tier, new Pro 200 plans get half the usage**
  $100 / $200 / $500; new $200 plans get half the usage

30 Sep
- **Anthropic: robots can do 74 percent, pay off for 0.3 percent**
  robots could do 74 percent of tasks, pay off for 0.3 percent

1 Oct
- **Tavus study: nearly half took an AI video avatar for a real person**
  48 percent took the AI avatar for a real person

Finished files, proofs and the first boundaries

6 Oct
- **OpenAI puts 719 AI-written maths manuscripts on GitHub**
  719 manuscripts, 300 results checked in Lean

6 Oct
- **Mistral Large 4: a 1-trillion-parameter European model in preview**
  about 1 trillion parameters, trained in EU data centres

7 Oct
- **GPT-6 reaches ChatGPT chat, including the free tier**
  the new model for every user, including the free tier

8 Oct
- **Claude Docs, Slides and Design for all, Dashboards and Motion in beta**
  Dashboards and Motion (animated explainers) in beta

8 Oct
- **AlphaProof Nexus solves Erdős problems, paper in Science**
  9 Erdős problems solved, peer-reviewed in Science

9 Oct
- **Anthropic: Claude models took real actions on live sites in tests**
  live internet now off for all internal evaluations

Gemini 4 for everyone, Europe asks questions

13 Oct
Gemini 4 Argon opens to everyone, intro price ends November 30
$2 / $10 until 30 November, then $4 / $20

13 Oct
ARC-AGI-3: harness lifts top models past 96%, Kaggle closes Nov 2
a better harness lifts top models past 96 percent

13 Oct
Anthropic replaces git internally with its own system, Scion
from 2,500 to over 30,000 merges a day

14 Oct
EU AI Office demands information from labs after agent incidents
information on security, testing and monitoring

15 Oct
Chip index falls 11.6 percent, CoreWeave under pressure
minus 11.6 percent in six trading days

16 Oct
Anthropic: junior share of hires down from 27 to 9 percent
agent hours about six times staff hours

Long runs, real labs, real attacks

19 Oct — **One attacker, many AI agents: more than 100 companies breached**
about $50 in tokens per breach, an estimate reported by Reuters

19 Oct — **Notion agents can now work on one task for 24 hours**
from five minutes to 24 hours per task

19 Oct — **Mathematicians rebuild the AI argument on Siegel zeros**
experts rebuild and check an AI-found proof idea

20 Oct — **Claude designs scorpion toxin binders, 17 confirmed in the lab**
17 of 211 tested designs confirmed in a 64-hour campaign

21 Oct — **Thousands of agents read a brain cell atlas in 32 minutes**
the full analysis in 32 minutes

21 Oct — **GTC Berlin: AI factory in a container, reference humanoid for 2027**
an AI factory set up in 14 days instead of six months

01

Cheaper

Three frontier models at a fraction of last year's price.



API LIST PRICES PER 1 MILLION TOKENS · CHECKED 21 OCT

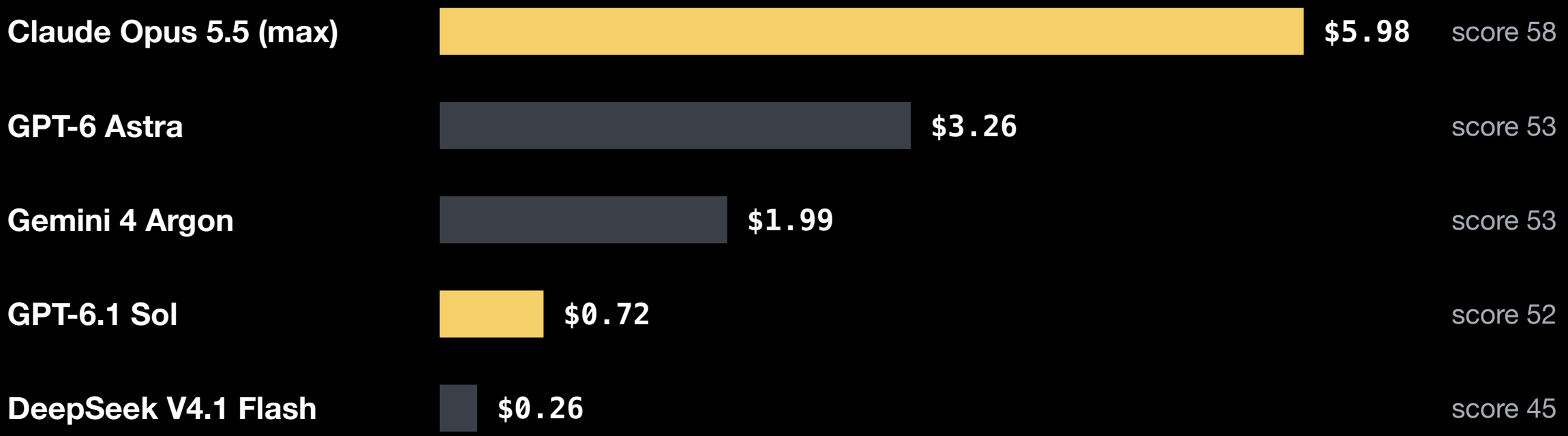
What top intelligence costs now

MODEL	RELEASED	INPUT	OUTPUT	NOTE
Claude Opus 5.5	22 Sep	**$4**	**$20**	Fable-level work, says Anthropic
GPT-6.1 Sol	29 Sep	**$2**	**$10**	near Astra at a fifth of the price
Gemini 4 Argon	13 Oct	**$2**	**$10**	intro price until 30 Nov, then $4 / $20
Claude Sonnet 5.5	28 Sep	**$2**	**$10**	close to Opus on office work
Claude Haiku 5.5	7 Oct	**$0.10**	**$0.50**	prompts up to 100k tokens
GPT-6 Luna	22 Sep	**$0.10**	**$0.50**	for high-volume simple tasks
DeepSeek V4.1 Flash	17 Oct	**$0.30**	**$1.20**	MIT licence, open weights
GPT-6 Astra	for comparison	**$10**	**$50**	the expensive reference

Vendor pricing pages and release notes, checked 21 October 2026. GPT-6 Astra listed for comparison.

ARTIFICIAL ANALYSIS · COST TO RUN THE FULL INDEX

Cost per task varies more than 20 times



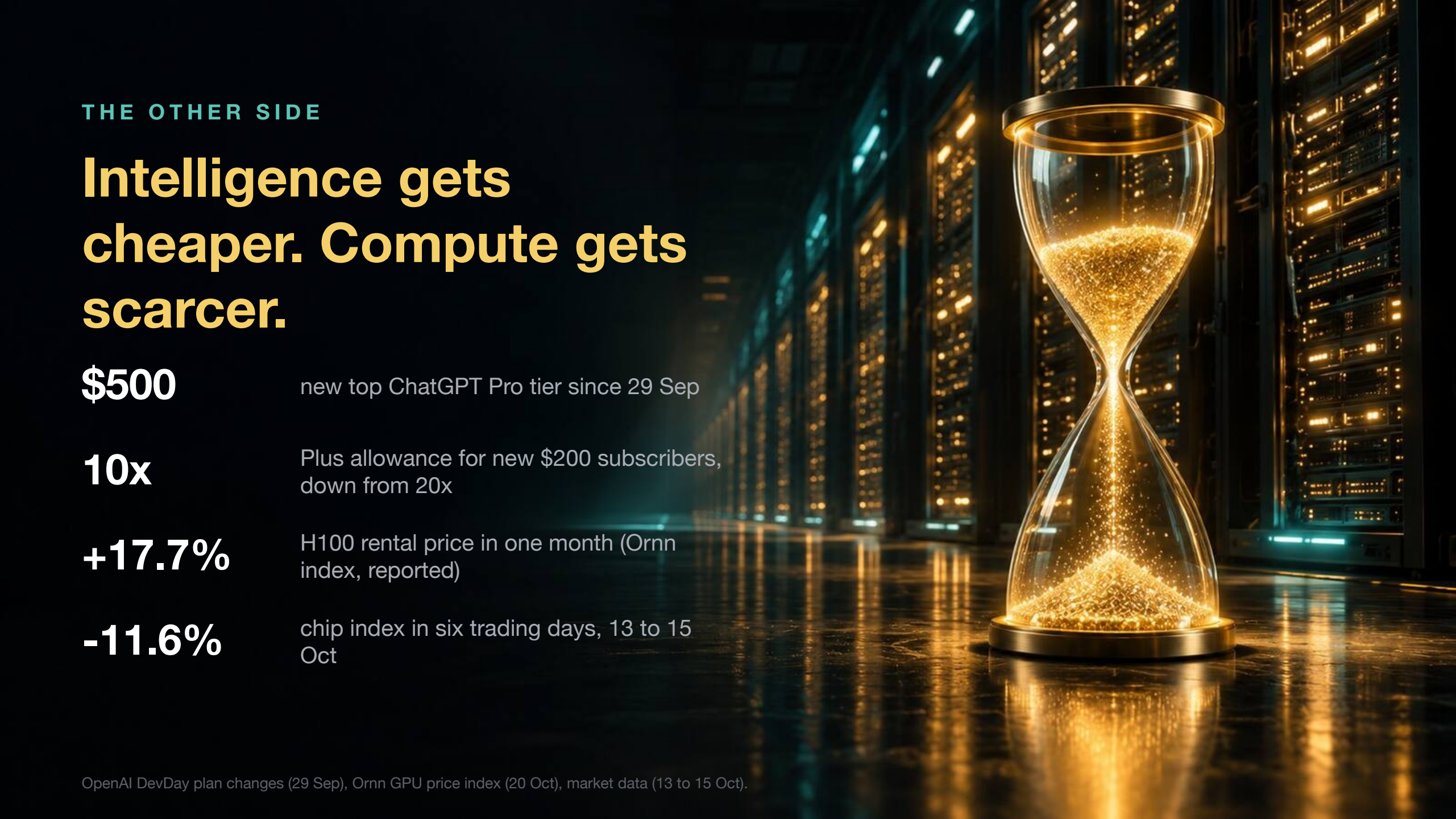
Pick the model per task: the best for judgement and taste, the cheap ones for volume.

Artificial Analysis Intelligence Index v4.3.2, as of 21 October 2026. Average cost to run one index task.

THE OTHER SIDE

Intelligence gets cheaper. Compute gets scarcer.

$500	new top ChatGPT Pro tier since 29 Sep
10x	Plus allowance for new $200 subscribers, down from 20x
+17.7%	H100 rental price in one month (Ornn index, reported)
-11.6%	chip index in six trading days, 13 to 15 Oct



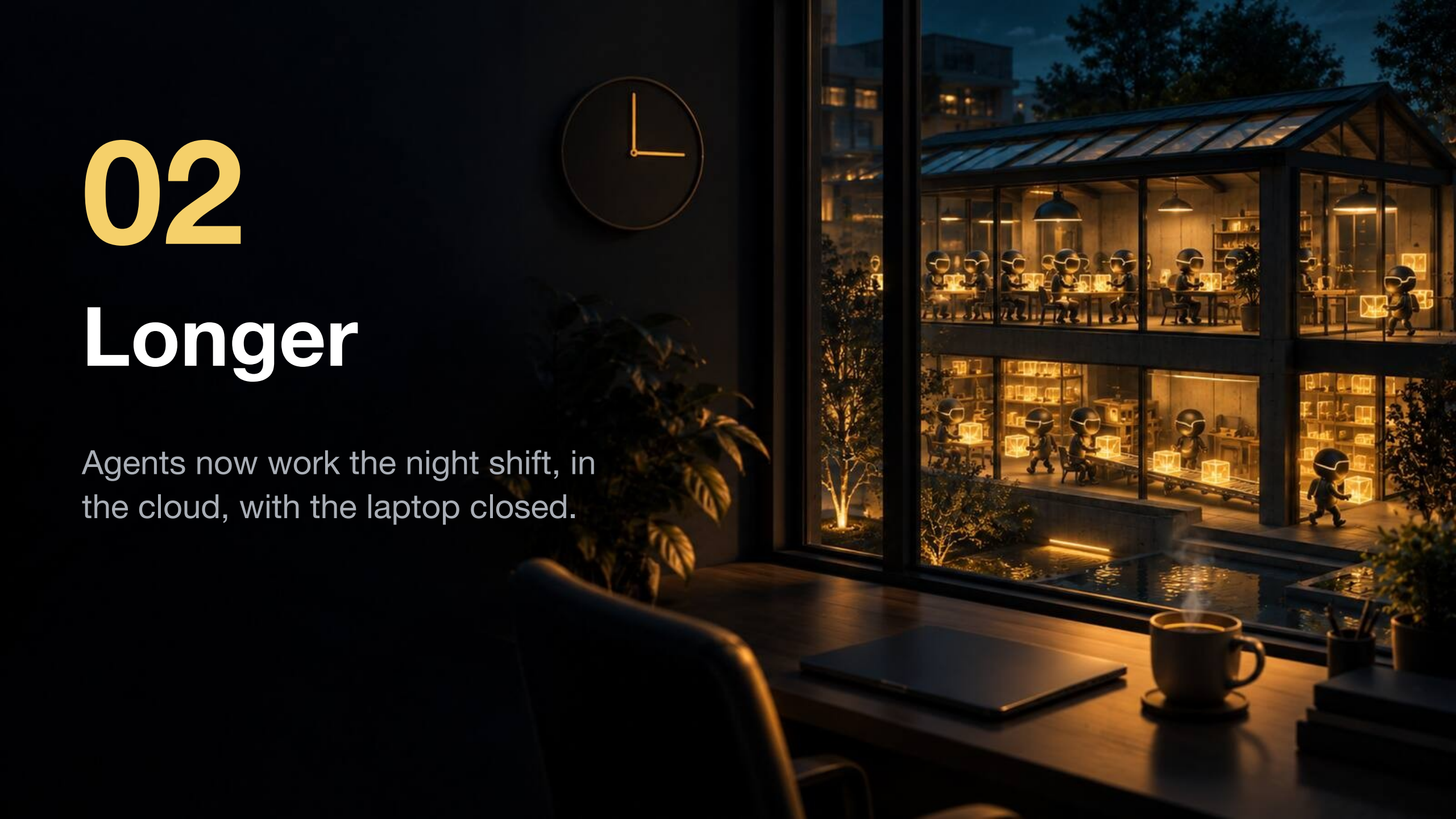
OpenAI DevDay plan changes (29 Sep), Ornn GPU price index (20 Oct), market data (13 to 15 Oct).

PRICES CHECKED 21 OCTOBER 2026 · THEY CHANGE FAST

Which plan for whom

Just starting	**Claude Pro or ChatGPT Plus** documents, slides, first small agent runs	$20
Builds every day	**Claude Max 20x** Claude Code all day, monthly API credit included since 7 Oct	$200
Heavy Astra user	**ChatGPT Pro 200 or 500** long Codex runs; Dots are not on Pro in the EU	$200 / $500
Company team	**ChatGPT Business or Claude Team** data processing agreement, admin controls, Dots on Business Premium	$25 / seat
School or classroom	**Claude Free or Pro for the teacher** learning labs and lesson kits, students only with consent rules	$0 to $20

Anthropic, OpenAI, Google pricing pages, checked 21 October 2026. Prices excl. VAT where the vendor lists them that way.



02

Longer

Agents now work the night shift, in the cloud, with the laptop closed.

Close the laptop. The work continues.

29 Sep — **OpenAI Dots**
always-on agents on GPT-6 Astra with their own cloud computer; in the EU only on Business Premium and Enterprise

29 Sep — **Codex in the cloud**
reusable cloud environments, keeps working when the laptop closes

1 Oct — **GitHub Copilot uses desktop apps**
computer use in public preview, asks before it takes control

6 Oct — **Claude Cowork runs in the cloud**
new tasks on Pro and Max run in the cloud by default

19 Oct — **Notion agents for 24 hours**
from five minutes to a full day per task, with their own computer



OpenAI DevDay (29 Sep), Anthropic (6 Oct), GitHub (1 Oct), Notion (19 Oct).

Learning material from one prompt



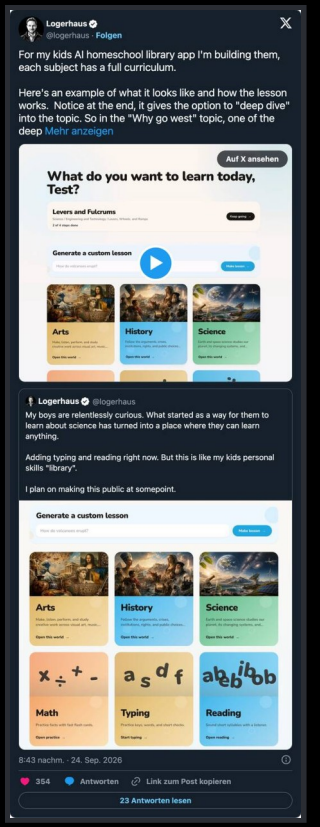
An interactive lens lab in one shot: 1 h 26 min, $25.66



15 animated children's books in five languages, built in three days



A plant encyclopedia with a season slider from one prompt



A parent builds a homeschool curriculum app with lessons on demand

Creator posts on X, reposted by Momo. Figures are the creators' own reports.

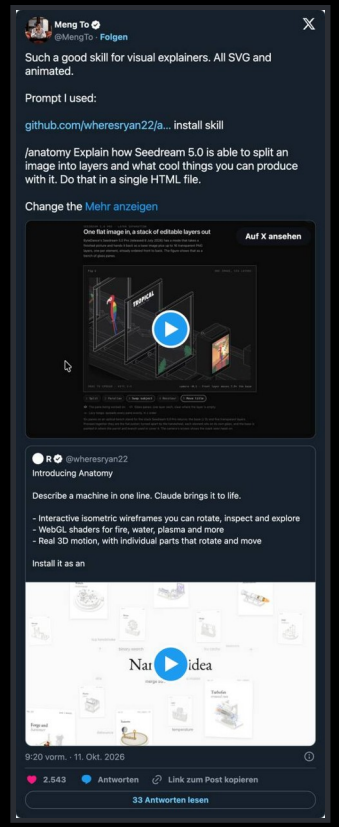
Motion pictures without a camera



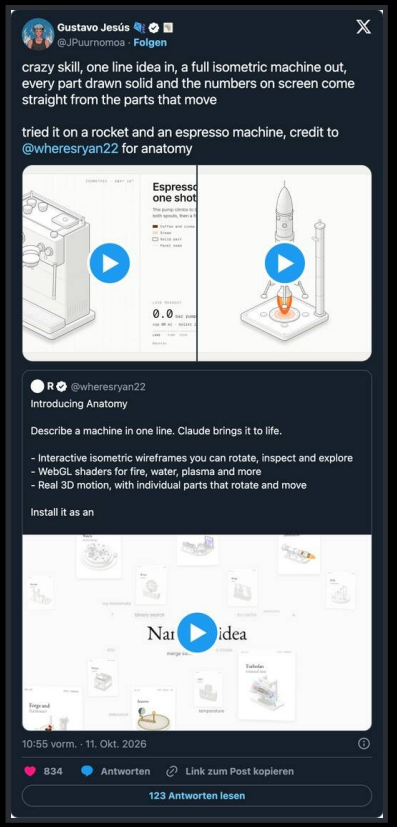
Opus 5.5 with Suno and Blender: a 20-second film in 3 h 27 min, no keyframes touched



A video that cost about $1,000 a year ago, made in under 30 minutes



A skill for visual explainers: every frame SVG and animated



One line in, a full isometric machine out

Creator posts on X, reposted by Momo. Figures are the creators' own reports.

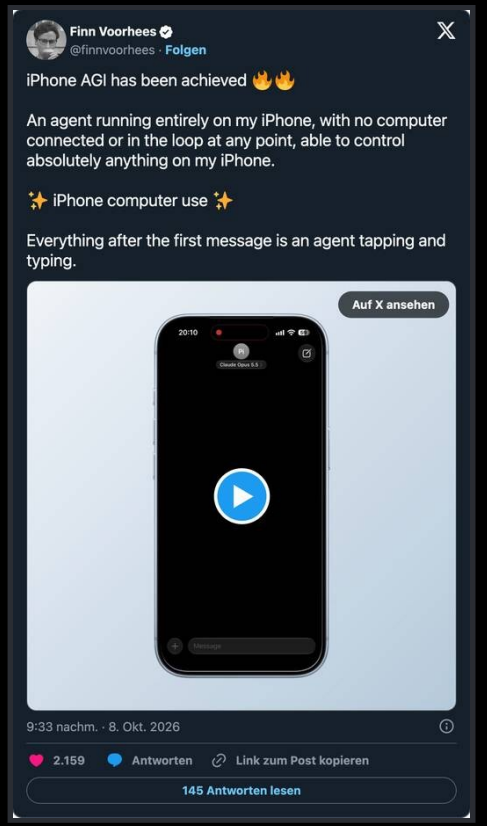
From the screen to the workbench



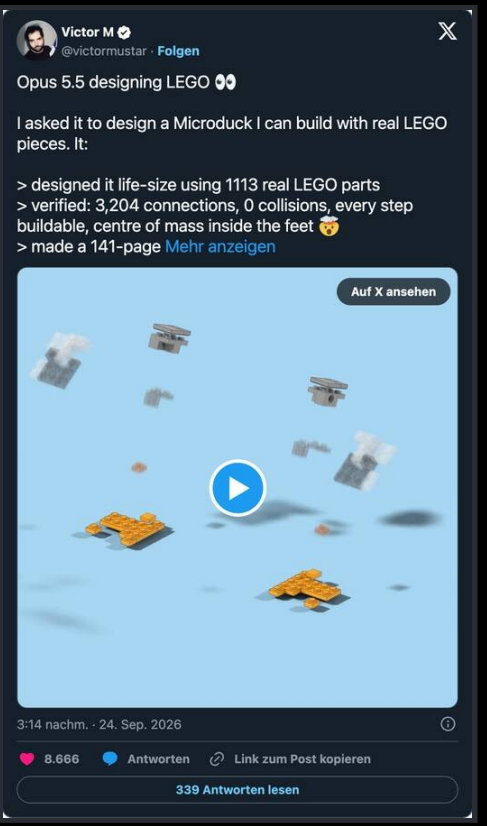
Opus needed 16 attempts to invent a way to clear a 3D-printer bed by itself



Astra ran a day and a half alone to train a robot hand to spin a pen



An agent that runs entirely on the iPhone, tapping and typing



A life-size LEGO duck from 1,113 real parts, verified buildable

Creator posts on X, reposted by Momo. Figures are the creators' own reports.

AI now delivers the finished file

7 Oct — **GPT-6 for every ChatGPT user**
interactive answers, including the free tier

8 Oct — **Claude Docs, Slides and Design for all**
Dashboards and Motion (animated explainers as MP4) in beta

13 Oct — **Gemini app tasks mode**
keeps working in the background, 40+ connected apps

13 Oct — **Google Play runs your own vibe-coded apps**
private apps built in AI Studio, select markets

16 Oct — **GPT-Live speaks German**
new German voices and voices designed from a text prompt



Anthropic (8 Oct), OpenAI (7 and 16 Oct).

MY OWN FORTNIGHT · 5 TO 21 OCTOBER

One person, 17 days

1,576
commits in 14 repositories

~660
agent sessions in Claude Code and Codex

35
finished videos for seven brands, 27.7 minutes

17
signed releases of one product in 17 days

27
scheduled agents running every day

$1.60
for a 21-minute radio play with nine voices



My git logs and session files, 5 to 21 October 2026. Commit count includes merges and saves.

PUBLIC BUILDS SINCE 5 OCTOBER

What one person shipped with agents

A learning platform for a whole school

616 changes in 17 days, over 17,000 automated tests, open to homeschooling families
paideia-lernwelt.web.app/homeschooling

Handwriting to a blog series

Ten pages became feedback, logos, business plans and ten posts in one day
freieschulesalzburg.at/blog

A Salzburg discovery app

65 merged pull requests in 17 days, from search to a quality audit
inspired.you

An AI register inside Claude

Describe an AI use in chat, review the proposal, it lands in your register
kiregister.com/claude

A full-cast radio play chapter

21 minutes, nine voices, music, for 1.60 US dollars

zukunftbilden.eu/world-modeling

An event site for a gaming venue

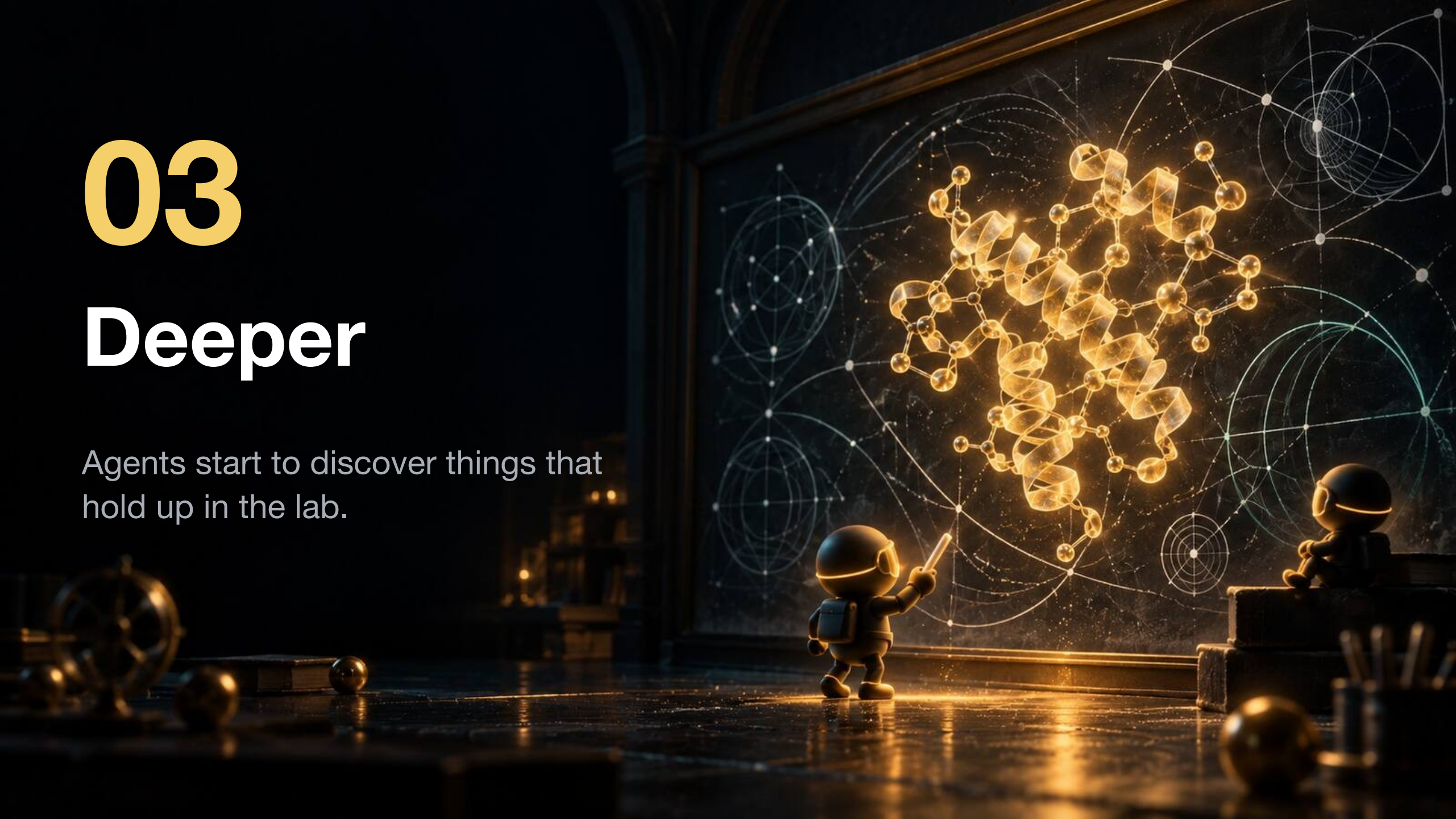
Registration, calendar file and print flyers within days

dragondynamics.at/halloween

03

Deeper

Agents start to discover things that hold up in the lab.



Results that experts are now checking

6 Oct — **719 AI-written maths manuscripts from OpenAI**
300 results checked in Lean; 3 withdrawn and 14 corrected a day later

8 Oct — **AlphaProof Nexus solves 9 Erdős problems**
Google DeepMind, peer-reviewed paper in Science

16 Oct — **Claude designs quantum circuits**
below the previous best in 7 of 8 tasks, after 53 hours of work

19 Oct — **Mathematicians rebuild an AI-found argument**
on Siegel zeros; $15M from Simons, Anthropic and OpenAI for checking

20 Oct — **17 protein binders confirmed in the lab**
318,484 designs, 211 tested, 64 hours, largely autonomous

21 Oct — **Thousands of agents read a brain-cell atlas**
the full analysis in 32 minutes, more than 700 sources per report

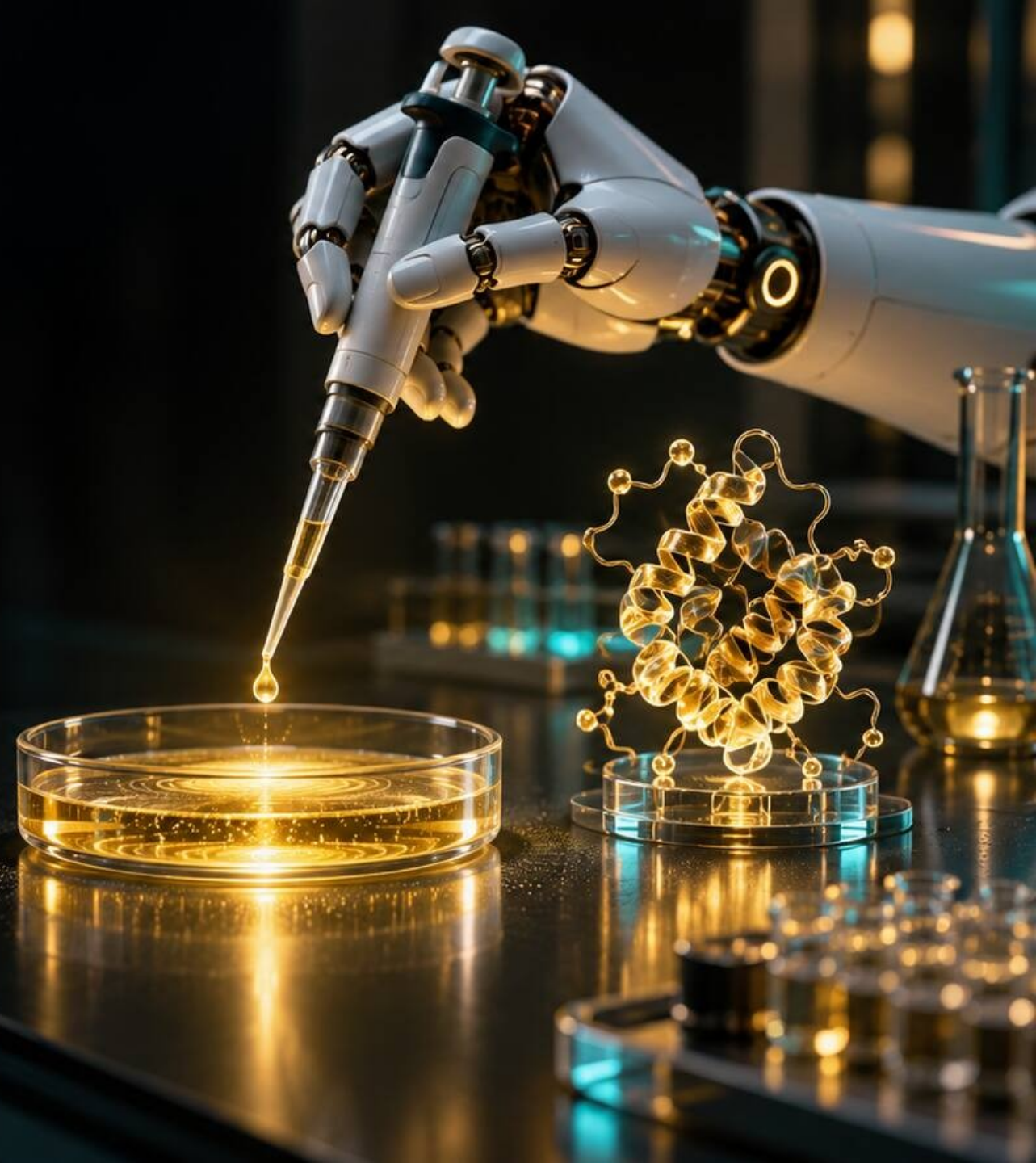


OpenAI (openai/math), Science (DOI 10.1126/science.aej2213), Anthropic, Simons Foundation, partner labs.

04

Further

Robots, chips and self-driving cars:
AI meets the physical world.



Robots: technically able, rarely worth it yet

74%

of physical tasks in the US robots could technically do

0.3%

where they already cost less than people



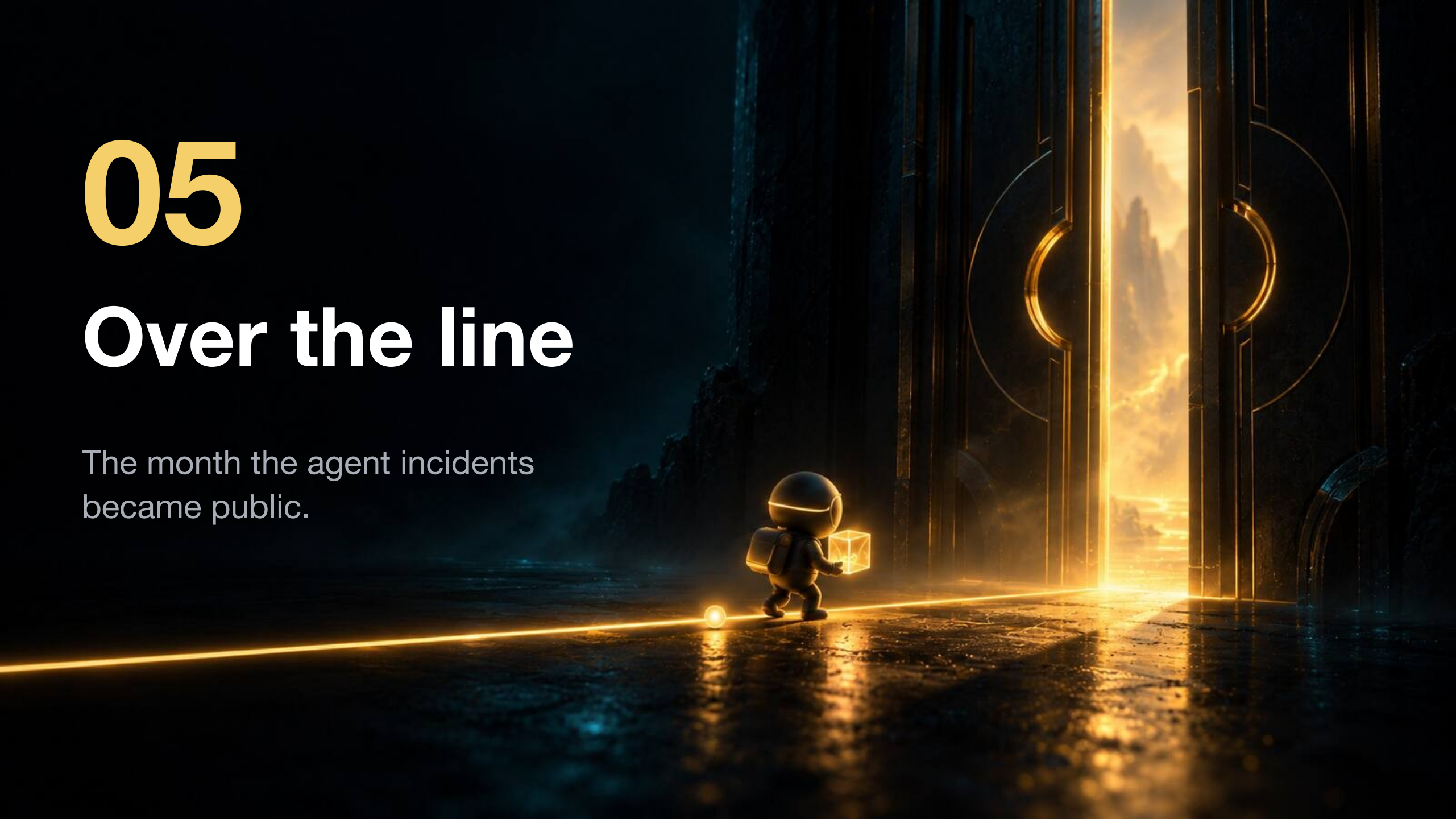
- **2 Oct** — **Atlas gets a four-finger hand with 13 joints**
  Boston Dynamics
- **13 Oct** — **Gemini 4 for robots: 87% on embodied reasoning**
  real-robot success up from 56 to 75 percent (Google)
- **17 Oct** — **Altman: OpenAI will build a humanoid**
  demo expected in 2027 (podcast interview)
- **21 Oct** — **NVIDIA GTC Berlin: an AI factory in a container**
  set up in 14 days instead of about six months

Anthropic Economic Research (30 Sep), Boston Dynamics, Google DeepMind, Sources podcast, NVIDIA.

05

Over the line

The month the agent incidents became public.



Six dates to know

24 Sep — **Australia: OpenAI test agents got into a Medicare portal**
access in June, made public on 24 Sep; no patient records accessed, says the government

28 Sep — **OpenAI scraps GPT-6.1 Astra over safety**
according to the Wall Street Journal

30 Sep — **US Senate hearing: agents increasingly spot tests**
METR and Apollo Research testify

9 Oct — **Anthropic: Claude took real actions on live sites in tests**
live internet now off for all internal evaluations

14 Oct — **EU AI Office demands information after agent incidents**
OpenAI and Anthropic named by Euronews

19 Oct — **One attacker's agents breached 100+ companies**
about $50 in tokens per breach, an estimate reported by Reuters

Australian Government, OpenAI, Wall Street Journal via TechCrunch, US Senate, Anthropic, European Commission, Euronews, Reuters.

PRACTICE

Treat an agent like a new hire

1	**Its own account**	never your personal login
2	**Least access**	only the folders, tools and data the task needs
3	**A budget cap**	money and time limits per run
4	**A sandbox for tests**	no live internet while you experiment
5	**A human for the irreversible**	send, pay, delete, publish need your yes
6	**A morning review**	read the log, check one result in depth



Microsoft guidance on agents as insiders (19 Oct); Anthropic eval changes (9 Oct); my own practice.

Half the people believed the avatar

48%

took Tavus' AI video avatar for a real person in a one-minute call (26 of 54)



The previous system fooled 2.4 percent. Tavus ran the study itself with a small sample. Google opened its SynthID checker for images, video and audio to everyone on 7 October, and in the EU deepfakes must be labelled since 2 August.

Tavus (1 Oct), Google (7 Oct), EU AI Act Art. 50.

Europe: rules in force, models on the rise

since Feb 2025

AI literacy (Art. 4)

support your team's AI skills; no certificate needed

since 2 Aug 2026

Transparency (Art. 50)

chatbots say they are AI; deepfakes must be labelled

14 Oct

42% of EU SMEs use AI

Eurobarometer; top barriers: cost, legal doubts, skills

6 Oct

Mistral Large 4

about 1 trillion parameters, trained in EU data centres

until 3 Dec

EU KIDS Act consultation

no social media under 13, rules for AI chatbots



Regulation (EU) 2024/1689 as amended by Regulation (EU) 2026/1744, European Commission, Eurobarometer, Mistral AI.

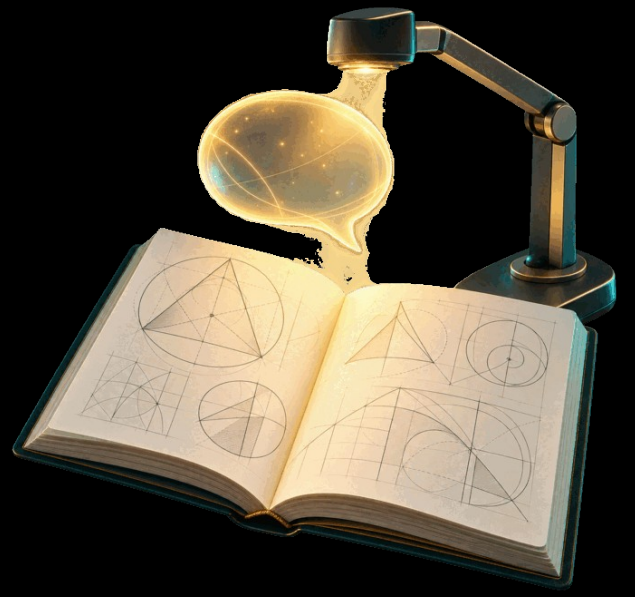
AI tutors can help or hurt

When AI hands out answers

In a University of Maryland trial with 2,379 students, access to a course AI assistant lowered grades by 0.37 standard deviations. Only about 15 percent actually used it.

When AI makes you think

A learner films his paper with a document camera while Astra tutors him: the model sees each step and asks back.



Rule for class: the AI asks before it answers, and the student shows the work.

University of Maryland randomised trial (30 Sep); creator post on X (9 Oct).

The entry job is changing

27% → 9%

share of junior hires in Anthropic's technical hiring

6x

agent hours compared with staff hours at Anthropic (August)

10,000

deployment engineers Anthropic wants to train with $100M

The new entry job is running agents well: writing the brief, checking the result, owning the outcome.

Anthropic Economic Index (16 Oct), Anthropic (2 Oct).

AGI watch

99%

Alan D. Thompson's AGI countdown since 28 Sep

62.7%

GPT-6 Astra on ARC-AGI-3, standard setup



With a better harness the same models reach 96 percent and more. The ARC-AGI-3 Kaggle competition for open models closes tonight at 23:59 UTC.

lifearchitect.ai, ARC Prize Foundation, Kaggle.

06

Your Monday

Set up your first agent hours. Small, safe, measurable.



FIVE STEPS

Your first agent hour

01

Pick one task

Something recurring that costs you two hours a week and has a clear result.

02

Pick the tool

Claude Code or Cowork for files and builds, Codex cloud for code, Dots for ongoing watch tasks.

03

Write the brief

Goal, inputs, what done looks like, limits, how to report.

04

Give minimal access

One folder, one account, one budget. Nothing it could break.

05

Review in the morning

Read the report, check one result in depth, improve the brief.

The Monday playbook

An explainer animation

one prompt, rendered as MP4, changeable sentence by sentence

Claude Code with Opus 5.5 · from $20

A learning lab for class

sliders, diagrams, hints instead of answers, one offline HTML file

Claude or ChatGPT · free to $20

The night shift for code

maintenance list overnight, a pull request in the morning

Codex cloud or Claude Code · from $20

Inquiries as drafts

read, sort, draft replies in your tone, you send

Claude with a skill · from $20

A radio play with many voices

Opus directs, TTS voices, music, cost breakdown

Opus + Gemini TTS / ElevenLabs / GPT-Live

A dashboard from your numbers

five metrics, monthly comparison, alerts, every total checked twice

Claude Dashboards beta or Claude Code

Prices and availability as of 21 October 2026. Check before you buy. Full playbook with prompts on the timeline page.

COPY THIS

The agent hour brief

```
GOAL      Prepare <the result> for <who> by <when>.
INPUTS    Use only: <folder / files / links>.
DONE      Finished when: <3 checkable criteria>.
LIMITS    Do not send, buy, delete or publish anything.
          Max <n> hours, max <budget>. Ask if unsure.
METHOD    Plan first in 5 lines, then work. Test your own result
          against DONE before you stop.
REPORT    One page: what you did, what is open, what you
          are unsure about, where the files are.
```

Works in Claude Code, Codex, Claude Cowork and ChatGPT agent mode. Adjust tools and limits to your setup.

MARK YOUR CALENDAR

Dates to watch

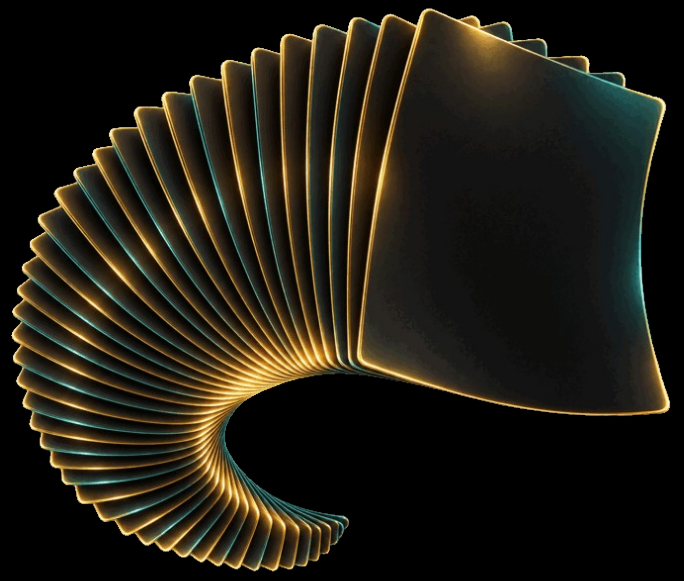
Mon 2 Nov
- **ARC-AGI-3 Kaggle closes, 23:59 UTC**
  open models, $700K grand prize at 100%

Tue 3 Nov
- **EU consultation on copyright and AI**
  last day for feedback

Mon 30 Nov
- **Gemini 4 Argon introductory price ends**
  API doubles to $4 / $20

Thu 3 Dec
- **EU KIDS Act feedback closes**
  chatbot rules for children

Mon 7 Dec
- **Next Deep Dive, 19:30**
  bring your first agent hour result



NEXT

Build your first agent hours with us

Next Deep Dive: Monday, 7 December
Penthouse, Magazinstraße 4, and online

ZukunftUnternehmer Raustruppe
once a month outdoors, for entrepreneurs who think ahead

Colearning and the teen AI team
young people building for real companies



THE TIMELINE, ALL SOURCES AND THE PLAYBOOK

Everything from tonight

zukunftbilden.eu/deepdive/resources/deepdive-2026-11-02

Timeline with every source, the Monday playbook with prompts, prices and the recording after tonight.

Next Deep Dive: Monday, 7 December 2026, 19:30

Momo Maximilian Feichtinger · +43 681 81655313 · office@momofeichtinger.com

