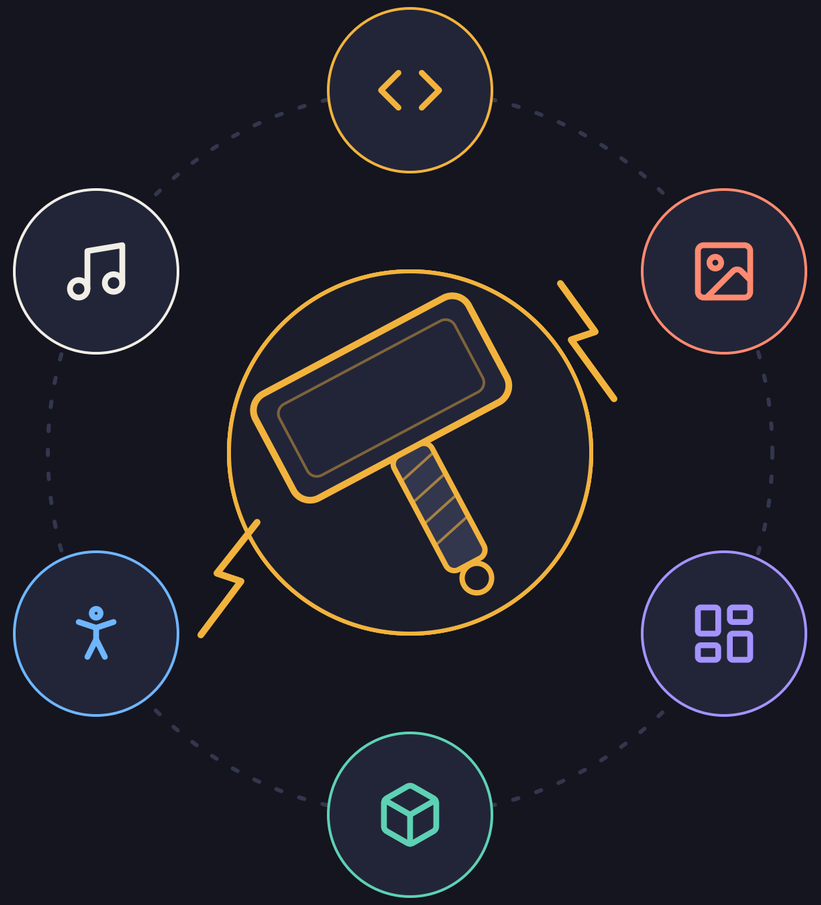
TALK · SERIOUS GAMES × AI

Thor's Hammer

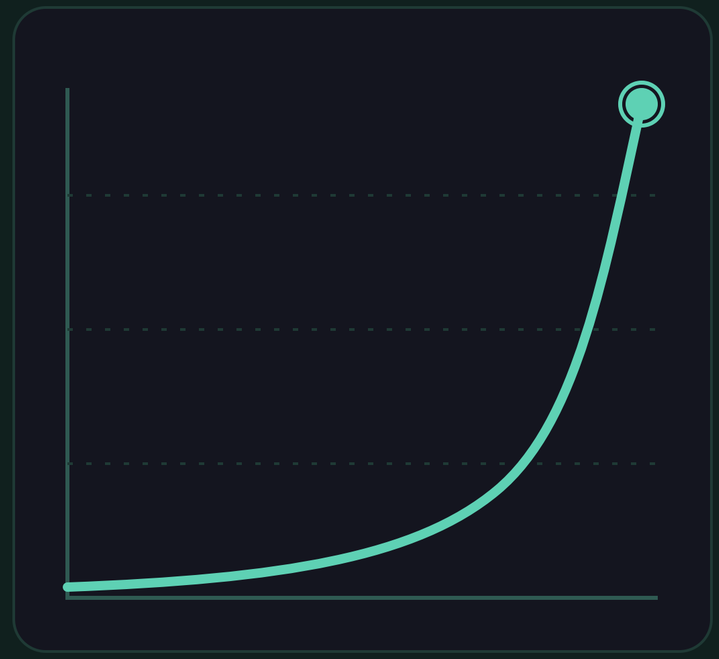
AI and LLMs in game development: what changed, how we use it, and what to do with it.



01 · HOW WE GOT HERE

Ten years, one curve

From a research paper in 2017 to agents that build games in 2026.



01 · MILESTONES

From a research paper to a colleague

Jun 2017

Transformer

Google's paper "Attention Is All You Need". The engine inside every model since.

2020 · 2021

GPT-3, Copilot

Plausible text. Code autocomplete inside the editor.

Nov 2022

ChatGPT

About 100 million users in two months. Everyone starts pasting.

Mar 2023

GPT-4

The first model whose game scripts often just work.

2024

Taste and reasoning

Claude 3.5 Sonnet builds UIs; o1 thinks before it answers.

2025

Agents

Claude Code, Codex: AI that edits whole projects and runs them.

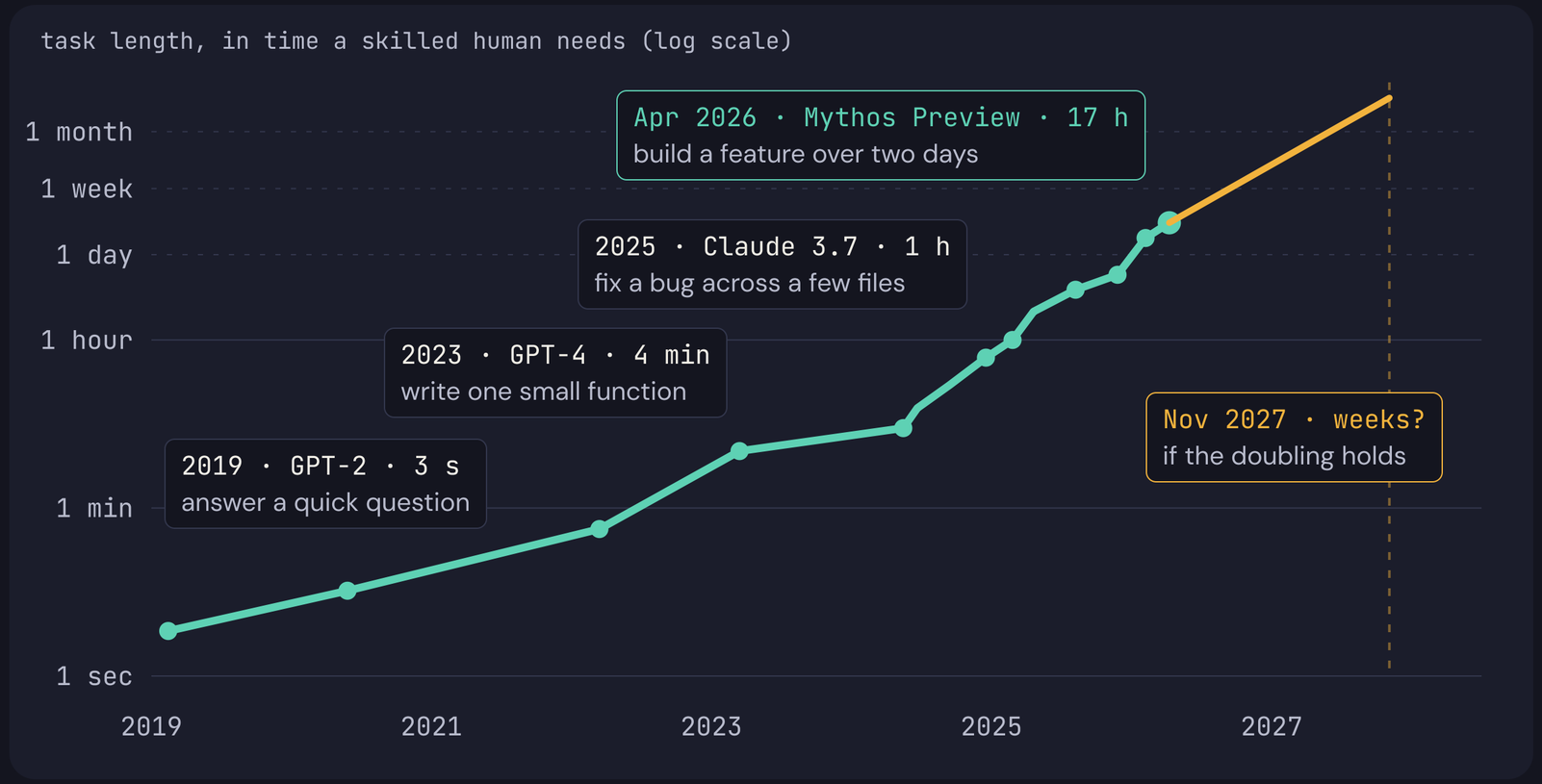
Sep 2026

Astra, Opus 5.5

A rough brief in, a working game out. Most of the time.

01 · THE CURVE

AI can work alone on longer and longer tasks



THE QUESTION

How long a task can AI finish on its own?

Length is the time a skilled human needs for the same task. It counts if AI succeeds at least half the time.

APRIL 2026

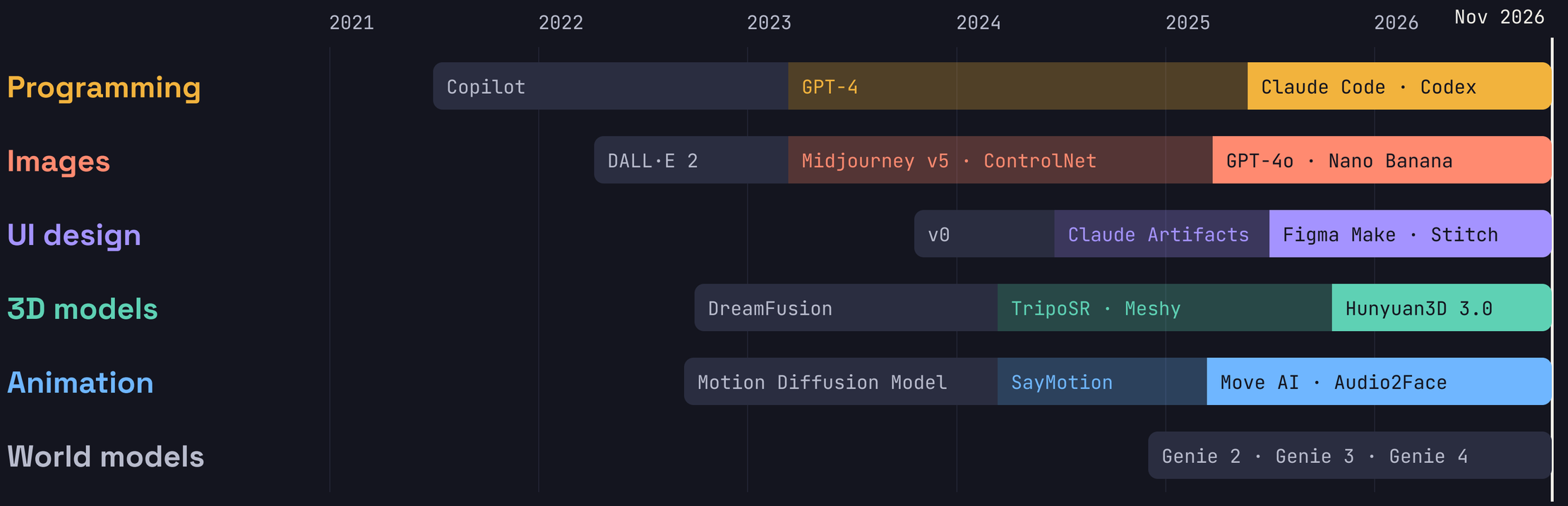
17 h

about two working days (4 min in 2023). Doubles every ~4 months.

METR Time Horizon 1.1 · metr.org/time-horizons

01 · CRAFT BY CRAFT

When each craft became usable for games



01 · TODAY

The frontier, November 2026

OPENAI · 3 SEP 2026

GPT-6 Astra

Pitched as OpenAI's best model yet for software engineering and using a computer. So strong at hacking that access was gated at launch.

ANTHROPIC · 22 SEP 2026

Claude Opus 5.5

Top-tier quality, about 40% cheaper and 30% faster than Opus 5. Leads agentic coding tests like Terminal-Bench.

GOOGLE · 13 OCT 2026

Gemini 4 Argon

A million tokens in and out: a whole codebase, design doc and asset list in one prompt.

THE CODING TEST EVERYONE USED

2% → 94%

SWE-bench, 2023 → 2026. Solved, then retired.

CHATGPT, EVERY WEEK

1.2B

people, September 2026

WHAT THAT MEANS FOR US

Work for hours alone: write, run, test, fix. Drive the editor. Decent UI and text, still not great.

01 · A YEAR FROM NOW

What works today, and what may work next

	TODAY · NOV 2026	NOV 2027 · MY PROJECTION
Works alone for	about two working days	several weeks
Code	a playable prototype from a rough brief	a vertical slice from a design doc
Art & 3D	concept art, props, UI mock-ups	characters and rigs get close
Testing	you play every build yourself	agents play, report bugs, tune balance
Worlds	Genie: minutes of walkable demo	short playable prototypes

WHAT THE FORECASTERS SAY

Anthropic: “a country of geniuses in a datacenter” around 2027.

OpenAI: an AI researcher by March 2028.

AI 2027 authors: automated coder, median mid-2028.

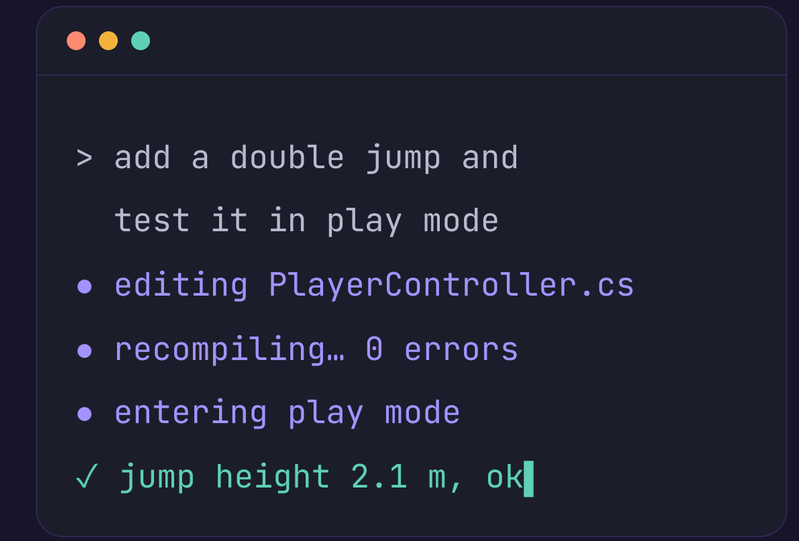
CAUTION

Extrapolations are not promises. Chips, power and money can slow it. Nobody measures “is it fun”.

02 · HOW WE USE IT NOW

The agent moved into the engine

Four ways of working with AI, and where each one breaks.



02 · FOUR WAYS OF WORKING

How we actually use it

2023

Copy, paste, pray

Paste the error into ChatGPT, paste the script back, hope it compiles.

`never saw your project`

2024

Assistant in the IDE

Copilot, Cursor: sees your files, suggests as you type.

`you still glue it all together`

2025

Agent in the terminal

Claude Code, Codex, Gemini CLI: reads the repo, runs commands, fixes its own errors.

`blind to the scene`

2026

Agent in the engine

MCP and the Unity CLI: builds scenes, plays the game, looks at screenshots, runs tests.

`can break things fast`

02 · MCP AND THE UNITY CLI

The agent can now touch the engine

YOU

A short brief

“add a double jump”

AGENT

Claude Code, Codex, Cursor…

plans, writes, checks its own work

ENGINE EDITOR VIA MCP / CLI

Unity · Unreal · Godot

- create objects
- edit scripts
- read the console
- recompile
- enter play mode
- take screenshots
- run tests
- build

THE LOOP

write → compile → play → look → fix ↺

MCP is a standard plug, like USB-C for AI tools.

02 · UNITY, NOVEMBER 2026

Two ways in: an MCP server or the CLI

Inside the Editor

Aug 2025

Unity AI (6.2): Assistant with Ask, Plan and Agent modes; generators for sprites, textures, sound.

2026

AI Gateway: run Claude Code, Codex, Gemini or Cursor in the Assistant window.

Both still in beta.

	MCP SERVER	UNITY CLI
How	the agent connects and calls the server's tools	the agent runs commands in a terminal
Example	MCP for Unity (CoplayDev) community · ~50 tools · ~16k stars	`unity test · build · recompile` official · beta since July 2026
Context	every tool description takes up the agent's context	looks up commands with `--help` when needed
Editor	must be open, with the bridge package installed	drives an open Editor, or starts its own for tests and builds
Best at	live scene work: objects, components, prefabs	the loop: compile, test, build, with clear exit codes

ALSO

unity mcp · the CLI serves MCP too

Unreal 5.8 · official MCP, experimental

Godot · community MCP

02 · HONEST TRADE-OFFS

Where it helps, and where it still fails

Benefits

Prototypes in hours, not weeks

It closes its own loop: compile, test, screenshot, fix

Glue code, editor tools, build scripts

Any engine, any language

Drawbacks

You own code you didn’t write, and can lose the overview

Speed-ups are real but uneven: experts were 19% slower in a 2025 trial

Subscriptions, usage limits, cost

Editor access means it can wreck a project: commit often

Still fails

Scene and prefab files: huge text it guesses at

Judging what looks and feels right

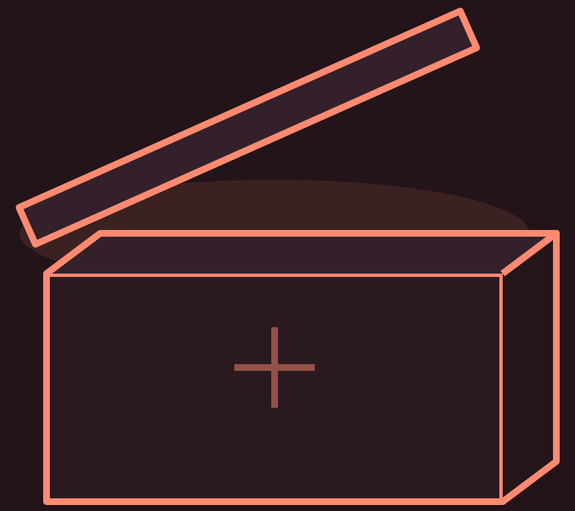
Less common systems: old Unity GUI, Blueprints, niche plugins

The official engine tools are still beta

03 · PANDORA’S BOX

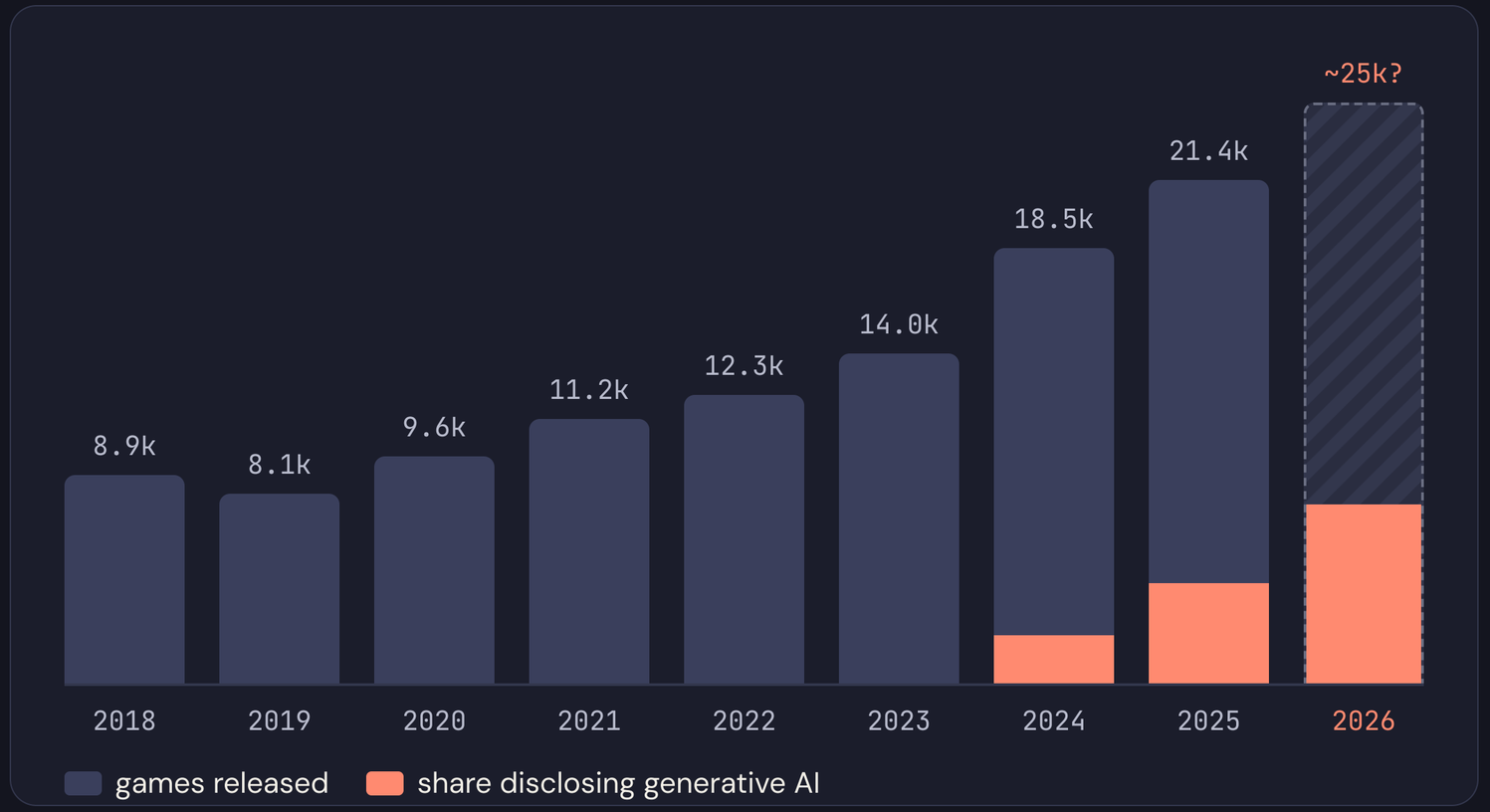
Everyone can make a game now

That is the good news. It is also the bad news.



03 · THE FLOOD

Steam is drowning in new games



NEW GAMES PER DAY, 2026

68

up from about 24 a day in 2018

2025 RELEASES

49%

had 10 reviews or fewer by December

SteamDB (June 2026 snapshot) · GamingOnLinux · Sulka Haro's census of 53,597 releases

03 · WORD OF THE YEAR

Slop

slop *noun*

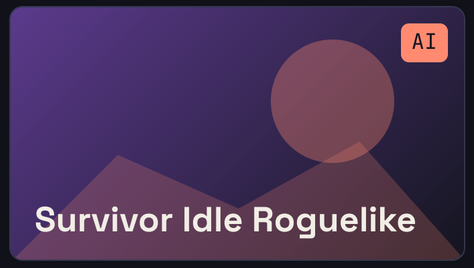
Low-quality digital content, usually mass-produced with artificial intelligence.

MERRIAM-WEBSTER · WORD OF THE YEAR 2025

Macquarie 2025: “AI slop”

Next Fest forums: “vibecoded slop”

PC Gamer: “haystack slop is a thing now”







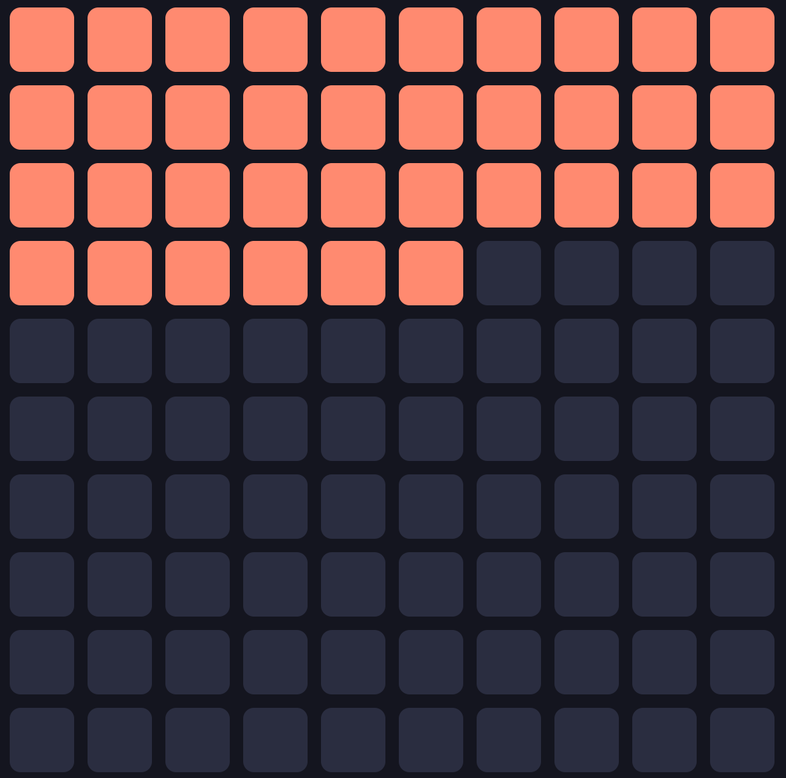






03 · STEAM NEXT FEST, OCTOBER 2026

One in three demos disclosed generative AI



36% 2,593 of 7,145 demos

JUNE 2026 FEST	TOP 25 DEMOS
19.5%	8%

Most AI-labelled demos sit in the long tail. The games people actually played mostly didn’t disclose it.

03 · OUT OF THE BOX

What’s already happening

Cloned before release

A 50-second clip

Freya Holmér posts a Tetris-with-rotation prototype. Within days, up to four clones hit the app stores. One took a day of prompting.

404 Media · Jul 2026

Build in public, in reverse

Needle in a Haystack

A prototype gets 50M views. About ten Steam lookalikes follow in days; a clone becomes a top Next Fest demo.

Dexerto · How To Market A Game · Oct 2026

Hiding the game

Playtests go private

The Moonlighter studio moves its playtest to Discord only. Another studio pulls its Next Fest demo, fearing clones.

Game Developer · Oct 2026

Decompiled and remade

Ports in weeks

An “100% AI-generated” PC port of N64 Smash Bros. in 25 days. Decomp.dev shuts down: AI forks broke the community.

Time Extension · VGC · 2026

Art backlash

Awards taken back

Clair Obscur loses its Indie Game Awards Game of the Year over AI placeholder textures. Crimson Desert apologises and audits its assets.

Engadget · PC Gamer · 2025–26

Whose work is it?

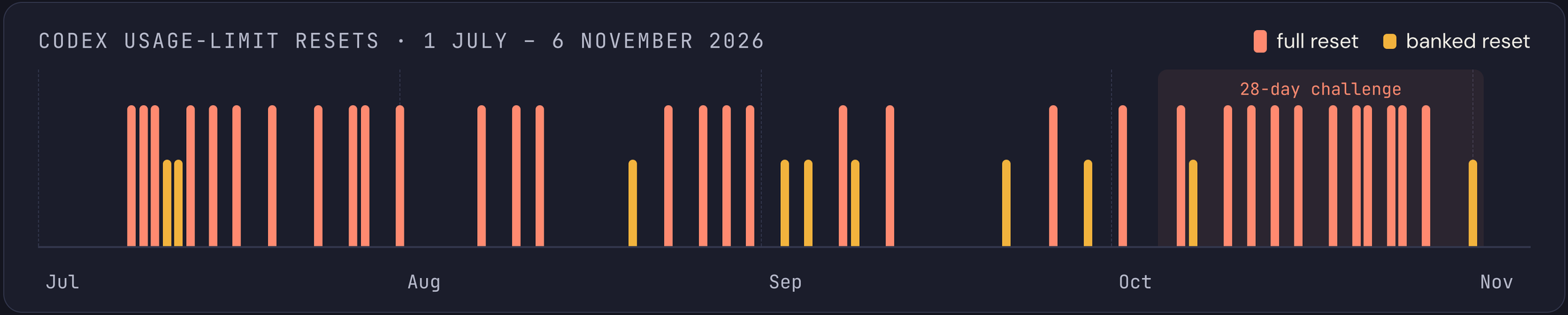
Courts and consent

Disney and Universal sue Midjourney. Anthropic pays authors $1.5B. 52% of developers say gen AI harms the industry.

GDC 2026 survey · court filings

03 · THE OTHER HOOK

AI labs gamify their own users



SINCE 9 JULY

43

resets, one every ~3 days

23 in the nine months before

WHAT DEVELOPERS SAY

“This feels a lot like free spins in slot machines.”

Others call it the gacha of coding, hit 100% every week, and watch their baseline creep up. Karpathy: ~16 hours a day directing agents. The Valley calls it “token anxiety”.

MY TAKE

The first hit is free.

Random rewards, daily events, limited-time bonuses: our playbook. Hook people and companies, then set the price.

whenaireset.com · codex-resets.com (unofficial logs of Tibo’s posts) · Hacker News “Codex Resets” · No Priors · Marketplace

04 · THOR'S HAMMER

Six months of building with it

What I learned making games and apps with AI since April.



04 · SINCE APRIL

Nineteen projects, from prototype to store



Paint Clouds

Paint the sky with procedural clouds



Living Pages

Children’s drawings come alive on paper (AR)



Cool Places

29 puzzles about shade, heat and cities



Power Places

Learn electricity by powering a village



One More Sol

Mixed-reality space game in your room



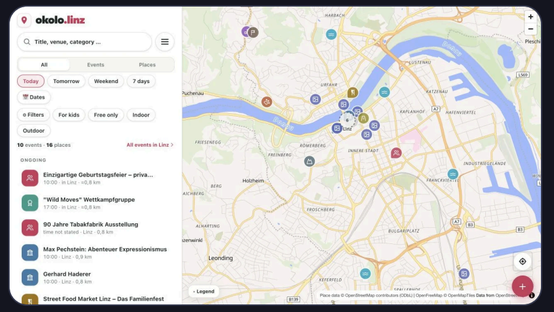
Handsfree Inc.

Comedy VR puzzles: your hands go to work



Spomen

Keep your family’s stories and voices



Okolo

What can we do around here this weekend?

04 · WHAT CHANGED THIS YEAR

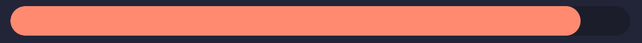
From micromanaging to briefing

SPRING 2026

Micromanage

Great for fast prototypes. But I fixed every second step, and the UI and text were weak.

HOW OFTEN I STEPPED IN



SUMMER 2026

Guide

Whole features landed. I still reviewed everything and steered every design decision.

HOW OFTEN I STEPPED IN

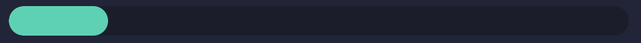


SINCE OPUS 5.5 & ASTRA

Brief

A rough spec gets past 90%. Sometimes it surprises me with ideas I didn’t have.

HOW OFTEN I STEPPED IN



04 · NOT JUST “MAKE ME A GAME”

Use it for everything around the game

Prototype fast

Ten rough versions of an idea in a week. Keep the one that is fun.

Simulate balancing

Let bots play 10,000 rounds overnight and show you where the economy breaks.

Sandbox mechanics

A throwaway scene to test one ability, one variable, one new rule.

Build your own tools

A level editor that fits your game, in hours instead of months.

Explore branches

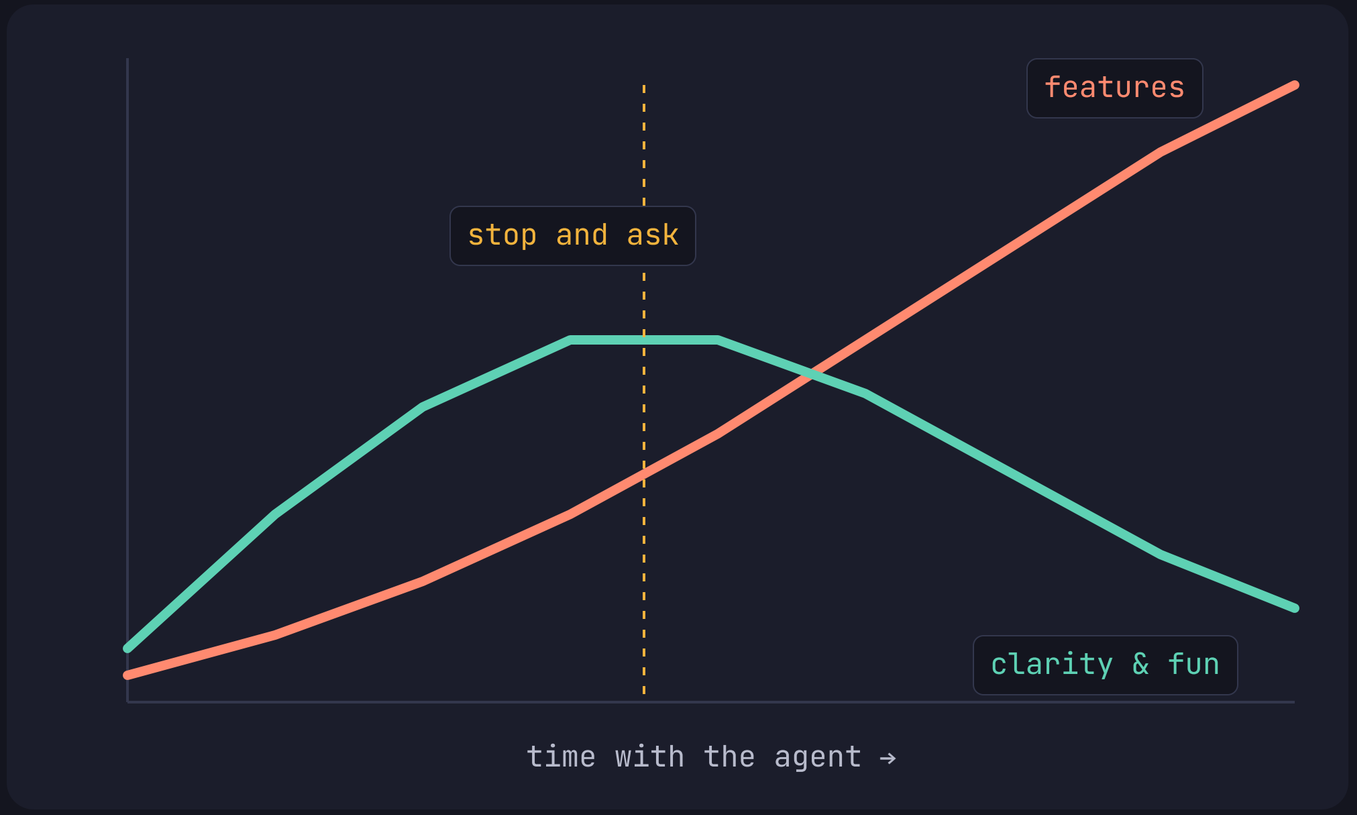
Try the paths players could take that you never had time to build.

Keep your craft

Use it for code and iteration. The art, writing and ideas can stay yours.

04 · THE TRAP

AI is great at adding. Design is removing.



BEFORE THE NEXT FEATURE, ASK

- What is my game about?
- Why is it fun?
- What are the 2–3 core mechanics?
- What can I cut?

04 · PICK THE RIGHT TOOL

Choose the stack for the game, not the habit

HOW WELL AGENTS HANDLE IT · MY EXPERIENCE



Why the gap?

Models are as good as their training data. There is far more public web code than Unity UI code, and Blueprints are graphs saved in binary files.

AI can work in any stack, so you are no longer locked into the one you learned.

04 · IT’S A TOOL

AI didn’t invent slop. People did.

BEFORE AI

Slop is old

Asset flips, shovelware, the hundredth clone of a hit. AI is just a faster shovel in the same hands.

NOW

The bar went up

Good studios use it to move faster and keep their own art and ideas. Not using it at all is now a disadvantage.

AN IDEA USED TO COST

years

→

NOW IT COSTS

weeks

Build it, see if people like it. If they don’t, move on to the next one.

THE TAKEAWAY

You've been handed Thor's hammer.

Build slop with it, or learn to build something good.

Learn it | Use it | Find where it helps | ... and where it doesn't



Questions?