

**How building with AI  
can double the  
throughput of your  
engineering team**



How building with AI  
can **triple!!!** the  
throughput of your  
engineering team

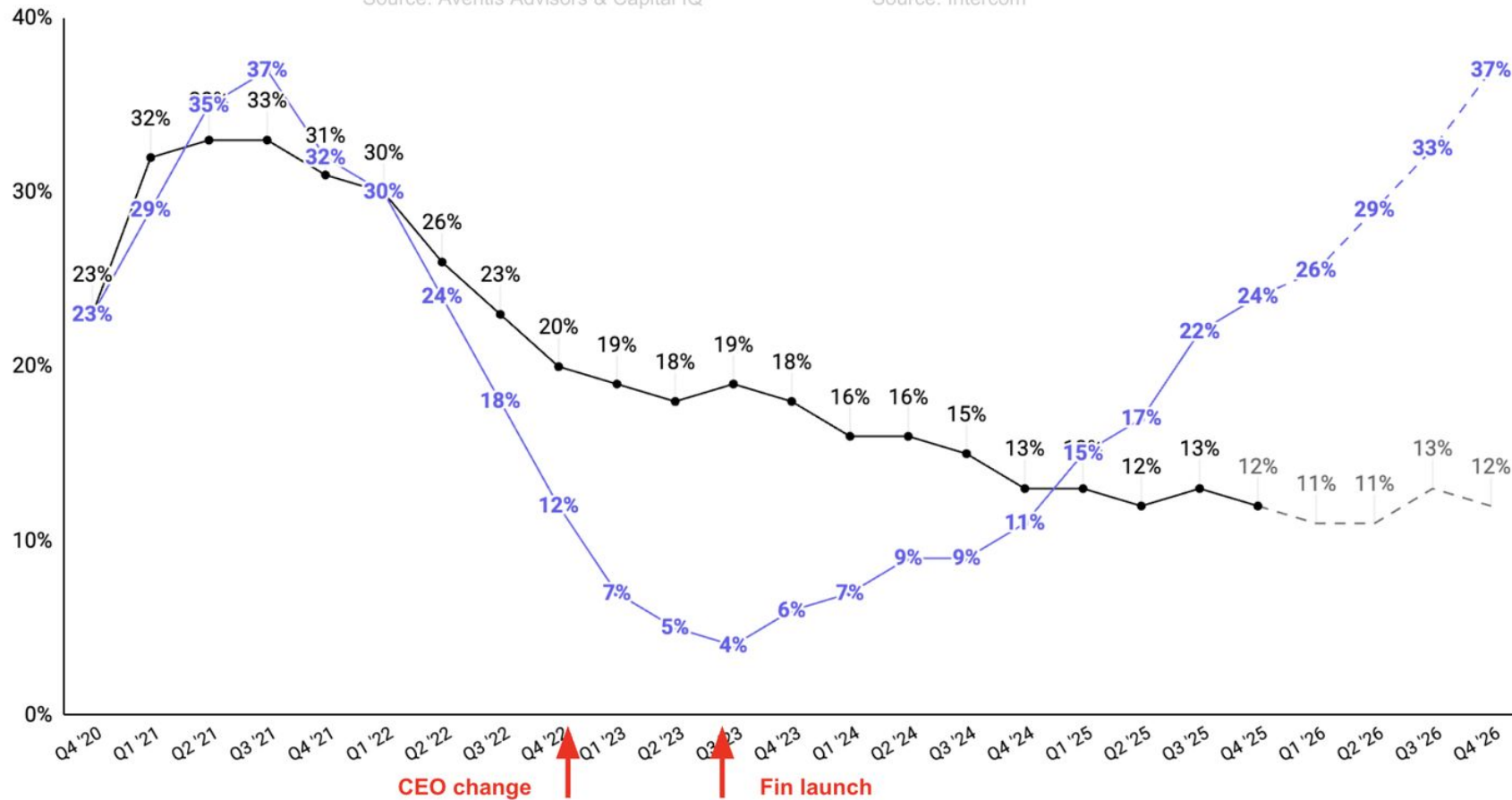


● SaaS growth rate

● Intercom growth rate

Source: Aventis Advisors & Capital IQ

Source: Intercom



## *Software? No Way. We're an A.I. Company Now!*



By [Sarah Kessler](#)

Feb. 14, 2026,  
8:00 a.m. ET

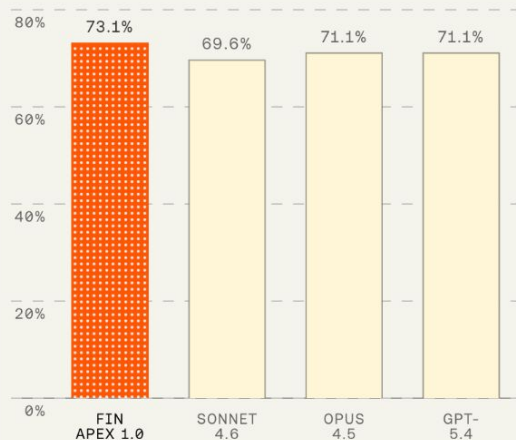
When ChatGPT came out in November 2022, Eoghan McCabe was about a month into his second stint as C.E.O. of Intercom, the customer service software company he had co-founded a decade earlier. Growth had ground to a near halt, and now he had another problem.

McCabe foresaw an existential crisis: If chatbots could answer customer queries, and if they replaced customer service agents, companies “wouldn’t need help desks — they wouldn’t need the software anymore.”

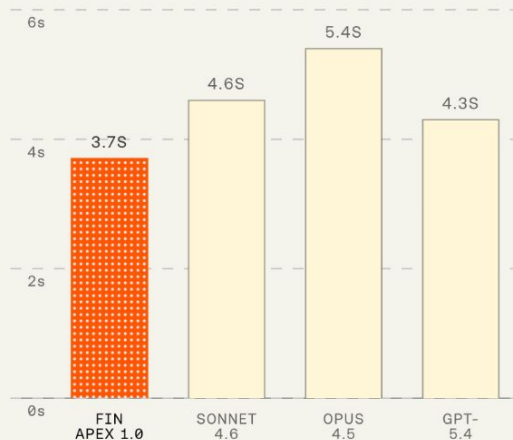
He decided that Intercom needed to become an artificial intelligence company, or risk becoming obsolete.

In production, **Fin Apex 1.0 outperforms frontier models** with higher resolution rates, fewer hallucinations, and faster responses.

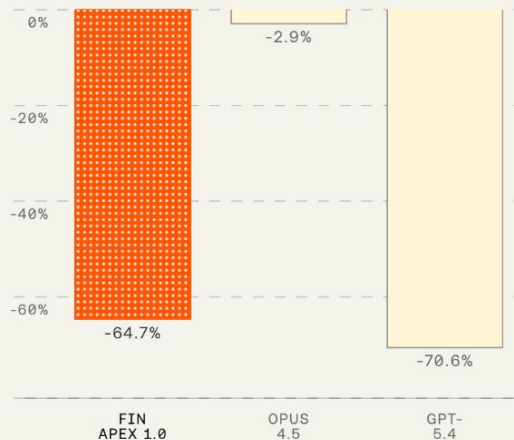
**2.8%** higher resolution rate



**0.6 seconds** faster time to first token

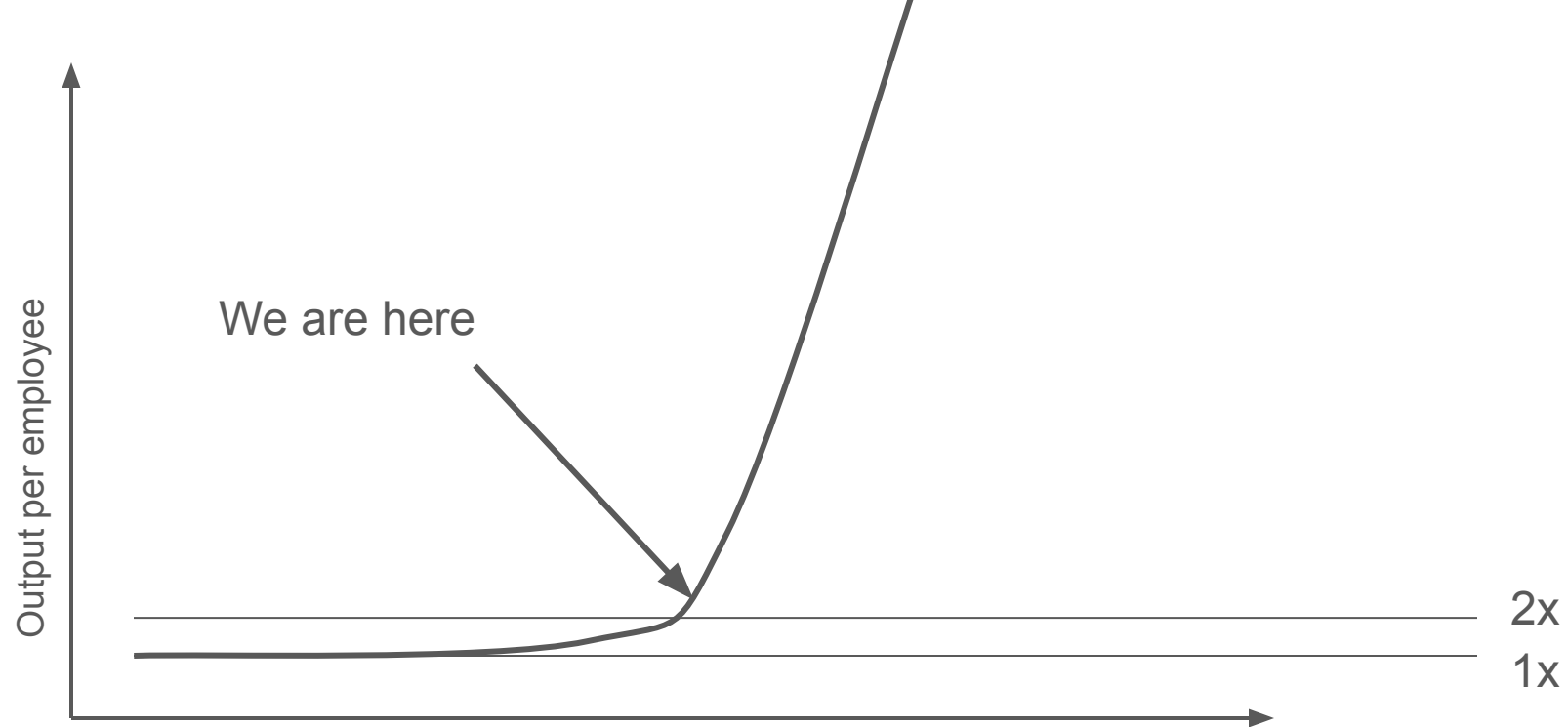


**-65%** reduction in hallucinations\*



\*Compared to Sonnet 4.6

**2x**





## 2x

Darragh Curran June 13 2025

*What follows is a version of an email I sent our entire R&D team about an explicit goal and deliberate action we'll take to become twice as productive through our embrace of AI.*

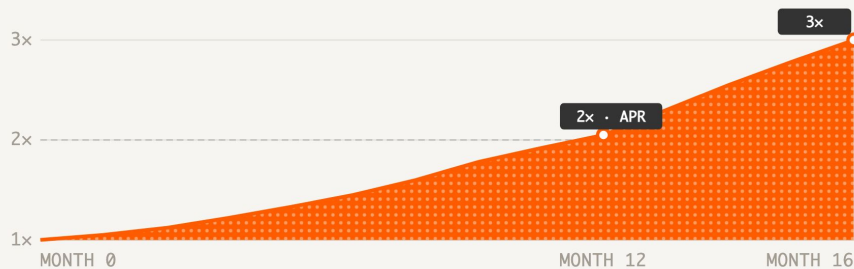
Advancements in AI are precipitating the most transformative shift in software engineering of the past 25 years. It's dramatically expanding what software can do and transforming how we build it. We are all novices in this new world, and that's filling me with wide-eyed excitement.

If we were to literally hit pause on further advancements, I'm convinced any engineering team just leveraging the already existing tools effectively should expect at least **double their current productivity** – a 2x improvement. Yet most people and teams in the industry at large are not getting close to this today, they aren't trying, and they probably don't believe it's possible, and even if they do, behavior change is hard and the forces or incentives aren't clear yet.

# We hit 2x — and kept going

R&D PRODUCTIVITY MULTIPLIER · 16 MONTHS (ILLUSTRATIVE)

**3x** HIT 2x IN UNDER 12 MONTHS —  
TRIPLED OVER 16, NO PLATEAU  
YET



Curve shape illustrative. Hit 2x early April (~2 months early); ~3x over 16 months. Exact series TBC.

**-54%**

**PRODUCT DEFECT BACKLOG**

Source: 2x blog (9-month figure)

**~2x**

**FASTER IDEA-TO-SHIPED**

>2x product changes shipped · blog

**÷2**

**COST PER MERGED PR**

Roughly halved · outline

Unweighted PRs   TrueThroughput (Weighted)   Mean vs p90



## DESIGN 2X VISION

**01. Defining what to build**

**03. Ensuring quality at scale**

*Agents execute the middle*

**02. Building the rails agents run on**



**Mario Kostelac**

@mariokostelac



We are at inflection point.

Wake up!



**tobi lutke**



@tobi · Jan 8

I shipped more code in the last 3 weeks than the decade before. The top AI models / agentic systems right now are an entirely different thing to what people used until the beginning of December.

6:46 PM · Jan 8, 2026 · **216** Views

# Change is hard

- Update job descriptions and expectations.
- Talk **CONSTANTLY** about the urgency of adopting AI
- Reward great work financially, socially, publicly
- Give people room to learn, enablement, support, access to tools
- Be very specific about what you want to see and how it is to be done

# AI-First Expectations Per Engineering Level

This doc establishes a set of base expectations that apply to every engineer, followed by additional level-specific expectations.

It is in addition to the expectations defined in [Engineering Levels v3.2](#).

## Base expectations (all levels)

Every engineer, regardless of level, is expected to:

1. **Work AI-first by default.** Claude Code is the primary way of working, not an occasional supplement. AI is used across the full development lifecycle: planning, coding, debugging, reviewing, etc.
2. **Use the platform broadly.** Go beyond basic code generation. Leverage platform skills across the team's remit - debugging, CI, code review, investigation, deployment - to operate more independently and effectively. Share what works - effective workflows, useful skills, patterns you discover.
3. **Demonstrate accelerating throughput.** We expect to see a meaningful and ongoing uplift in PR throughput as AI leverage deepens. This is not a one-time step change; it should be a continuing acceleration as skills and workflows improve.
4. **Take on broader and more ambitious work.** Use AI to bridge knowledge gaps and take on work that would previously have been considered too unfamiliar or too broadly scoped. The boundaries of what one engineer can own should be expanding.

# **Platform Strategy**

**Standardise on a single, skill-driven AI platform that utilises Intercom-specific guidance.**

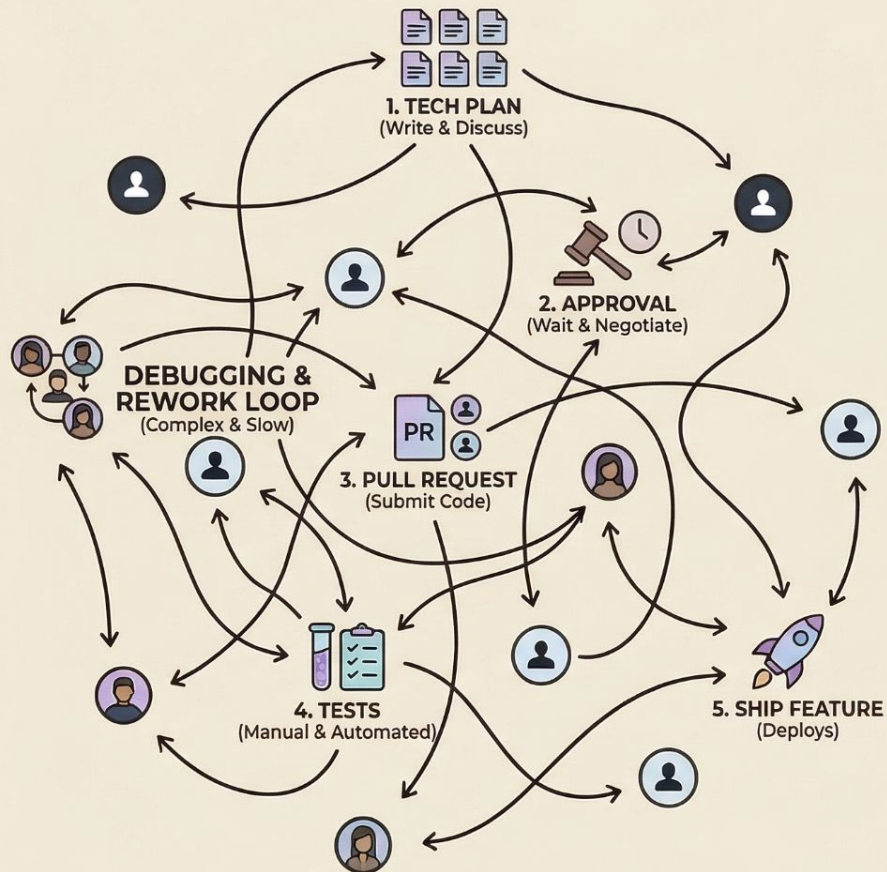
**A single extensible platform that all of R&D can leverage.**

**That has access to all Intercom tools, and the necessary skills to use them well.**

**Utilising human-curated Intercom specific engineering guidance for code generation and review**

**Combined with a flywheel of continuous improvement of the guidance**

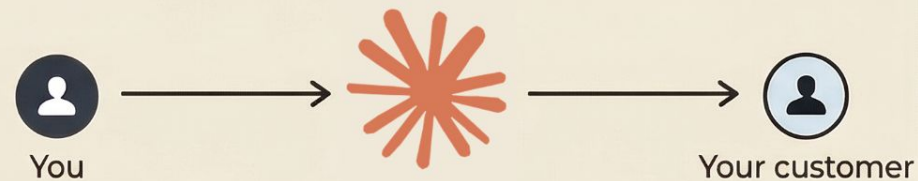
## THE OLD WAY



VS

## THE NEW WAY

VS



# **All technical work is becoming agent-first.**

Agents must be able to perform anything a human can do.

Documentation is infrastructure.

Engineering is changing. Agents write, test, and review code, while engineers focus their time on writing specs, validation, and improving the agents.

# Run Less (AI) Software

We focus efforts on evergreen, Fin-specific capabilities, and aggressively deprecate internal tools when first-class vendor features are available.

# **Give Agents Problems, Not Tasks**

Agents should be given problems and figure out the necessary tasks on their own.

Habitually and urgently fill in the gaps.

# Intercom AI maturity

1. Writes code / responds to incidents / bug reports by hand
2. Uses AI auto-complete / asks Claude questions
3. Agents are writing the code.
4. Rarely reads code. Doesn't really open an IDE any more.

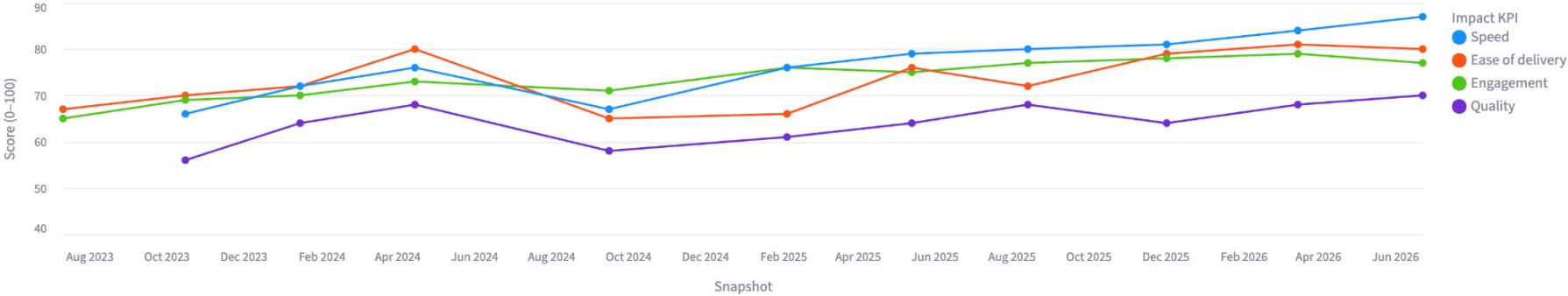
# Intercom AI maturity

5. Writes skills for others to use.
6. Writes self-improving high-quality skills with evals for others to use. Verifiably and repeatedly making others highly productive. Mines session data to find gaps in skill coverage.
7. Optimising our environment for agents (environment work-practices, software architecture, infrastructure, etc.)

# DevEx — DX Snapshot Surveys

Org-wide Developer Experience sentiment from DX snapshot surveys. Source: `rd.dx_snapshot_*` (imported from the DX data cloud).

## Impact KPIs over time

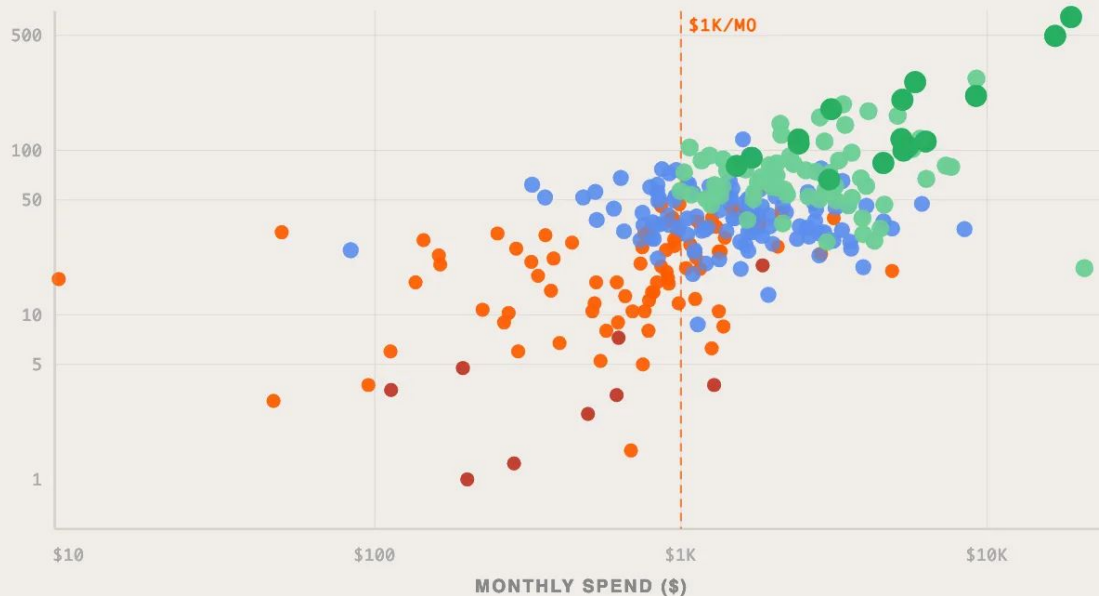


Org-wide trend across all snapshots. Pick a snapshot below to drill into the full detail.

SPEND (\$) VS WEIGHTED PRS MERGED PER ENGINEER (LOG-LOG)

**>\$1K** MONTHLY SPEND  
CORRELATES WITH HIGHEST  
THROUGHPUT

■ ELITE ■ STRONG ■ DEVELOPING ■ EMERGING ■ MINIMAL



Each dot = one engineer (287 of 297 with >0 PRs shown). Log-log axes. Color = tier (Elite/Strong/Developing/Emerging/Minimal). Vertical line at \$1K/mo threshold.

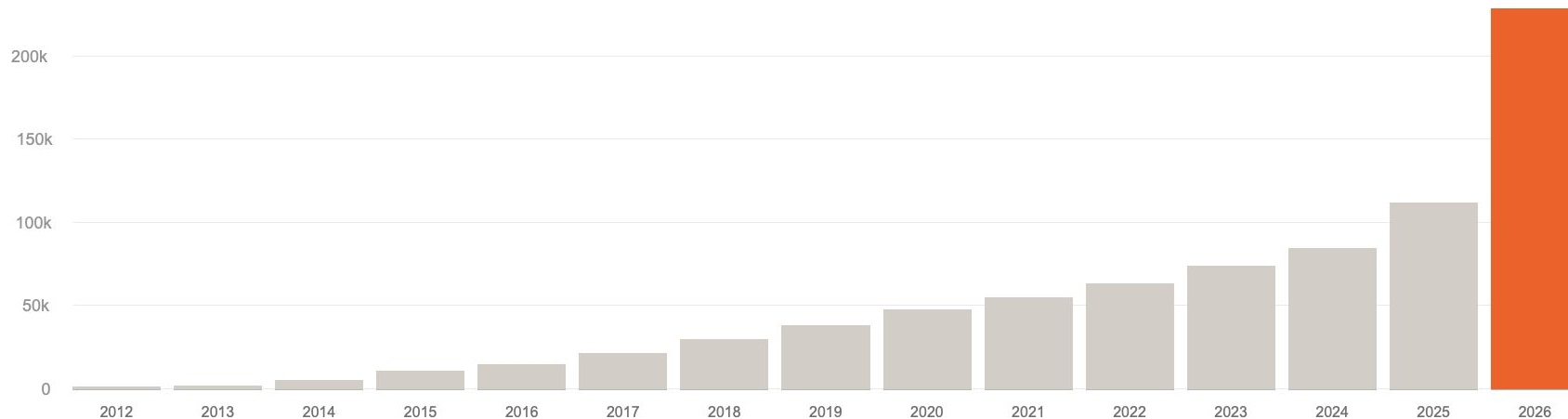
<https://ideas.fin.ai/p/2x-nine-months-later>

# From 1,142 tests to 228,626

Tests in the suite on August 31st each year

116,632

Added in the last 12 months



FULLY LOADED \$/PR

**45%** DECLINE IN  
COST PER PR



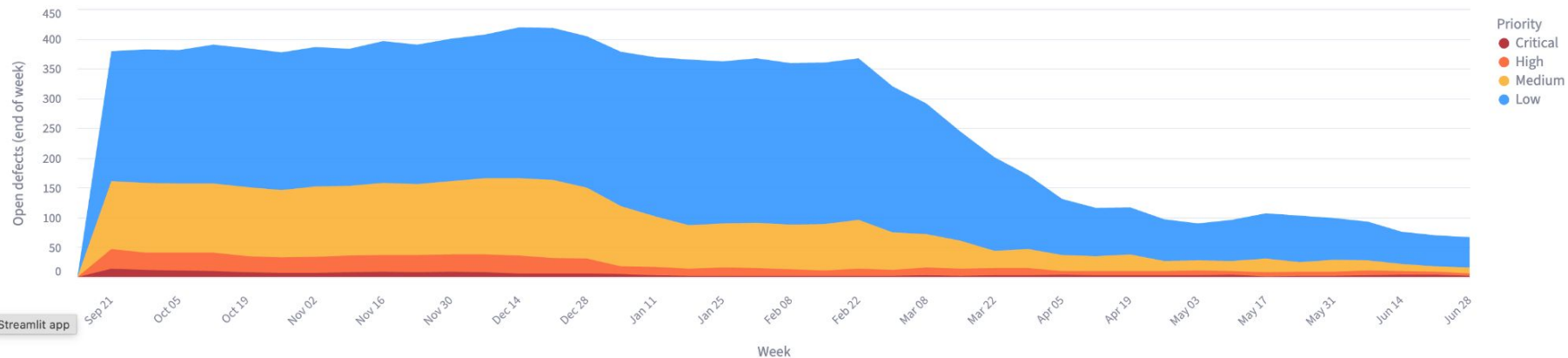
Fully loaded cost (payroll + AI) per merged PR. Dec spike from reduced holiday productivity. Increasing PRs/head drives cost down.

<https://ideas.fin.ai/p/2x-nine-months-later>

# Defects



Open Defect Backlog by Priority



68d

Median time to ship (all time)

64d

Median time to ship (90d)

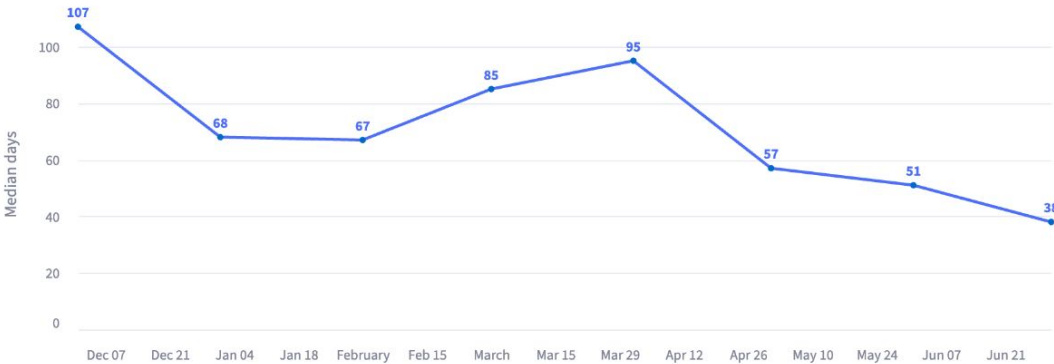
68d

Median time to ship (filtered)

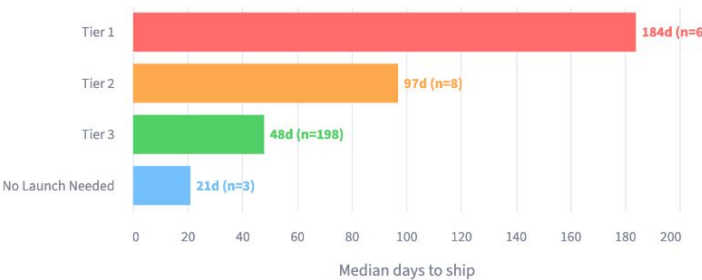
10.6

Feature throughput / week (90d)

### Median time to ship features



### By Launch Tier ?



## Good skills:

- Do one thing really well.
- Are created using a high-quality feedback loop.
- Use scripts for the deterministic parts.
- Have evals that are regularly used and maintained.
- Use progressive discovery to avoid bloating context.
- Have guardrails. Best-effort attempts by Agents that go off-script often result in bad outcomes.
- Are cleaned up after iteration. It's really common for skills to accumulate cruft over time, such as references to an old way of doing something – this information is usually irrelevant.
- Are regularly refreshed with new learnings. The skill itself can be instructed to do this, or periodic analysis of session data can achieve this.

[github.com/intercom/2x-skills](https://github.com/intercom/2x-skills)

## **tool-misses**

Find missing CLI tools and BSD-vs-GNU failures, fix them via Homebrew.

CLAUDE-CODE-TOOLS

## **fix-flaky-tests**

Diagnose and fix flaky tests across frameworks and CI.

TEST-TOOLS

## **thermo-nuclear-code-review**

Ruthless structural review — dramatic simplifications, not style nits.

CODE-REVIEW-TOOLS

## **pr-tools**

Open well-formed PRs; attach screenshots and recordings.

CREATE-PR · ATTACH-GITHUB-ASSETS

## **Other stuff**

---

**>1000 Claude Code weekly users - basically ALL of Fin**

**Deployment agents, exception agents,**

**Remote agents for one-shot tasks on demand**

**Automatic approvals > 40%**

**Single person team/product experiments**

**Product Engineers & Designers & Product Managers -> Product Builders**

# Friction, today

1. Claude Code writes a lot of tests.
2. Claude Code's English is shite.
3. Open-ended exploration still full of red herrings
4. Rapid growth means focus on costs - trade-offs are hard.
5. Over-reliance on agent output grinding teams down.

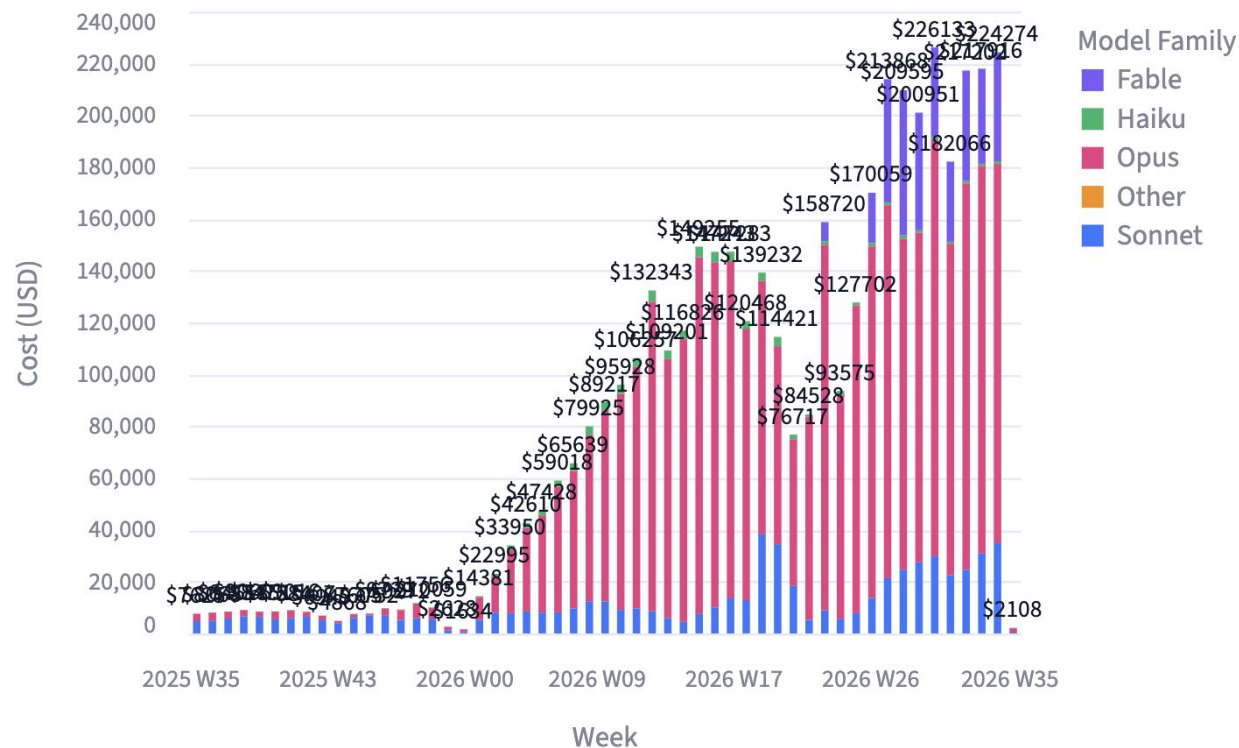
Cost per User

\$6318.79

Total Tokens

50,703,557,013

### Claude Code Costs by Model Family (Week)



Total Invocations

413,070

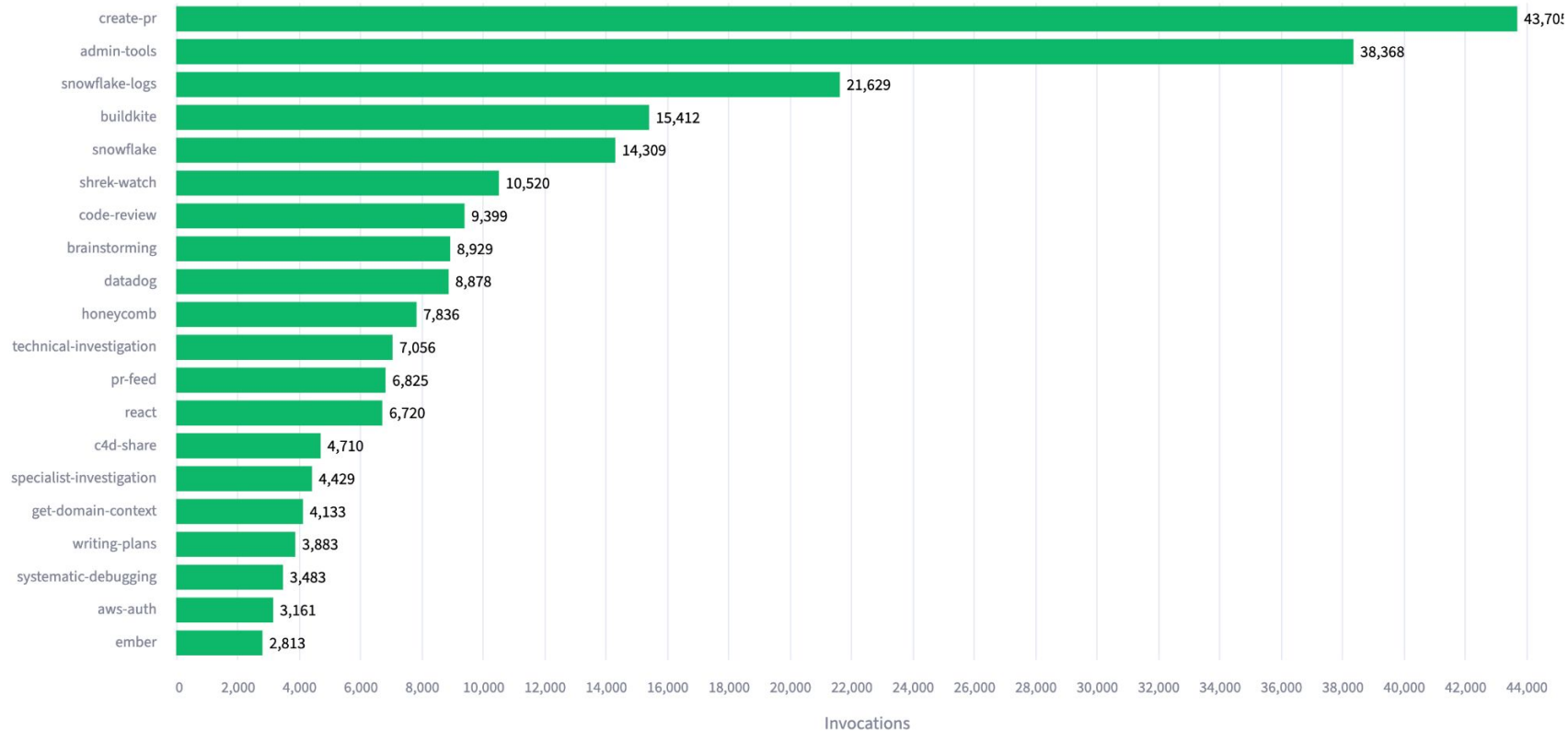
Distinct Skills

5,188

Users Using Skills

1,167

### Most Popular Skills



[brian.scanlan.ie](http://brian.scanlan.ie)

[github.com/intercom/2x-skills](https://github.com/intercom/2x-skills)

[ideas.fin.ai](http://ideas.fin.ai)

