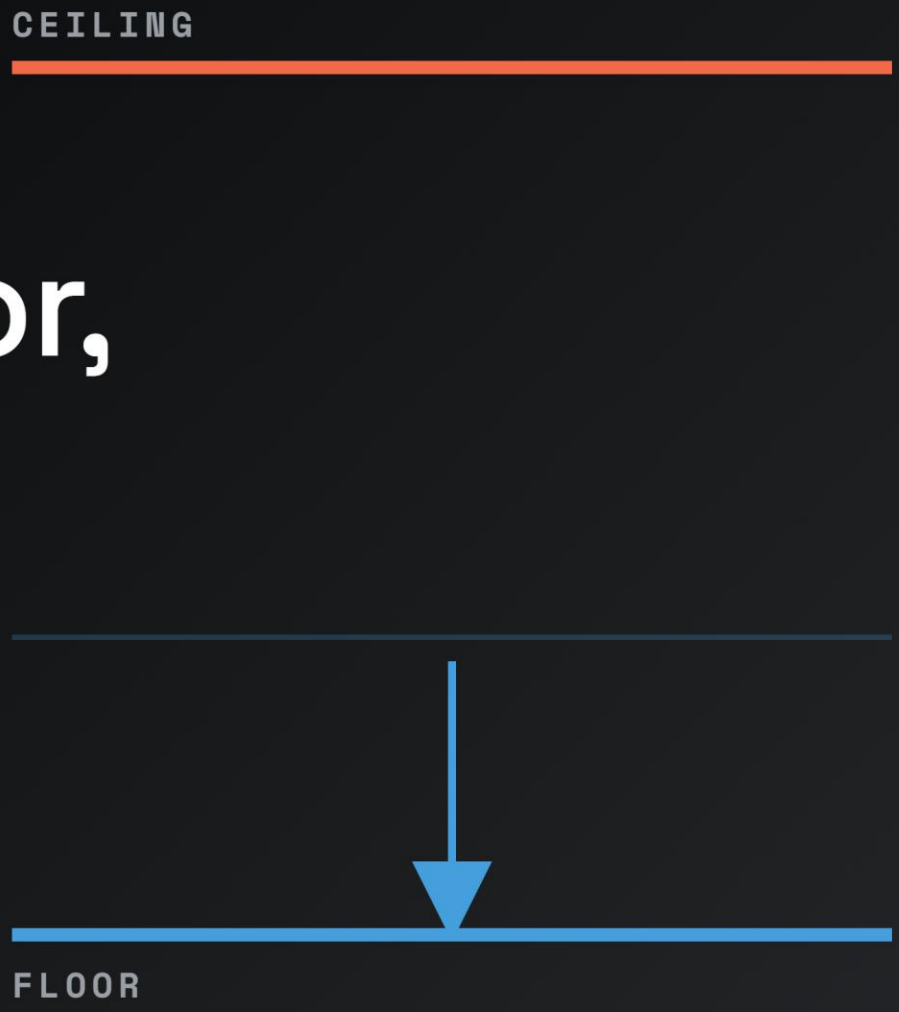
Visible Builders · AI After Hours

AI lowered the floor, our work now is the ceiling

Marylin Alarcón
Lead R&D Engineer, SoftServe



To get started, please scan this code



A

or



B

marylin.me/afterhours/vote

Look at both, then vote A or B

A looked finished.
A loses your data.

TRY IT fill in A → Continue → Back

A



← Back



0 of 3 saved

B



← Back



3 of 3 saved



Marylin Alarcón

Lead R&D Engineer, SoftServe

I ship products that use AI, **built with AI.**

15+
years building software

2018
working with AI in production

PyLadies
Medellín organizer

frontend → backend → full stack → tech lead → **AI**

web marylin.me in linkedin.com/in/marylinalarcon gh github.com/marylin

AI will not replace you.
It will expose you.



Most of my code is written by my agents.
I call them my minions.

THE AI GURUS WILL TELL YOU

"With AI you can build anything in 15 minutes."

TRUE. SO WHAT IS LEFT FOR US?

The ceiling.

Tonight

1

The floor dropped

2

The trap you can't see

3

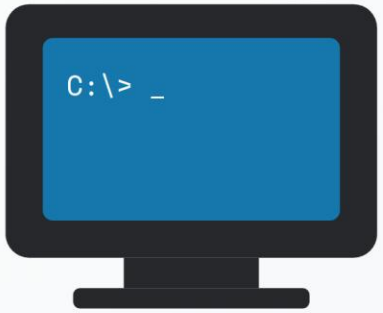
Four decisions that raise the ceiling

guardrails · evals · cost · when not to use an agent

You leave with a checklist for your next AI build.

The floor dropped

1996



One simple app:
months, and a whole team

the floor

today



App A, the one you picked:
15 minutes

Who did AI help more?

Below-average performers

+43%

better work with AI

Top performers

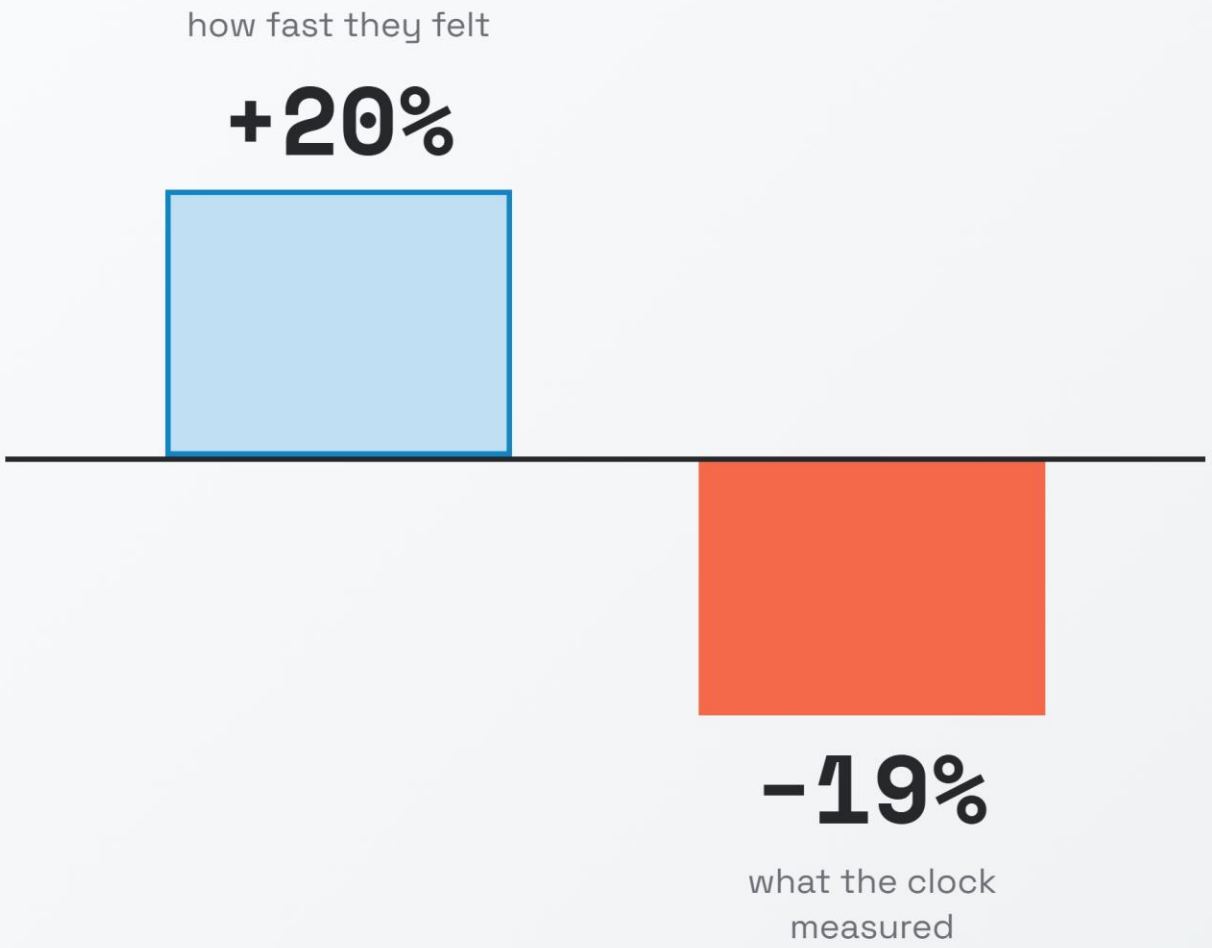
+17%

better work with AI

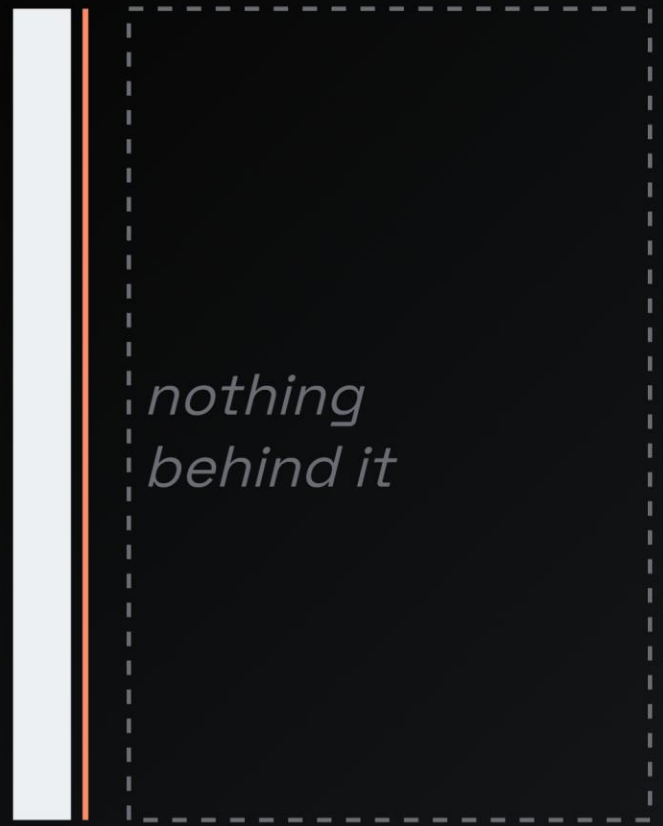
The less you knew, the more AI helped.

758 consultants doing real tasks · Harvard Business School + Boston Consulting Group

It felt faster.
It was slower.



Experienced developers, in their own code · METR, 2025



what you see

The finished illusion

When something looks done,
we stop looking.

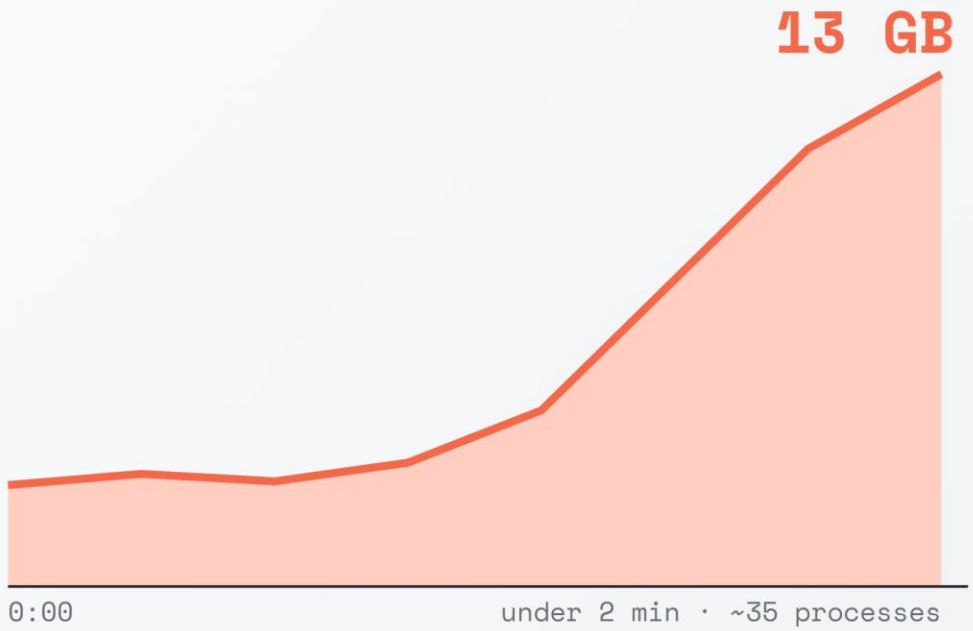
RULE 1 · GUARDRAILS

A rule the AI can ignore is a suggestion

RULE

~~Don't raise the memory limit.~~

written in a prompt



Now: a hard block

checks every command before it runs

RULE 2 · EVALS

I replayed 108 real emails before switching it on

Please stop writing to me.	would send — Thanks for reaching out! Here is your portal link...
I want to cancel. Now.	would send — Happy to help! Here is your portal link...
Leave me alone.	would send — Thanks for reaching out! Here is your portal link...
[insult]	would send — We appreciate your message! Here is your portal link...

Examples reworded, not the real emails

100 of 108
emails would be sent with nobody checking

24 of 25
”stop writing me” emails would get a cheerful reply

0
abusive emails recognized as abuse

THE PROMPT MOST OF US WRITE

Build me the million-dollar app.
Don't make mistakes.

Looks done ✓

October 7 was saved as July 10

THE SAME MILLION-DOLLAR APP, MY WAY The skeleton. The real prompt runs for pages, and these six parts are always there.

GOAL	A booking app that 100 local businesses would pay $20 a month for.	
DONE MEANS	A customer books, pays, and gets a confirmation. Dates stored as YYYY-MM-DD in the customer's timezone.	
DON'T TOUCH	The payment code. The database schema.	guardrail
PROVE IT	Book for November 13 and show me the stored date and the charge.	eval
BUDGET	20 model calls. Tell me before you need more.	cost
STOP AND ASK	If anything is unclear, and before deleting anything.	your call

Is done ✓

November 13 → 2026-11-13

RULE 3 · COST PER INFERENCE

INVOICE

API usage · one agent run **$5,000.00**

Authorized by you **?**

Total due Monday

You made 300 requests I never asked for.

"The logs show 61 submission attempts from me, not 300."

A real case, shared publicly by the user in September 2026. His account of it.

In code, a loop with no limit: **3** calls in the demo → **3,000** with real users

RULE 4 · WHEN NOT TO USE AN AGENT

Sofía, a friend's AI agent

26,000

lines of prompt

JUNE

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

booked: **07/06** (day/month)

JULY

1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

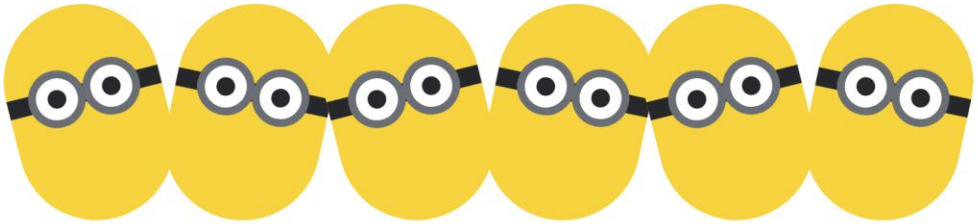
US server read: **07/06** (month/day)

June 7 became July 6.
Nobody was there.

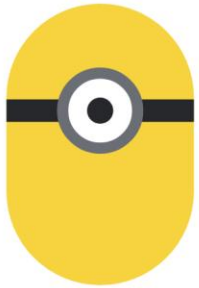
RULE 4 · WHEN NOT TO USE AN AGENT

Everyone does everything

one task



One does everything

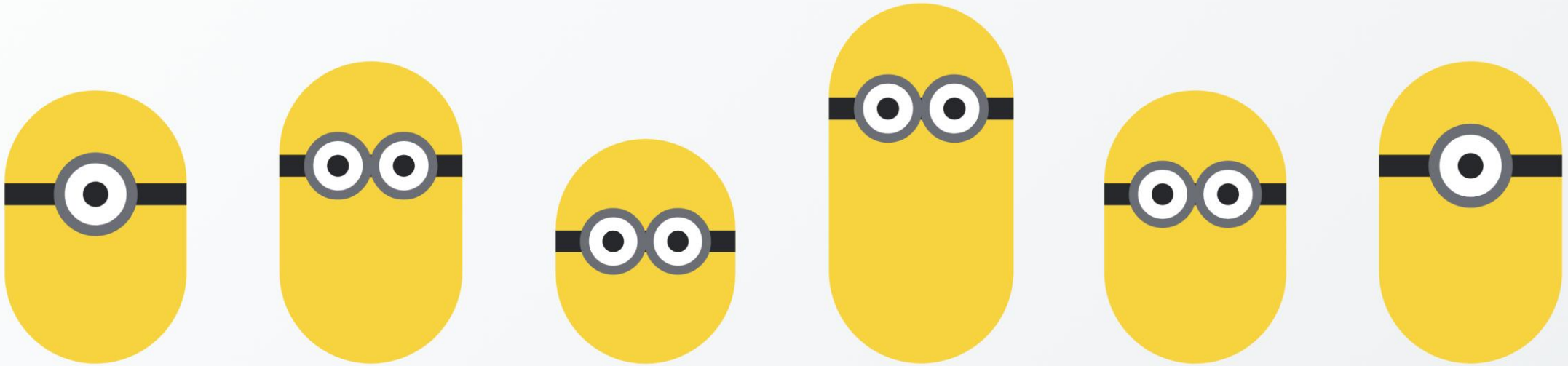


leads · WhatsApp · listings · exchange rate · visits · reminders · follow-ups

If it can be a function, don’t use an agent.

My minions

Me · the setlist → **Lead agent** → one job each



Stuart
`interface`

Dave
`logic`

Bob
`data`

Kevin
`code review`

Jerry
`testing`

Carl
`error resolution`

1 The floor dropped

You saw it: app A, 15 minutes.

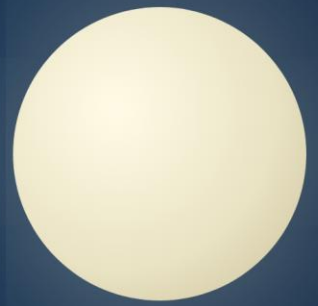
2 The trap is invisible

You lived it: almost everyone picked A.

3 The ceiling is yours

Taste, rules, tests, and deciding when not to.

THE PROMPT MOST OF US WRITE



”Minions, tonight we steal the moon.”

GRU

THE SAME JOB, MY WAY

GOAL

Bring the moon home tonight.

The whole moon, in one piece.

DONE MEANS

Moon in hand. Nobody hurt.

Everyone back before sunrise.

TASTE

Look great doing it.

Gru's version: "I don't know what I'm doing, but I look great doing it."

GUARDRAIL · DON'T TOUCH

The shrink ray points at the moon.

Only the moon. Never at a minion.

EVAL · PROVE IT

Shrink something small first.

Show me it worked before the real shot.

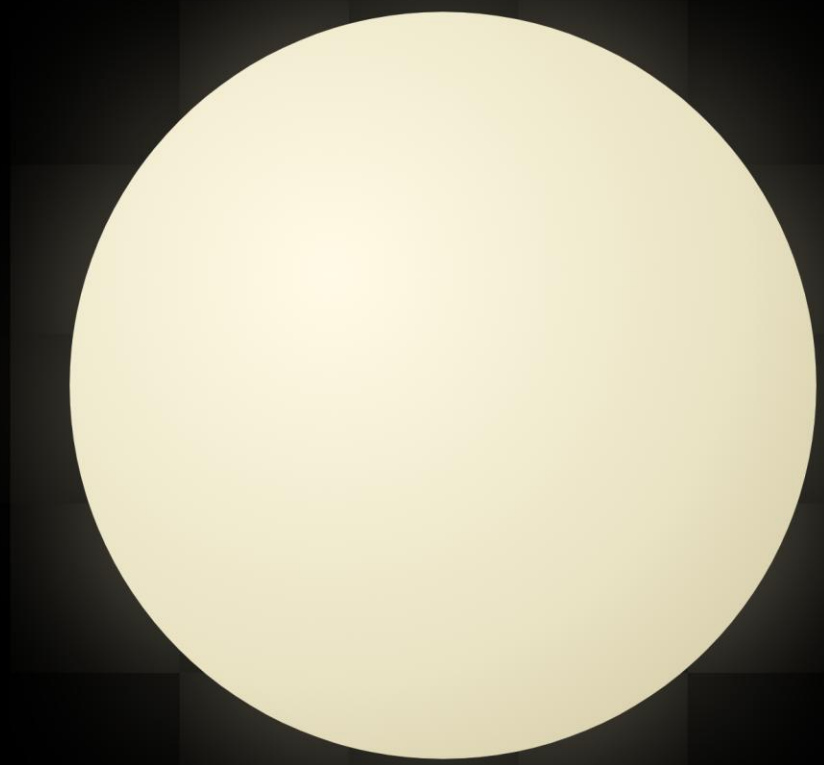
COST · BUDGET

One loan from the Bank of Evil.

Every shot costs. Ask before the second one.

NO AGENT · NOT FOR MINIONS

The bedtime story. I read that one myself.



We stole the moon.

¡Muchas gracias!

Now ask my minion:

- the slides
- useful links
- 15 minutes with me, one on one

An AI answers right away. I answer... eventually.
Feeling brave? Ask it for 18,000 glasses of water.

WhatsApp · +1 208 605 7990