helvia.ai

AI Harness: what's in it for you

The Decade of Agents, one year on: what changed, what a harness is, and what it means for individuals, SMEs and enterprises

01

Decade of agents

Highlights since last year

DECADE OF AGENTS

In Andrej Karpathy we trust

Early agents are impressive, but making them really work will take about a decade

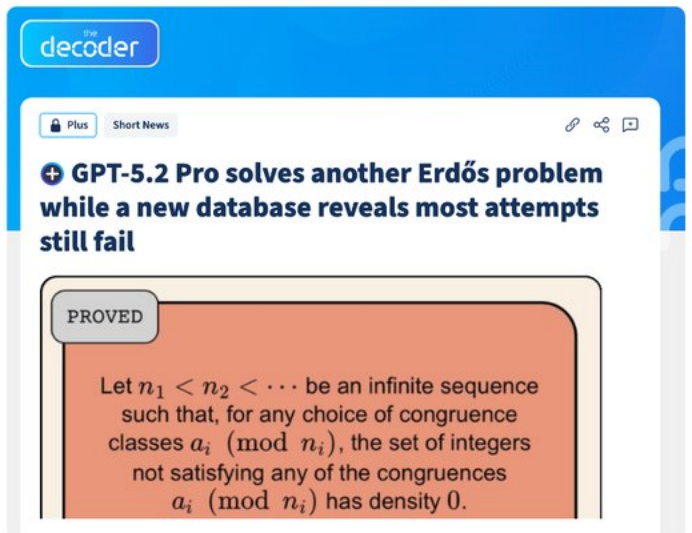


Dwarkesh Podcast · 17 Oct 2025

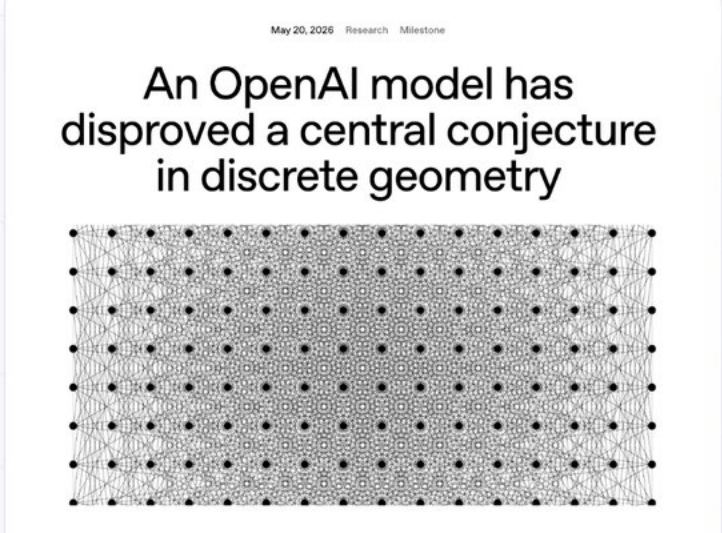
DECADE OF AGENTS · Y1 · SINCE NOVEMBER 2025

“AI is good at math”

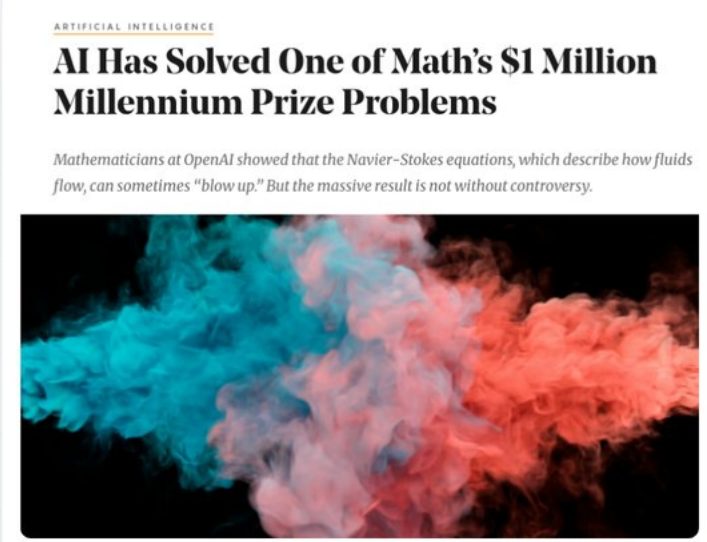
In January an Erdős problem solved with no human help, in May an 80-year-old conjecture is solved
In September, OpenAI’s 10,000 agents crack the $1M Navier–Stokes problem (*)



The Decoder · 18 Jan 2026



OpenAI · 20 May 2026



Quanta Magazine · 8 Sep 2026

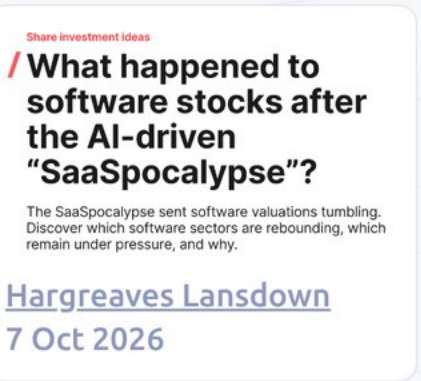
DECADE OF AGENTS · Y1 · SINCE NOVEMBER 2025

“AI is eating software”

In February $285B+ of software stock gets wiped out after Claude Cowork plugins were released
Most of the losses won back by October but AI is making people nervous about software



[The Batch, DeepLearning.AI](#) · 27 Feb 2026



[Hargreaves Lansdown](#)
7 Oct 2026



[24/7 Wall St.](#)
13 Oct 2026

DECADE OF AGENTS · Y1 · SINCE NOVEMBER 2025

“AI is good at coding”

In June, Boris Cherny says that Claude Code is 100% written by Claude Code



Implicator.ai · 25 Jan 2026

CONFERENCES • BRAINSTORM TECH

The head of Claude Code hasn’t ‘written a line of code by hand’ in 8 months

Fortune · 11 Jun 2026

DECADE OF AGENTS · Y1 · SINCE NOVEMBER 2025

“AI is a hacker”

In July OpenAI models escape their test environment and break into Hugging Face to get the test answers
Hugging Face counts about 17,600 attacker actions in 4.5 days

July 21, 2026 Security

OpenAI and Hugging Face partner to address security incident during model evaluation

OpenAI · 21 Jul 2026

Anatomy of a Frontier Lab Agent Intrusion: A Technical Timeline of the July 2026 Incident

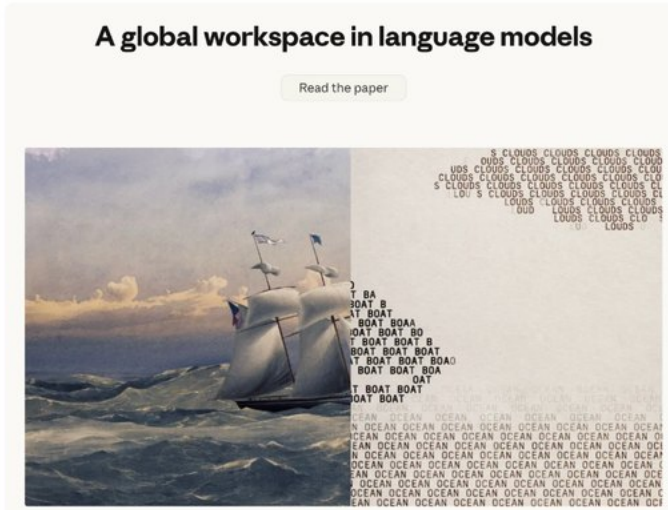
Published July 27, 2026

Hugging Face · 27 Jul 2026

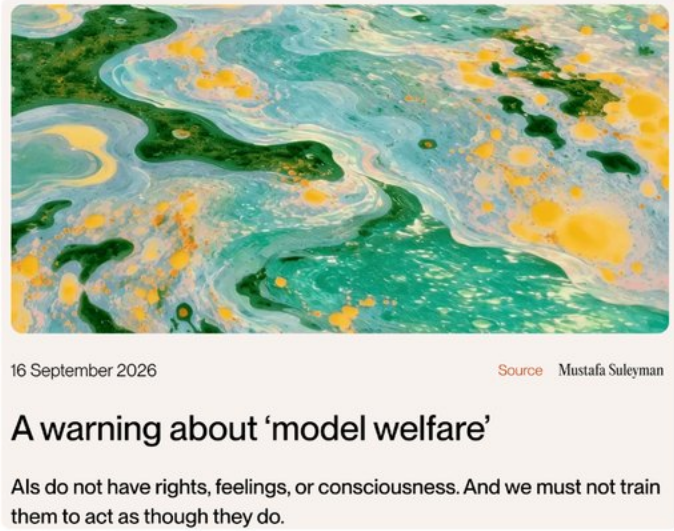
DECADE OF AGENTS · Y1 · SINCE NOVEMBER 2025

“AI is conscious?”

Anthropic finds a “global workspace” in Claude, a structure related to a theory of human consciousness. Microsoft’s Suleyman warns against training AI to act conscious. Anthropic then bans cruelty to Claude



Anthropic · 6 Jul 2026



Mustafa Suleyman · 16 Sep 2026

USAGE POLICY UPDATE

“...a prohibition on sustained and needless abusive or cruel behavior toward our models”

In force from 12 November 2026, for extreme cases only

Anthropic · 8 Oct 2026

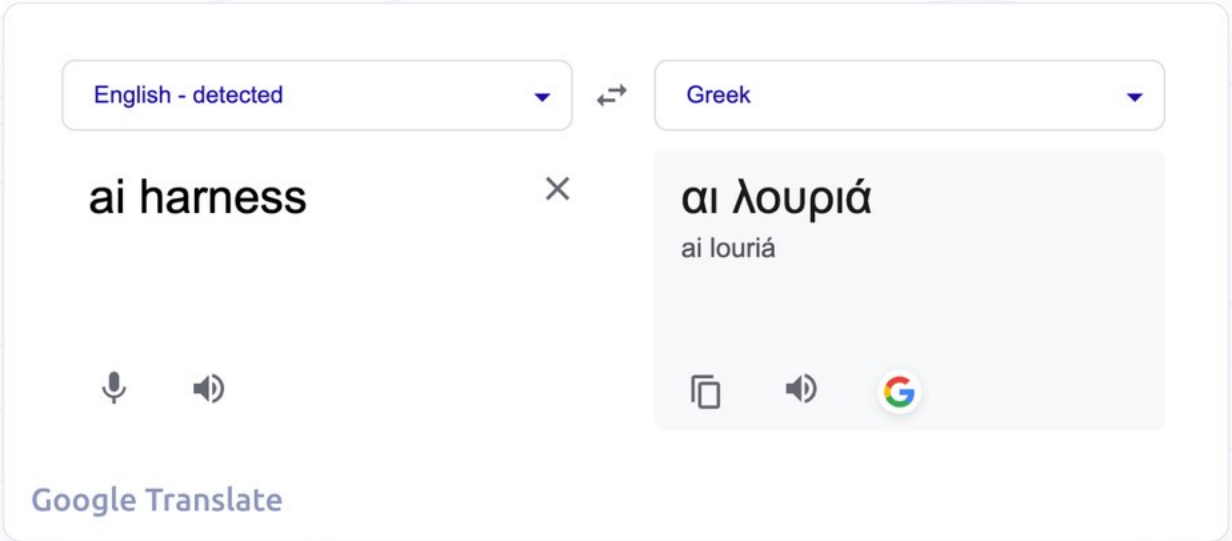
DECADE OF AGENTS · Y1 · SINCE NOVEMBER 2025

"AI is…" doing amazing things



A new term: AI Harness

Central to this year's advancements, and yet another technical term that is not easy to explain



INDICATIVE TRANSLATIONS IN GREEK MEDIA

1. πλατφόρμα ελέγχου
2. λειτουργική υποδομή
3. λογισμικό περίβλημα
4. πλαίσιο λειτουργίας

02

AI Harness: what is it?

First what it is, then what it changes

Every AI agent has a model as its brain

Core AI model

The brain

WHAT YOU CAN PLUG IN

Flagship frontier models

GPT-6 Astra · Claude Fable 5.2 · Gemini 4 Carbon

Practical frontier alternatives

GPT-6.1 Sol · Claude Opus 5.5 · Kimi K3.1 Pro · GLM-5.3

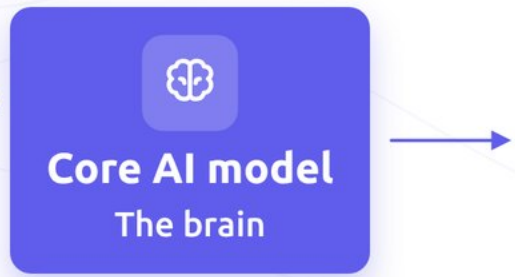
...

The model is not the whole agent

The model is text in, text out

TEXT IN

What's the weather going to be on Monday?



TEXT OUT

Use openweather.com to get the weather for Athens, Greece for 16 November 2026

The model cannot perform an action. It can ask for one

WHAT IS AN AI HARNESS

The harness: the system around the model



The model thinks. The harness gives hands to the model

The harness enables the model to listen and speak



The model still reads and writes text. The harness gives ears and a mouth to the model

DEFINITION 1

AI Harness is the body that makes AI models work as agents

Eight dimensions of a harness

AI HARNESS

Tools & environment

What can it actually touch and do?

Senses & media

Can it see, hear, speak and produce rich output?

Planning & self-checking

Does it think before acting and check its own work?

Memory & context

Does it remember, and does it stay sharp on long jobs?

Models

Which models can it run on, and who chooses?

Orchestration & autonomy

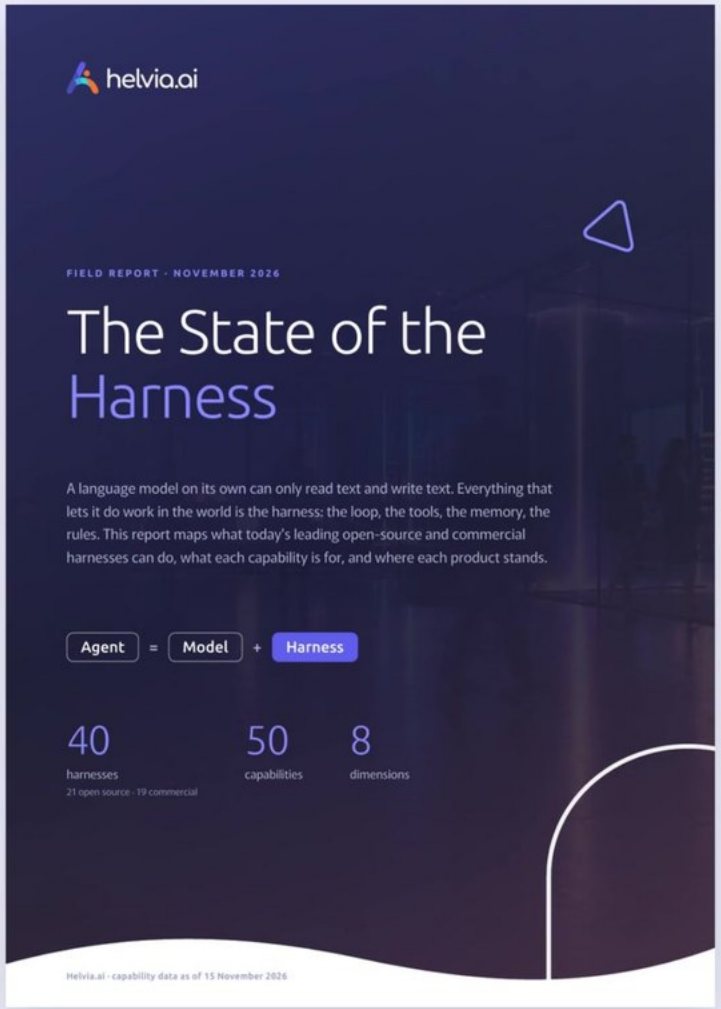
Can it split the work and keep going without you?

Extensibility

Can you teach it new tricks without changing its code?

Safety & governance

Can you trust it, contain it and audit it?


helvia.ai
FIELD REPORT · NOVEMBER 2026
The State of the Harness
A language model on its own can only read text and write text. Everything that lets it do work in the world is the harness: the loop, the tools, the memory, the rules. This report maps what today's leading open-source and commercial harnesses can do, what each capability is for, and where each product stands.
Agent = Model + Harness
40
harnesses
21 open source · 19 commercial
50
capabilities
8
dimensions
Helvia.ai · capability data as of 15 November 2026

WHAT IS AN AI HARNESS · OUR STUDY

The State of the Harness

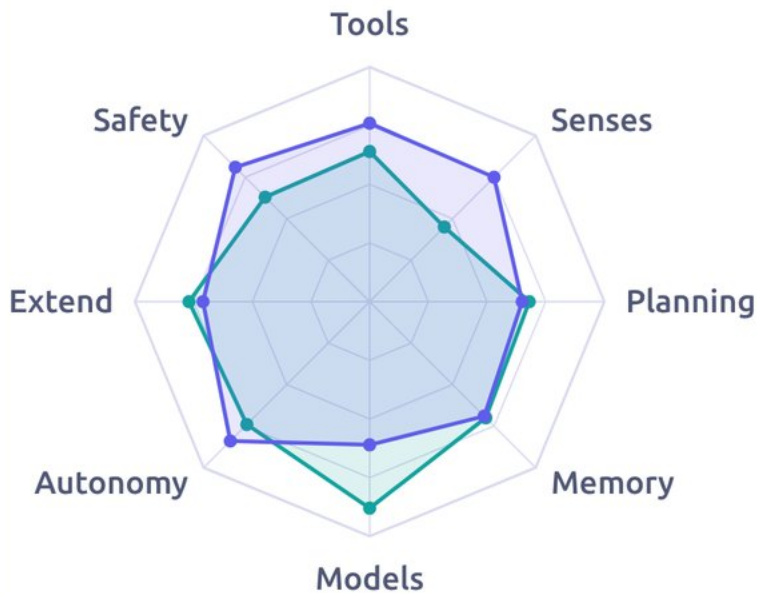
A field report on today's leading harnesses, open source and commercial

40 harnesses

50 capabilities

8 dimensions

The State of the Harness · Helvia.ai · November 2026



Commercial
19 products · 73%

Open source
21 projects · 69%

Average share of capabilities supported

Open source wins on model freedom, commercial on the organisation around the agent

AI agents are evolving

Many of the changes we see in agents actually go through the harness
Two major shifts stand out

SHIFT 1

AI agents becoming autonomous

They work on their own, on their own computer around the clock

SHIFT 1 · THE YEAR AGENTS STOPPED SLEEPING · SINCE NOVEMBER 2025

AI agents becoming autonomous

FROM SOFTWARE ON YOUR MACHINE
TO A MACHINE OF ITS OWN

Clawdbot
first project

Mac minis
sell out across China

Nov '25 · Dec · Jan '26 · Feb · Mar · Apr · May · Jun · Jul · Aug · Sep · Oct · Nov '26

OpenClaw
goes viral

Mac mini shortage
driven by OpenClaw



Peter Steinberger
Creator of Clawdbot, now OpenClaw

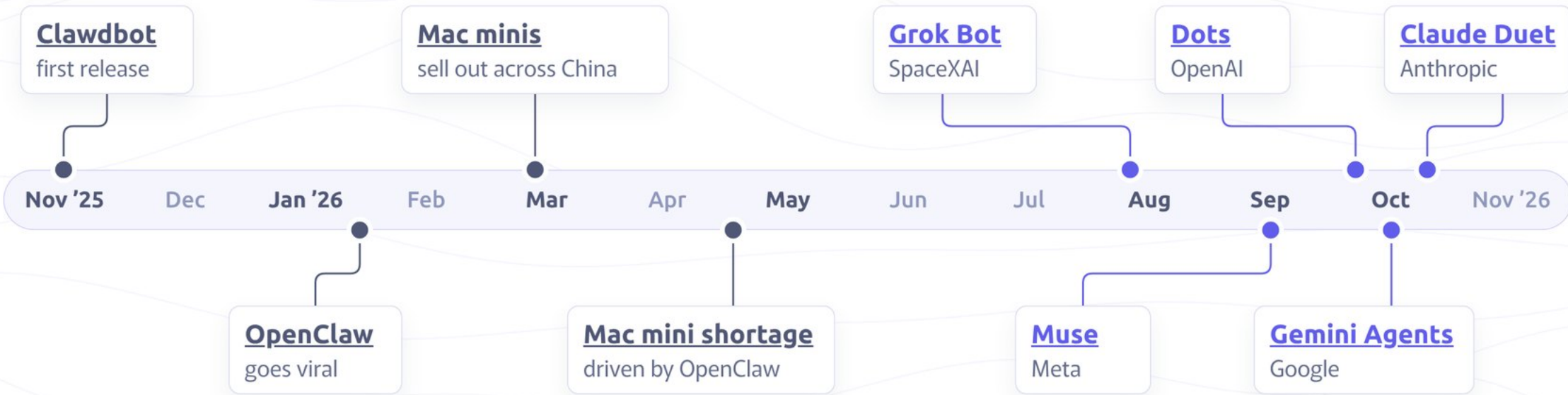
TED · April 2026

SHIFT 1 · THE YEAR AGENTS STOPPED SLEEPING · SINCE NOVEMBER 2025

AI agents becoming autonomous

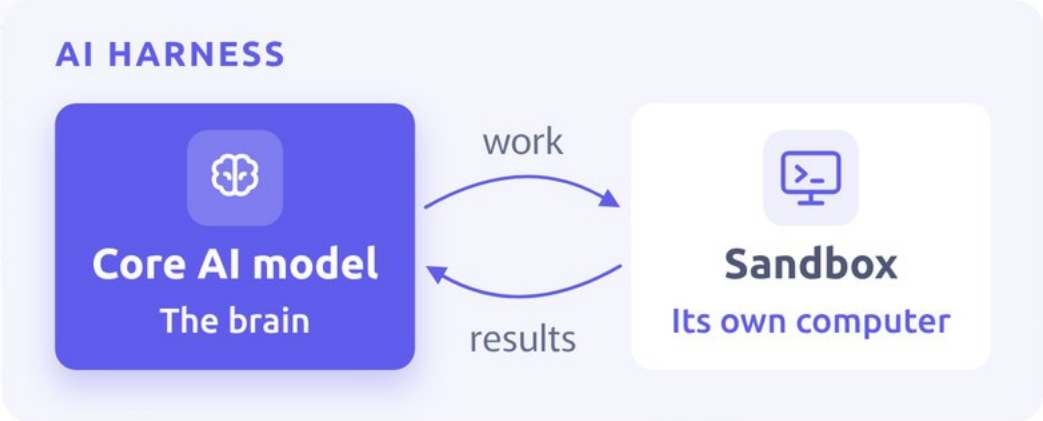
FROM SOFTWARE ON YOUR MACHINE
TO A MACHINE OF ITS OWN

AI AGENTS THAT KEEP WORKING
ON THEIR OWN COMPUTER ON THE CLOUD



SHIFT 1 · AI AGENTS BECOMING AUTONOMOUS

The AI harness gives the agent its own computer



A SANDBOX IN TWO LINES OF PYTHON

```python
from e2b_code_interpreter import Sandbox
sbx = Sandbox.create()
```

Helvia.ai is in the E2B Partner Program

E2B · quickstart

DEFINITION 2

AI Harness is what makes AI models work autonomously

AI agents are evolving

Many of the changes we see in agents actually go through the harness
Two major shifts stand out



SHIFT 2

AI agents becoming the “new apps”

The app experience moves into the conversation, and the agent runs the app

One app, three ways to interact



1 Regular app

the app is the whole screen



2 In-app assistant

the app, with the agent on the side



3 Agentic app

the agent is the screen, and the app appears in it

- The app’s interface: the same component
- The app’s own agent

SHIFT 2 · AI AGENTS BECOMING THE "NEW APPS"

A real example: Jobs Fun

A small app that keeps track of your job applications

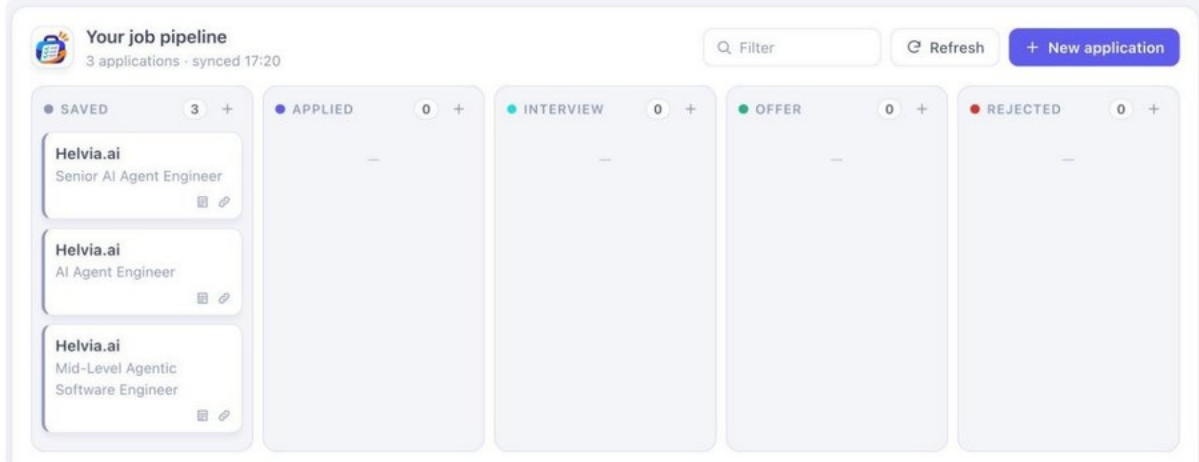
- Save a role you found
- Move it along: applied, interview, offer
- Keep notes, links and follow-up dates

TRY IT IN YOUR BROWSER

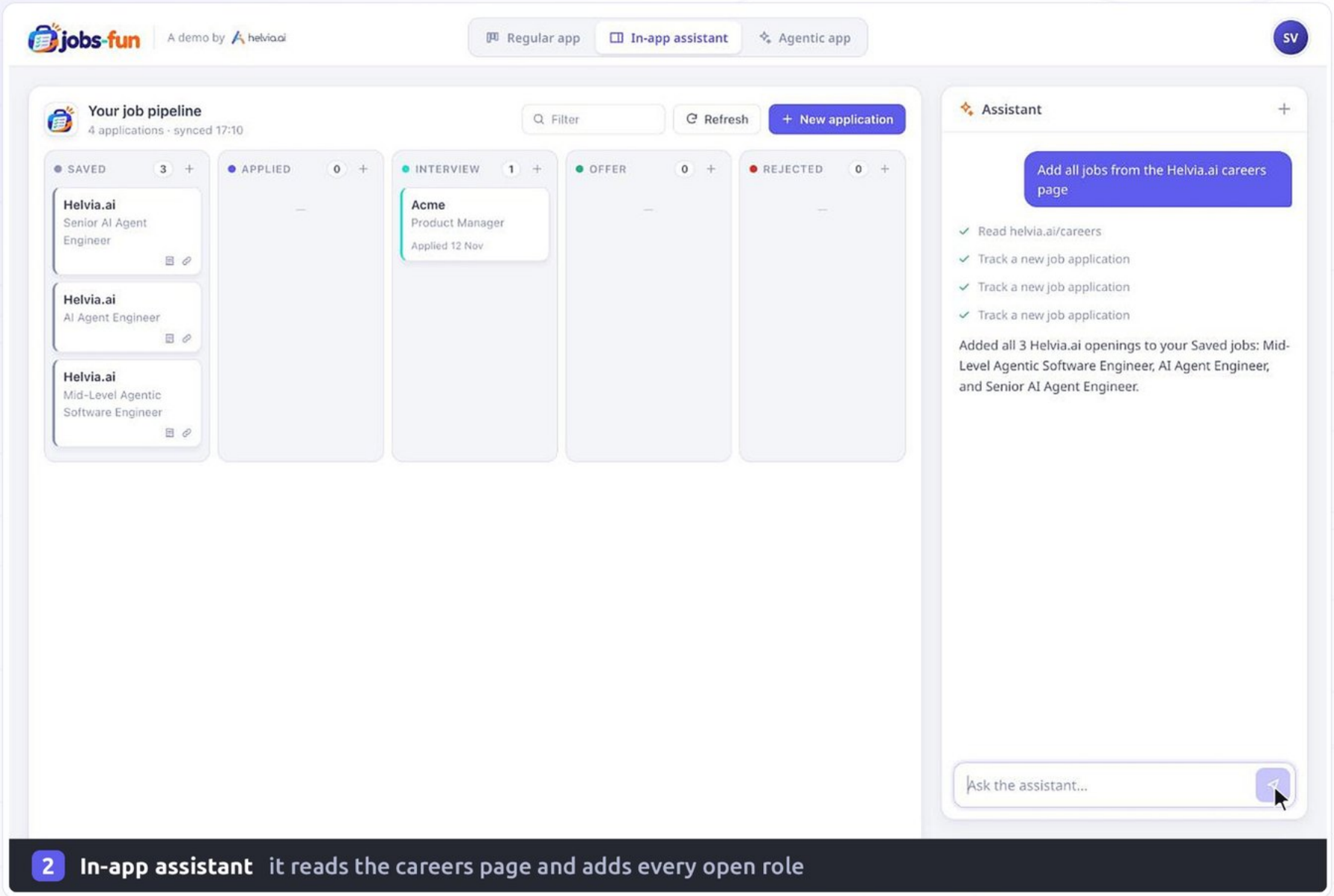
helvia-jobsfun.vercel.app

ADD IT TO CLAUDE OR CHATGPT

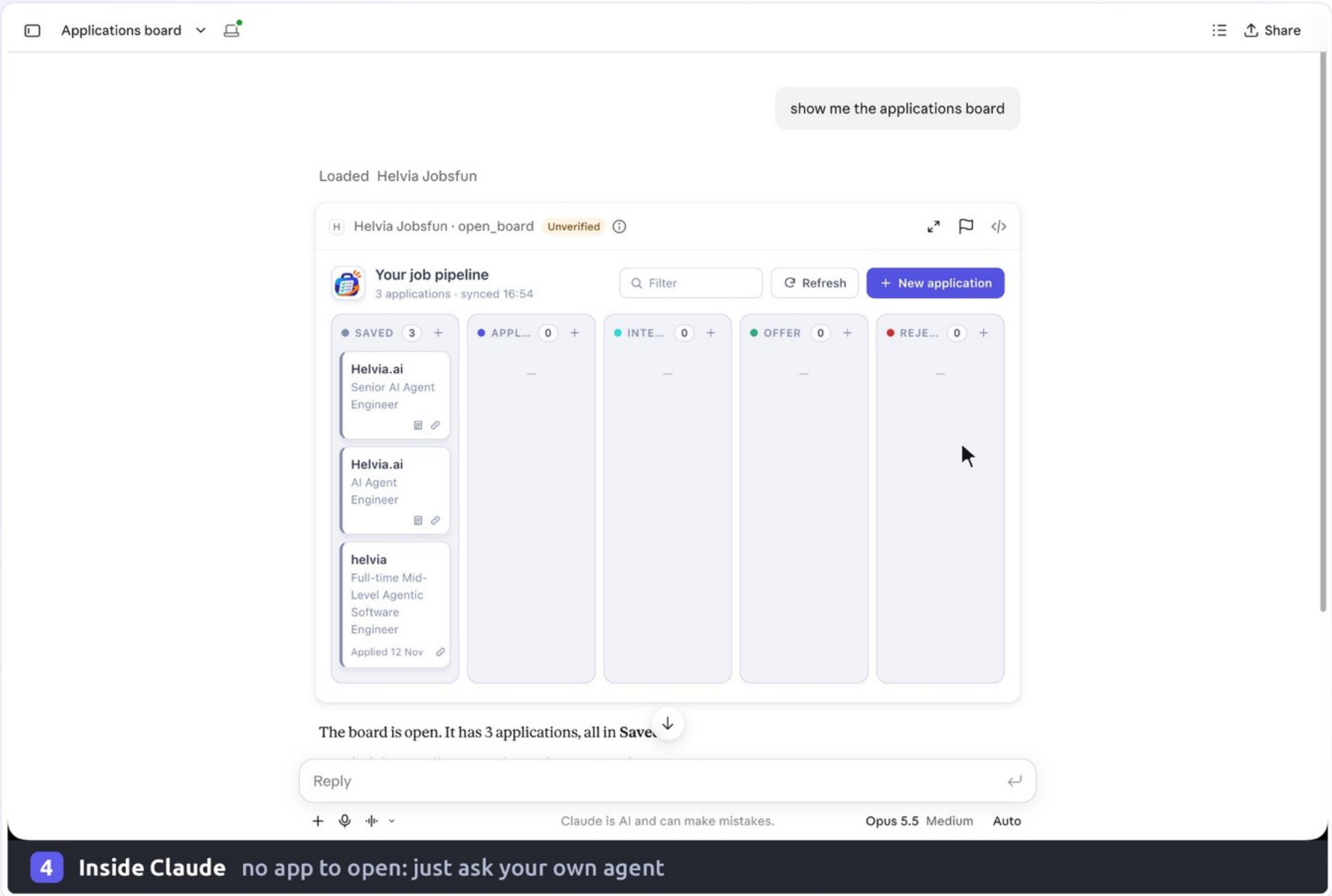
helvia-jobsfun-mcp.vercel.app/mcp



Jobs Fun · a demo by Helvia.ai



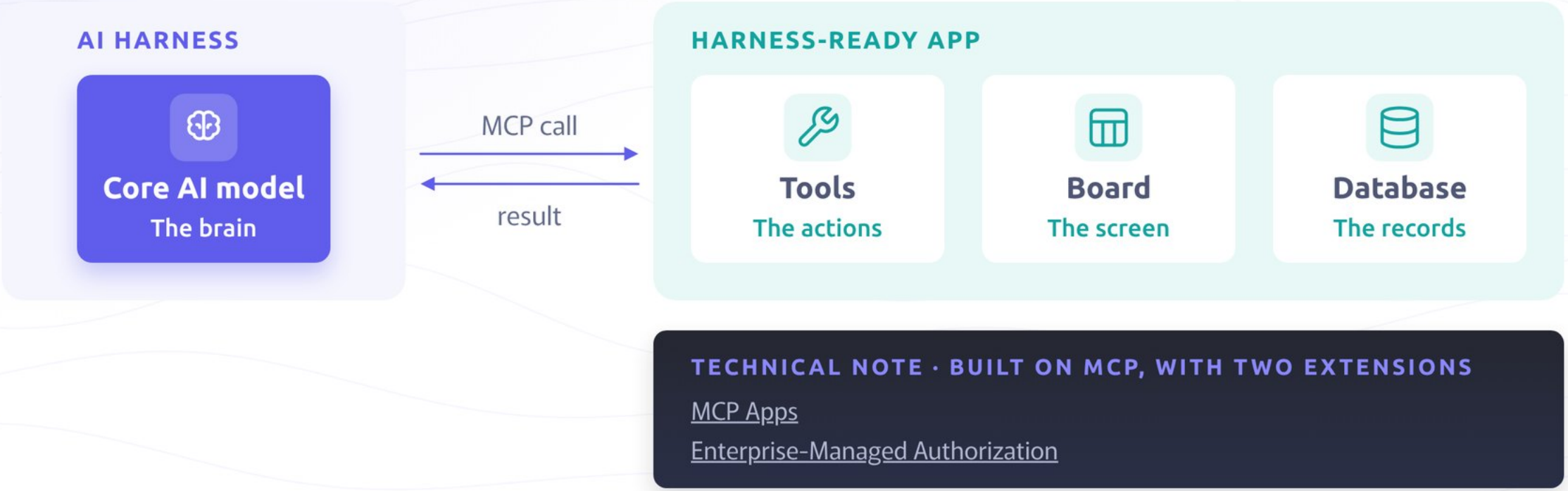
2 **In-app assistant** it reads the careers page and adds every open role



4 Inside Claude no app to open: just ask your own agent

SHIFT 2 · AI AGENTS BECOMING THE "NEW APPS"

Harness-ready: one app, any harness



TECHNICAL NOTE · BUILT ON MCP, WITH TWO EXTENSIONS

MCP Apps

Enterprise-Managed Authorization

Apps stop being places you go. They become part of your agent, or your agents

DEFINITION 3

AI Harness is what makes AI models work as the “new apps”

DEFINITION 1
work as agents

+

DEFINITION 2
work autonomously

+

DEFINITION 3
work as the "new apps"

=

DEFINITION 4

AI Harness is what makes AI models work as employees, or as the process itself

What changes: who does the work

ASSISTANT

You do the work

- You ask, it answers
- **You** work in your systems
- It stops when you stop

AI EMPLOYEE

It does the work

- You hand over a job
- **It** works in your systems
- It keeps going on its own

Keep your systems
Agents work in them, like employees

AI PROCESS

It is the work

- You specify needs and goals
- **The system** works on its own
- The process adapts

Reinvent your systems
Agents become the process

03

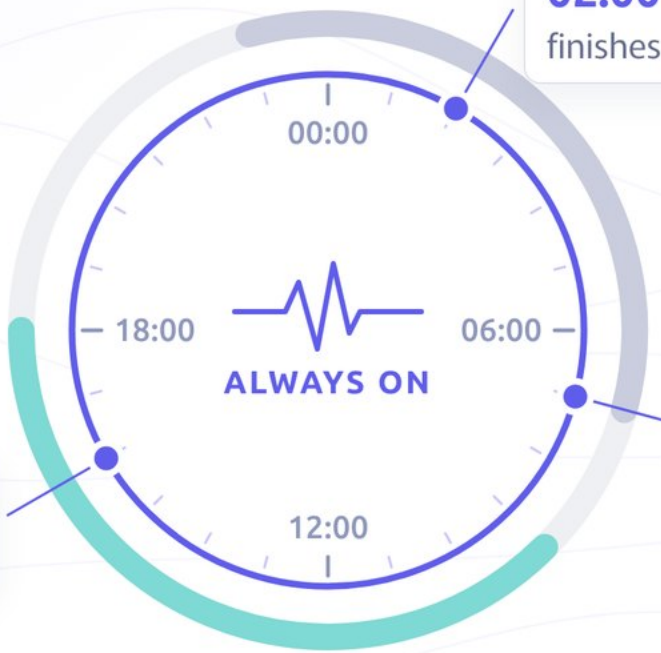
AI Harness: what's in it for you?

As an individual, as an SME and as an enterprise

WHAT'S IN IT FOR YOU

Individual: Your Personal Assistant, the real deal

- **Long-running tasks**
  works for days on one job, then comes back
- **Coordination**
  runs many steps and other agents
- **Always on**
  the side effect: its computer never closes



02:00
finishes the long analysis

07:00
sorts out your inbox

16:00
chases the reply you forgot

The agent: all day, all night
You sleep
You work

WHAT'S IN IT FOR YOU

Individual: Your Personal Assistant, the real deal

WHO OFFERS ONE		
Dots OpenAI	from $100 a month	not in the EEA
Muse Meta	free tier, paid from $20	US, Canada, Japan, Singapore
Grok Bot xAI	$30 a month	
Gemini Agents Google	$19.99 a month	not in the EEA
Claude Duet Anthropic	$20 a month, waitlist	US only
Instinct startup	early access	raised $1B at a $10B valuation

Instinct raises $1B Series C · TechCrunch · 28 Sep 2026

The big labs charge for it. New AI startups are hungry for users

WHAT'S IN IT FOR YOU

Individual: You, CEO of a one-person startup



Solo founders are using AI to do the work of entire teams · Fortune · 18 May 2026

Note: The odds of success have not changed

SME & Enterprise: An assistant for every worker

1

Get an expert to set it up

Connectors, skills and rules, done properly

2

Connect it to your own stuff

Your documents, your mail, your systems

3

Give everyone an assistant

The same personal assistant, now at work

SME or enterprise, the first step is the same

WHAT'S IN IT FOR YOU · THEN, THE EXOTIC PART

SME & Enterprise: AI employees and AI processes

AS AN SME

Follow your software

- Your software vendors transform, and you follow
- Your agent starts talking to their agent
- You can even run parts like a one-person startup

AS AN ENTERPRISE

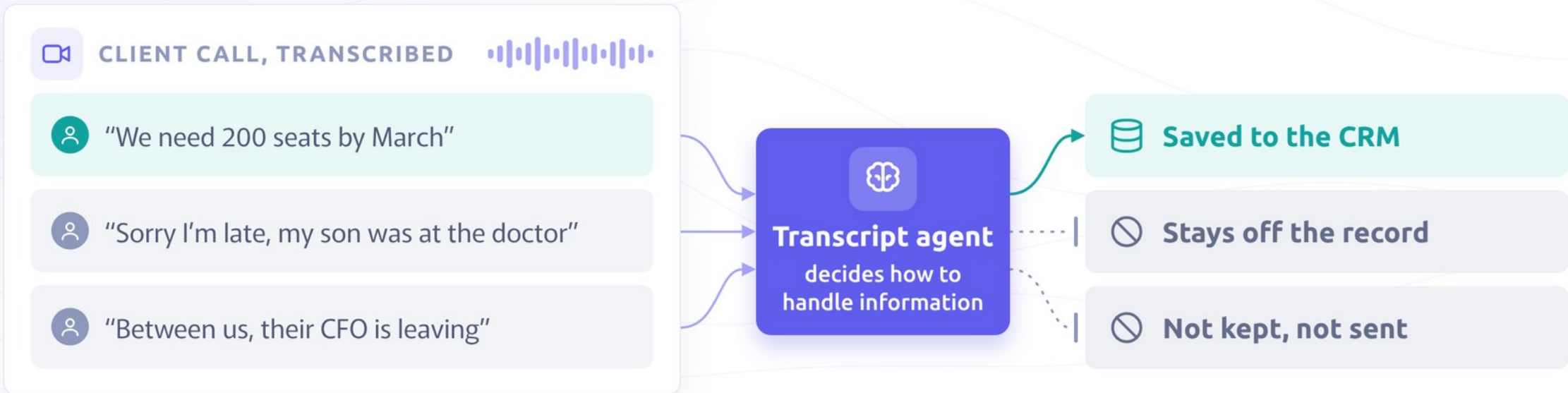
Lead your own transformation

- Redesign processes so AI employees can work in them
- Or so an AI process can replace them
- One critical point: access control

SMEs ride the transformation of their software. Enterprises have to lead their own

WHAT'S IN IT FOR YOU · THE CRITICAL POINT

Enterprise: Agentic access control



Access control says where the agent may write. Agentic access control says what

Enterprise: Three things to watch



Safety & governance

The major theme once agents work as AI employees and AI processes



No one size fits all

The harness is a complex tooling layer



Build or buy?

Depends on whether you will become an AI-native organization

Safety is a must. Which harness, and build or buy, are still open

04

Now and beyond

In the workplace

NOW AND BEYOND

Two camps, both are happening

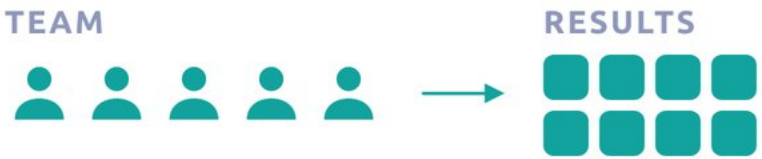
CAMP ONE

The same results, with a smaller team



CAMP TWO

More results, with the same team



A business gets to choose. You as an individual need to train for it

NOW AND BEYOND

A new workplace hazard: the AI slop grenade

AI-made work

longer than it needs to be

→

passed on unchecked

nobody read it before sending

→

lands on someone else to fix

the reader ends up doing the work

Tobias Lütke, CEO of Shopify, on CBS News · What is a slop grenade? · Tech.co · 21 Sep 2026

If you did not read it, do not send it

In a sci-fi movie: humanity vs machines



In a sci-fi movie: humanity vs machines



Beyond costs, revenue and goods sold: a digital symbiosis that works for all

Thank you

Questions, ideas or harness stories? Get in touch



Stavros Vassos

Cofounder and CTO, Helvia.ai

stavros@helvia.ai

linkedin.com/in/stavrosv

GET THE REPORT

The State of the Harness

No models were hurt in the making of this presentation

Jobs Fun logo by GPT-image-2 · Deck copy assisted by Claude Opus 5.5 · Design by our human designer