

Fixing AI & creating accidental consultants.

Accuracy for AI agents, companies, groups & people.

QUICK CHEAT SHEET 1 PAGE SUMMARY

YOU ASK:

You, or your group, get cheap accurate answers & maybe earn \$

YOU ANSWER:

Take on challenges from AI Labs, peers, debates & games

YOU SOLVE:

Know which AI answer or agent action is right before you act, fact check & get a multi-AI receipt, programming problems

YOUR 21 AI STAFF:

Your advisor, always at your side when you need something, controls your current staff of 21 AIs who use your knowledge base in your notebook to organize your life & find opportunities.

WHO PAYS

People

Free · \$5 · Plus \$9 · Pro \$19/mo

Companies

Team plan \$299/mo · APIs for your apps & AI agents with safe-to-act checks · browser extension for quick adoption · consulting jobs to our network of experts · brand feedback

Forums & hobby groups

1st group live, from \$10/mo

AI labs

Human correction data

\$100Ms expected, paid per 👍/👎 & correction

Companies, forums & hobby groups may outgrow individuals. Where to push first is your call. Branding is yours also. ReallySolved.com integrates a lighter version of every funnel (separate funnel sites are described on page 10).

PROBLEMS SOLVED @launch

- **Wrong answers on \$M decisions are expensive.** 1 AI on its own was off by \$1.34 million on the Section 179 tax limit.
- **Within 2 points of Claude Opus 5 in accuracy, at 1/40th the cost.**
- **We've created a ready to launch ecosystem (15 funnels for companies, forums & hobby groups, & individuals) to collect human correction data & get people consulting gigs as a byproduct.**
- **People can get paid 2 ways doing what they already do today.** They'll 👍👎 AI results & stop giving away their knowledge for free on Reddit, Twitter etc.
- Sam Altman, Elon Musk & Dario Amodei all say AI is moving too fast. **Mass multi-channel crowdsourcing is the only way to keep up.**

MOAT

- **21 questions about AI mistakes we don't think anyone else has fully solved, & we've answered every one.**
- **We've already tested 169 AIs mapped to ~500 (& growing) Accuracy & Safety Topics so we know which ones to ask about your subject before we ask. This giant matrix (some of it filled out by research) gets more accurate over time.**
- **We catch the AI mistakes the labs don't know to look for.**
- **If we can achieve network effect, a competitor with big money can't compete.**

LOW RISK

- **Tell us what your terms would be** to take over the promotion & running of the websites. Our current idea is: take 100% of quick ramp-up net revenues until you've recouped 3x your time & costs + equity & ongoing profit sharing from AI lab revenue & all other revenue.
- **Founded & funded by a former VC, not raising money & tiny burn.**
- **~80-92% margins after the AI & card fees, & our cost per answer goes down as we grow.**
- **Consumers already pay for a far thinner version of this.** The biggest app that just shows you several AIs earns over \$100M a year at ~\$30 a month, & not 1 of them settles the disagreement.
- **Our conservative bottom-up, slow ramp-up, analysis assigns a very low value to our correction data & still shows \$100Ms expected from AI labs**

Just 1 of many examples. 3 AIs. 3 wrong answers.

We asked what you are allowed to put in a 401(k) this year.

\$22,500

×

\$23,000

×

\$23,000

×

The real answer
\$24,500 for 2026

5 of the 7 AIs

we tested got 5% or less right, once the answer changed. We fix that.

On the questions people really ask, the best free AI we tested scores 94.4% & the worst 70.4%. A 24-point gap, & the person asking has no way of knowing which one they got.

Why the moat is **hard to copy**.

Every answer they correct & every note they add makes their own staff better at their job. That is the reason they come back, & the reason leaving costs them something.

The AI companies say the real advantage is not the model, it is the loop of people teaching it. That's exactly what our users who choose to earn money for correcting mistakes that frustrate them every day are doing.

A wrong answer,
User 👍👎 or types
a short correction.

→ That topic's score
updates →

~500 topics
& counting
changes in real time

Reddit, Microsoft et al collect the argument. **We provide the accurate answer.**

Why smarter AI doesn't **replace us.**

ACCOUNTABILITY FRESHNESS THINGS THAT WERE NEVER WRITTEN DOWN

WE KNOW WHO CREATED WHAT

"Why can't they just look it up?"

Where there's a clean source, the right answer is to go read the source, & the AIs that search the web live already do that. **Where we're worth something is the huge middle:** the rule is published but ambiguous, the official page hasn't caught up, the manual says one thing & everyone who does the job knows another, or the question is about what actually happens rather than what's written down. **That middle is most of what people ask, & lots of searching rarely finds.**

COMPANY PROBLEMS WE SOLVE

For companies: **your team's AI answers, checked before anyone acts on them.**

We estimate about **\$5,000 a year per employee** checking what AI told them.*

- **Out-of-date answers caught.** Tax rules, customs limits, HR & benefits: we tell the team the moment something they looked up stops being true.
- **A record of every answer**, with how it was settled & when. **The company decides if any of it is sold.**
- **\$299/mo for up to 15 people · 1 invoice · +\$15/mo for each extra person.** The forecast counts 5, 20, 100 & 500 companies at each audience size. Staff earn nothing on the seat; they can open their own personal account.

1 question. No job post, no bids - just 3 minutes in exchange for lunch money.

(maybe a VERY nice lunch)

To buy 1 answer today

5 steps

- 1 Post the job
- 2 Sift the bids
- 3 Check credentials nobody checked
- 4 Pay for 40 hours to answer 1 thing
- 5 Hope they were right

To buy 1 answer here

1 step

- 1 Ask it. They set the price.

Nobody sells 1 answer from verifiable experts who have deep domain knowledge on very specific granular topics, but not looking for consulting jobs. Right now they're often giving away this knowledge online - but now can almost effortlessly share it & get matched to opportunities. [More info](#)

* Staff time spent checking, plus what the mistakes nobody caught cost. We counted those mistakes at about a sixth of the usual rate, to keep the figure low.

Why this is buildable now, & was not 3 years ago.

6 things changed. Every 1 of them is somebody else's published finding, not ours.

1. THE ANSWER MOVES, & SOMEBODY PROVED IT
2. THE RULES CHANGE & NOTHING IS ANNOUNCED
3. THE EXAMS ARE LOSING THEIR GRIP
4. SOMEBODY NOW PAYS REAL MONEY FOR CORRECTIONS
5. SOMETIMES THE AI JUST GOES QUIET
6. WINDOWS NOW PICKS WHICH AI DOES EACH JOB. IT CANNOT TELL WHICH ANSWER IS RIGHT.

3 years ago there was no published proof an AI's answers drift, no way to see a rule change, & no buyer for the correction. All 3 exist now. The piece still missing is the crowd that does the correcting, & that is the piece we are asking you to grow.

We arrived at numbers that were wrong. 5 times. We caught every one ourselves.

This is the page most decks do not have. It is here because we promised we would show you the tests that went against us, & because a company selling accuracy that quietly buries its own bad runs is exactly the thing we are trying to replace.

- Our own question set was too easy, so we threw out our own numbers.
- Then the grading of that outside run turned out to be wrong too, & all in one direction.
- Which is why exactly 1 number from that run appears in this deck, instead of a table of them.
- And the real limit of our own method, which survives the correction: some questions defeat every engine at once, & asking more of them does not fix it.
- And the newest one: our confidence number sometimes said it was certain when it was wrong. We tested it, it failed the test, so it is off the screen.

& IT IS STILL HAPPENING, WHICH IS THE POINT

Our own testing keeps catching its own mistakes over & over again. We will catch more when the sites are live with more data coming in, making our results even better.

What we would rather you take from this than from any single number. All 5 of those errors were found by us, in our own work, & each cost us a claim we liked. That is the same job we are selling: checking the answer before somebody acts on it, & saying so when it changes.

A SECOND RUN, DIFFERENT QUESTIONS, SAME FINDING

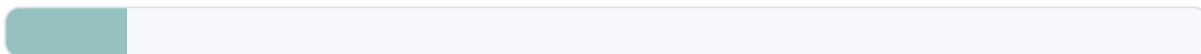
Claude Opus 5, 24 of 31 right (77%)



GPT 5.6, 18 of 31 right (58%)

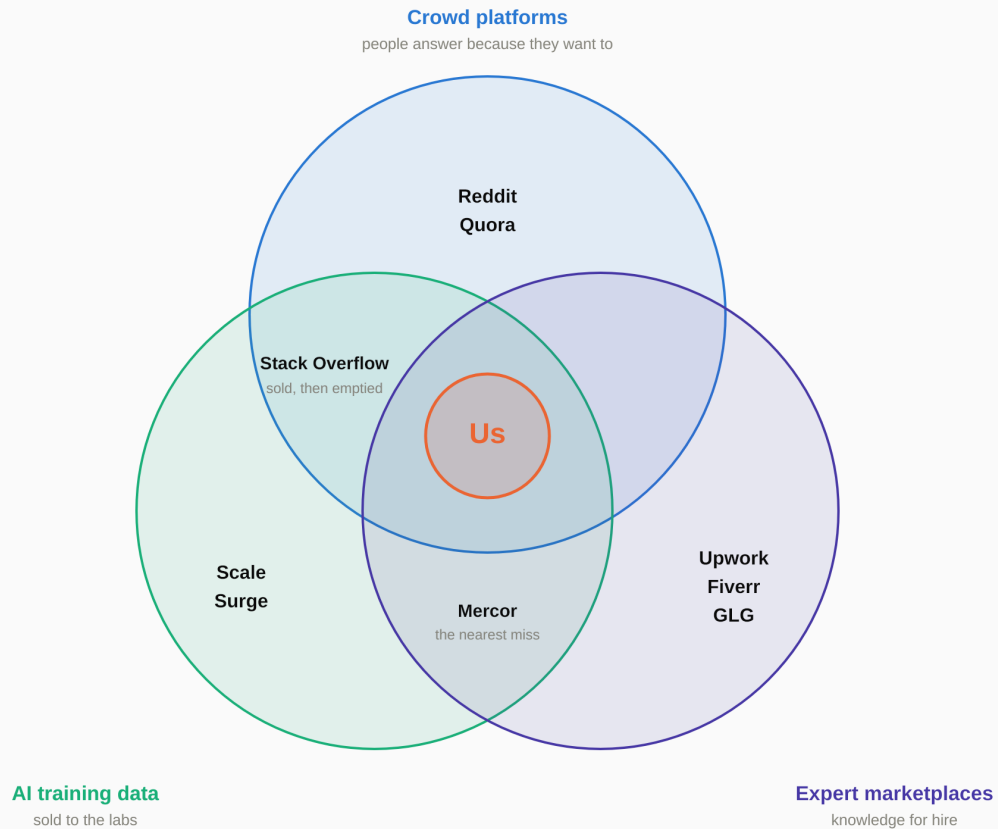


Gemini Flash, 3 of 31 right (10%)



67 points top to bottom, & 19 points between the 2 best. Same 31 questions, same day.

Help us grow the little circle in the middle



Reddit and Quora hold the crowd and sell it, but grade nobody. Scale and Surge grade, but buy their people by the hour. Upwork, Fiverr and GLG have the experts, and sell their time rather than their verdicts.

Stack Overflow's collapse did not start with AI. It started in 2014, when moderators started shutting down people's questions so nobody could answer them. We never reject a question.

More details on market fit

	Reddit 400M+ weekly users	Scale, Surge paid labelling	Stack Overflow volunteer Q&A	Upwork hired experts	Us pre-launch
People turn up without being paid per task	☒	☐	☒	☐	☒
Answers get settled and graded, not just discussed	☐	☒	<input checked="" data-bbox="1034 425 1058 448" type="radio"/> ¹	☐	☒
Takes no money from the AI engines it grades	<input data-bbox="678 492 702 515" type="radio"/> ²	<input data-bbox="858 492 882 515" type="radio"/> ³	☒	☒	☒
It is about a named AI's answer, not a human's question	☐	☒	☐	☐	☒
You can earn from what you know in minutes, with no bid, application or interview	☐	<input checked="" data-bbox="858 638 882 660" type="radio"/> ⁴	☐	<input checked="" data-bbox="1197 638 1220 660" type="radio"/> ⁵	☒

☒ has it ☒ partly ☐ does not

Stack Overflow marks an accepted answer, but that means the person who asked was satisfied. It is 1 person's approval, not a check. Reddit is paid about \$130M a year by OpenAI and Google, and Google's Gemini runs Reddit's own answer product. Meta bought 49% of Scale AI in 2025. OpenAI and Google moved their work elsewhere. They pay well and pay up front, but you have to be screened onto the workforce first. Upwork does let people earn from what they know but we do it with less friction.

21 questions we've answered about AI mistakes we don't think anyone else has fully solved & the 21 staff we give you to increase your effectiveness, knowledge, & efficiency while they automatically help organize your life.

What we've answered, & keep answering

- 1 How do you dramatically improve the accuracy vs. people asking several different AIs?
- 2 Does a cheap AI ever beat an expensive one, & on what?
- 3 How often can an AI spot a mistake in another AI's answer?
- 4 On average can a second AI produce a better answer by being fed the answer from another AI?
- 5 Can you tell, before you trust it, when an AI is about to be wrong?
- 6 Does asking the same question a different way change the answer?
- 7 How can you get an AI answer summarized so you don't have to read a book every time & get straight to the most accurate parts of the answer?
- 8 Is it better to tell an AI to read a page, or to let it use what it already knows?
- 9 Which AIs quietly refuse to answer, & what do they refuse?
- 10 Can we make a free answer as accurate as a paid one?
- 11 How do you fix the answer when your question to an AI inadvertently contains something that isn't true?
- 12 How far out of date are the answers you are getting today, & which AIs are more current on which topics?
- 13 Does searching the web fix an out-of-date answer?
- 14 Does asking more AIs make the answer better & which topics?
- 15 By how much are many different AIs better at which subjects, & at what cost differential?
- 16 What does it actually cost to check an answer properly?
- 17 Which AIs give the best life advice, & does the sparse published ranking hold up?
- 18 How do you know which parts of a long AI answer to believe?
- 19 When all the AIs agree & they're all wrong can anything catch it?
- 20 Does an AI know when its own information stopped being current, even if you ask it directly?
- 21 Can an AI be talked out of a right answer by being told the other AIs disagree?

17 of these we answered by running the tests ourselves, 4 by research & our own reasoning. Every answer comes with the number of questions behind it, the engines it ran on, the date it ran, & what result would have changed our mind. Partners see the answers.

The staff that puts those answers to work



Earn from what you know

Visibility Fixer · Job Scout · Question Monitor



Catch things in time

Out-of-date things you wrote · What email needs you today · Out-of-date answers you got · Scheduling Desk · Meeting Prepper · Refund Chaser



Check what is true

Fact Check · Self-contradiction checker · Company Reviewer · Translator · Life Coaches



Run your day

Note Tidier · Notebook Auto-Organizer · First Draft · Product Hunter · Drafting social posts · Warranty Manager · Form Filler

They read your notebook, & the advisor hands out the work. You choose which ones run, & nothing is ever sent anywhere until you press send.

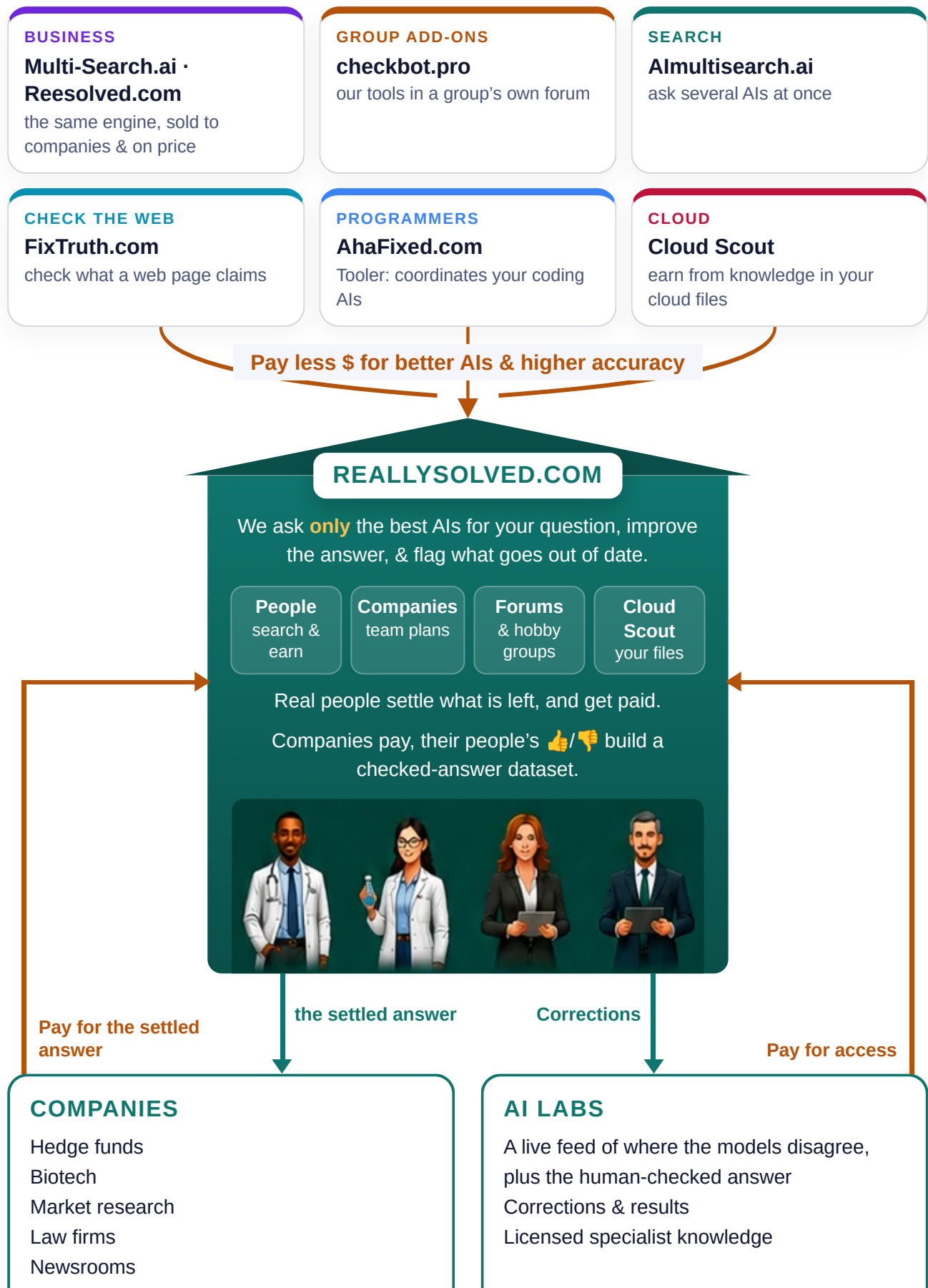
Why people keep paying

You wrote something a year ago. It is not true any more, & nobody told you. **We watch what you have written or been told, & tell you the moment something newer disagrees.**

- Asked what would keep them paying, people picked this over the search itself.
- Asked what it is worth, 1 person said about \$10 a month. A company watching its own old advice said \$300 to \$1,000 for the same watch. **Same feature, 30 times the price.**
- They doubled that again for a dated record of what they were told & when, because it proves they were right & have knowledge on the subject.

How the core works.

The same engine behind every door, & money arriving from subscriptions, companies & AI labs.
Website URLs are just placeholders for GTM partner to rebrand.



+ more

Behind the building: Certainize.ai (developers & API) · FixAI.group (independent AI safety council)

Patent pending, reviewed by a top-10 patent firm

1 engine, multiple doors.

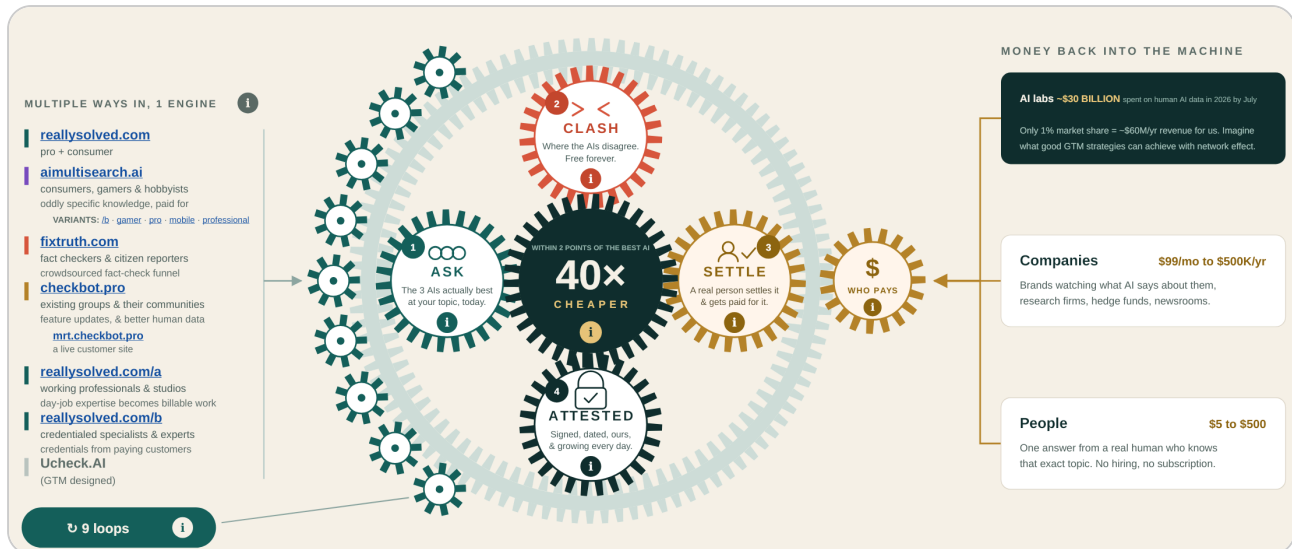
- Where the other doors are
- Infrastructure, not doors
- What a company is buying
- Why an AI lab would pay for this

THE MECHANISM

One engine, drawn as the machine it is.

Ask, clash, settle, attested. Built & running; opens to the public at launch.

Tap to open the live one, where every gear moves & opens →



THE SITE PORTFOLIO

BETA FUNNEL EXAMPLES

[ReallySolved.com](#) ↗

The "Serious Expert" funnel: attestation & a topic-specific reputation score. Our brand, for now.

[Multi-Search.ai](#) ↗

[checkbot.pro](#) ↗

Feature updates for existing groups. They get much better engagement & results while we get better human data for the AI labs.

[mrt.checkbot.pro](#) ↗

A live customer site.

[Reesolved.com](#) ↗

NEGLECTED

Real potential, less attention so far

[FixTruth.com](#) ↗

Fact-checking - the angry, X/Twitter-driven crowdsourced funnel.

[AIMultisearch.ai](#) ↗

Compare 3 AIs correctly - the viral, gamified, crowdsourced consumer funnel.

INFRASTRUCTURE

The engine underneath everything else

[FixAI.group](#) ↗

An independent AI-verification council - the engine Cyrus could use.

Certainize.ai ↗

The company funnel + API + AI labs + agent-support expert network.

AhaFixed.com ↗

Tooler, a coordinator for coding skills, plugins, MCP, CLI, agents, hooks, commands, routines & connectors, for any programmer or vibe coder. Built for people who already installed them all & still have problems: it trims the pile so the right ones show up when they are needed. Next: with their OK, the fixes coders make to their AIs become human correction data.

Free lists rank AIs. None covers your question's meaning, or attempts to use algorithms to give a more accurate answer, or says when an answer went stale.

What a buyer wants to know

On the free lists?

✓ Best at coding?

Yes

✓ Best at math?

Yes

✗ Best at my topic?

No list has it

✗ Improves the answer?

No list has it

✗ Has this answer changed?

No list has it

They score exam subjects, not what people type.

Stripe agreed to pay \$7.5B to see where AI spending goes, but hasn't closed as of this writing. We tell them who got it right.

We sell accuracy to people today, & to their agents next, on rails Cloudflare is building.

EXPANSION

3 more businesses, inside the same machine.

Each 1 is a real business on its own, & all 3 run on the engine that is already built.

Coopetition

Work with the big apps where it pays. Compete with them where it does not. Cooperate to grow the pie, compete to split it.

A p p	What they earn	Engage ment	What they charge	Se ver al AIs	Disag reeme nt summary	Improv es answer accuracy	Hum an settl es it
ReallySolved							
ChatOn							
Chat & Ask AI							
Nova							
ChatNow							

* GTM partner to help decide.

Not 1 of them settles it, pays a person to settle it, or keeps the record of who was right. They are the audience we plug into, not the company we race.

Enterprise extension or white label

Enterprise search tools like Glean, Microsoft 365 Copilot & Perplexity Enterprise hand a company's staff an answer, & nothing independently checks it. **We would like to be the engine they use. We could add value their big audiences can plug into & we keep the human corrected data.**

Proof-of-thinking

AI detectors guess whether a person or a machine typed something. **We can show the work instead:** where the AIs disagreed, what got corrected, & which real person checked it. Anyone can read that record, & it cannot be faked.

GTM TO DECIDE:

1. How we reach people sequence:

- A) our B2B2C doors (search, groups, testing)?
- B) back end inside others' Coopetition apps?
- C) compete with them head on?
- D) enterprise extension or white label?
- E) proof-of-thinking?

2. Price, & which features land in which plan. Marked * on this deck wherever it is still open.

Which plans Deep Check belongs in & what it costs on its own are both TBD.

3. What a company buys: 1 bill for a team, who can see what, & what leaves with them?

4. Whether we go after big companies?

5. Whether letting a group keep its own copy wins more groups than the lock-in it gives up?

6. Expert-supply channels: direct recruiting (Reddit, StackOverflow, GitHub), professional-association partnerships, corporate HR outreach, content-led (LinkedIn first), or a paid referral program. Which first, & how much budget?

The ones already asking **several AIs at once.**

Company	Price	Score, & how it is made	Self-reinforcing loops	Disagreement summary	Improves answer accuracy	Human settles it	Enterprise features	How far along										
ReallySolved									Free / \$5/mo / \$9/mo / \$19/mo* \$50 / \$90 / \$190 a year	Not an agreement score. A settled answer, from a proprietary blend of math + statistical analysis	9. Contests, disputes, badges, streaks, referrals & more	Yes. We write out where they disagree	Yes. The best AIs for your question + our own accuracy algorithm → ~82% on \$5, ~95% on \$9 & \$19	Yes. Real people settle it & get paid	Tell us which 3 to build first	Built, running & tested		
ConvergePanel									Free / \$99.99 /mo / \$169.99/mo / \$1,200 / \$2,040 a year	0 to 100, & their own page admits it measures agreement, not accuracy	None found	No	No	No	Audit trail, governance dashboard, peer review	Real positioning, thin proof		
Talkory.ai									Free / from \$5 / Enterprise custom no period published	A % plus an agreement count	None found. Each AI reviews only its own answer, which is not a loop	No	No	No	SSO, white label, data residency	Polished, feels like a small team		
MultipleChat									Free / \$20/mo / \$90/mo / \$240 / \$1,080 a year	Flags where answers disagree, no number	1. A role-assigned review debate	No	No	No	SSO, Trust Center, team admin	The most built out, 30,000+ users claimed		

Council AI	\$19.99 /mo / \$39.99 /mo / \$99.99 /mo \$240 / \$480 / \$1,200 a year	A number, method not published	1. A debate, models challenge each other	No	No	No	Works paces + an MCP server	Thous ands of users claime d
Suprmind	\$19/m o / \$45/m o / \$95/m o / \$195/ mo + custom \$228 / \$540 / \$1,140 / \$2,340 a year	A map of where they agree & disagree, no score	1. A debate, in 3 flavors: debate, red team, first principles	No	No	No	Enterpr ise tier	Real pricing tiers, active marke ting
AISCouncil	Free / \$5/mo \$60 a year	None	1. A debate, with peer review & a vote inside it	No	No	No	Open source	Small, mainte nance in progre ss
AiZolo	Free / \$9.90/ mo \$99.90 a year, their own annual price	None	None found	No	No	No	None	Small, built for search traffic
OneScales / MultiLLM.pro	Not verifiab le	Not verifiable	Not verifiable	No	No	No	Not verifiab le	Thin

* *GTM partner to help decide.*

Not 1 of them keeps a record of who turned out to be right.

& not 1 lets the expert keep what they earned or turns specialized hobbyists or professionals (not looking for a job) into opt-in instant consultants. They are all work-for-hire for the labs, paid by the lab rather than by the people asking, so the record stays with the lab.

Sources

60 features built & running. 7 left to build.

17



Answers & accuracy

- Search by photo
- A commentator in the tone you choose
- The Right/Wrong slider

15



Your notebook & staff

- The auto-organizer
- Imports from Notion & Evernote
- Your knowledge as a 3D picture

9



Community & experts

- A live group forum with our AI inside
- Expert profiles & a short personal address
- Track-record points

11



Companies

- A company plan with an admin page
- A shared team space
- Slack & Teams apps

3



Sites & sign-in

- 9 live sites
- 1 sign-in across all of them
- Google, email link or Microsoft

5



Plans & payments

- A 7-day free Plus trial
- Checkout for every plan
- A billing page people manage themselves

Still building

1. Being found & hired by companies
2. Watch my own documents
3. Called it first
4. Deep Check
5. What AI tells people about you or your company

Looking for a **GTM Partner**

THE OFFER

Tell us what your terms would be to take over the promotion & running of the websites. Our current idea is: take 100% of quick ramp-up net revenues until you've recouped 3x your time & costs + equity & ongoing profit sharing from AI lab revenue & all other revenue. Propose the deal that makes sense for your cost structure, current GTM staff, & resources.

Revenue progression: Companies & Consumer first, then anonymized human corrected data revenue from AI labs.

What flows to the GTM/Ops partner

Reaching... (paying members)	Subscriptions		Share 50%?*		Share 25%?*	
	Year 1	Every year after	Year 1	Every year after	Year 1	Every year after
1,000 + 5 companies	~\$41K	~\$77K	~\$250K	~\$490K	~\$370K	~\$740K
10,000 + 20 companies	~\$476K	~\$880K	~\$2.5M	~\$4.9M	~\$3.7M	~\$7.4M
100,000 + 100 companies	~\$4.3M	~\$8.0M	~\$25M	~\$49M	~\$37M	~\$74M
1,000,000 + 500 companies	~\$42M	~\$78M	~\$250M	~\$490M	~\$370M	~\$740M

* The AI-lab line, after members get their share. How much to share with members is your decision as the partner.

Mature content filters are your decision too: whether they ever launch, & whether they start on or off.

The AI-lab line is upside, not the plan. It is never counted inside the subscription figures, & nothing here depends on it.

No advertising money is assumed anywhere in this table. Every sign-up in it arrives through outreach, not through buying attention. That is the half we are asking a partner to own.

The margins, plan by plan

\$5/mo Certainizer™	\$9/mo Plus	\$19/mo Pro
---------------------	-------------	-------------

How much of every \$1 we keep	~80%	~85%	~92%
Profit from 1 subscriber, 1 month	\$4.02	\$7.69	\$17.40
What 1 subscriber is worth over a year	\$48.24	\$92.28	\$208.80
	\$5/mo Certainizer™	\$9/mo Plus	\$19/mo Pro
What we get back for every \$1 spent winning a subscriber			
at \$0.17	~284 : 1	~543 : 1	~1,228 : 1
at \$10 to win them	~5 : 1	~9 : 1	~21 : 1
at \$30 to win them	~1.6 : 1	~3 : 1	~7 : 1
at \$60 to win them	~0.8 : 1	~1.5 : 1	~3.5 : 1
How fast that \$0.17 comes back	~1 day	<1 day	<1 day

We are guessing at how fast users arrive & stick. We are hoping our GTM partner will have better guesses & the experience to know which of our strategies are good, which need fixing, & better strategies to share.

3 revenue streams are EXCLUDED from every figure above.

All 3 are built into the product. **Not a cent of any of them appears anywhere above, so every number here is a floor, not a forecast.** Deep Check pay-per-answer · bounty commission · companies buying verified answers. **The AI-lab corrections corpus is now forecast on its own, in section 6 of the forecast page. The company & AI-lab lines are the destination. The consumer subscriptions above are the bridge.**

The AI-lab line is the biggest of them all.

What the AI-lab line is worth

Where those come from: items times a price per item. A quick 👍 or 👎 sells for little, a written correction for much more, & both add up.

Our users can produce more of this data than the world currently buys, which solves the problem of AI being out of control.

👤 The people side costs us \$0 up front. The people who check answers are paid out of what that check earns, split 50/50* (* what we would promise users if the GTM team agrees). A share of the money, not a wage. Paying them by the hour instead is the 1 line that would sink a margin like the one above.

Every figure in this money block is our own estimate, built bottom-up, & is reproduced from the line-by-line forecast where each one is worked out in full. Built & running; opens to the public at launch.

What's in each plan

Unlimited on anything that only thinks. Metered on anything that looks something up, posts, or runs on a timer.

Accuracy climbs with the plan: ~60% asking 3 AIs yourself, unaided → ~82% on Free & Certainizer \$5/mo → ~95% on Plus \$9/mo → ~95% on Pro \$19/mo.

Feature	\$5/mo	\$9/mo	\$19/mo	What tests said it's worth stand-alone
What changes from plan to plan				
AIs that can answer	80	100	169	Plus is 85% cheaper · Pro saves ~\$7,000/yr, vs paying each AI
The top AIs: Fable, Opus, o3 & GPT-5.6 Sol	—	—	✓	
AIs from outside the USA	—	4	6	
Follow-up & out-of-date answers	5 facts	20 facts	50 facts	\$40-50/mo as a build check
Find an old question by meaning	—	30 days	all	
Search by photo	—	✓	✓	
Fact Check	✓	more checks	unlimited	
What email needs you today	✓	+ catch-up on threads	+ drafts replies	
Staff	starter set	+5 more	+4 more	
Notebook Auto-Organizer: files on your PC & cloud storage	—	—	150 entries a month	
The same on every plan				
Auto-routing to the best AI for your topic	✓	✓	✓	Free on every plan
Certainizer™ settles the disagreement	✓	✓	✓	
History kept while you're subscribed	✓	✓	✓	
Job Scout, 3 months, while hunting · in Certainizer	✓	✓	✓	\$15/mo hunting, \$0 the week they sign

People independently priced a single high-stakes answer between \$5 & \$20. No price shown for Fact Check: unanimous that nobody would pay for it, & that charging for it would change how they read the company.

Who buys companies like this

- **Stripe agreed to pay \$7.5B to see where AI spending goes, but hasn't closed as of this writing. We tell them who got it right. That has to be worth more.**
- **Lovable is worth \$13.3B & is openly shopping for companies to buy.**
- **Reddit & LinkedIn hold the argument & the credentials. Neither holds the answer. (potential acquirer)**
- **Microsoft says the moat is the learning loop, not the model. We could be a natural independent fit for them because they can't appear independent. (potential acquirer)**

Funded

Founded & funded by a former VC, not raising money & tiny burn (mostly patent attorney)

Running

Built & tested today, not a plan

~80-92%

Of every \$1 we keep, after the AI & card fees

Patent Pending

Reviewed by a top 10 patent firm

Founders & team



Bob Haya, Founder

Retired venture capitalist

Earlier in his career, Bob had 2 exits & co-invested alongside firms like Andreessen Horowitz, Kleiner Perkins, & Khosla Ventures, etc. See:

createlabs.com/portfolio/portfolio.html

Bob still collaborates with his billionaire VC partner & informal advisor Cyrus Hodes, both personal friends with contacts at the highest levels of today's top AI companies & the White House.

[Click on ABOUT US here](#)

The longer bio reallysolved.com/gtm/bio

Open to any ideas, including adjusting the business model, or maximizing the underlying engines for entirely different applications.

Already built several acquisition funnels with persona variations (more verticals & personas already spec'd). Patent pending, reviewed by a top-10 patent firm.

"I want another source of ground truth in the world, and I want to contribute to the gig economy."

"I'm easy to work with, Open to changing Branding...just blow it up!"

"Twice I walked into a government office with the wrong forms & procedures. Both times I had asked the best AI there is. Both times it answered confidently, & both times I was embarrassed & sent home. AI is helpful but useless if I can't tell which parts are correct & there's no way to fix it for the next person."



Mathew Svensson

Co-founder, B2B

Mathew spearheads the Hands-On B2B consultative services of ReallySolved & is focused on the workflows between departments & how to apply AI most efficiently. MSc in Technology Entrepreneurship, Technical University of Denmark. Over a decade helping companies build, streamline, & scale their operations. Recently got the highest possible grade on his AI related thesis, 12 (A+), and was told that all his sub thesis work would have been normally a thesis each on their own.

"I design intelligent organizations. Most companies think they have a technology problem. Usually it's a flow problem: information, decisions & work get stuck between teams, and adding AI on top doesn't unstick them. So I ask what kind of organization you want to become, and what's stopping you. I address flow first &

tools only where needed. I connect specialists so their work creates value. When I leave, your organization understands itself better, executes better, and can keep evolving without me."



Cyrus Hodes

Informal advisor

Founder of AI Safety Connect, which convenes governments, frontier AI labs & investors alongside the UN General Assembly & the international AI summits. Co-founder of The Future Society at the Harvard Kennedy School. Leads the Safety & Assurance of Generative AI project at the Global Partnership on AI, sits on the OECD expert group on AI futures, & runs the GRASP global risk mapping at the Mohammed Bin Rashid School of Government in Dubai. Venture partner at Lionheart Ventures. Formerly advisor to the UAE's AI minister.

Where the people come from

People do the work for money, status & a reputation they can use.

STEP 1

Consumers

Free multi-AI search. An answer costs us almost nothing.* **BUYER 1, IN SMALL AMOUNTS**

STEP 2

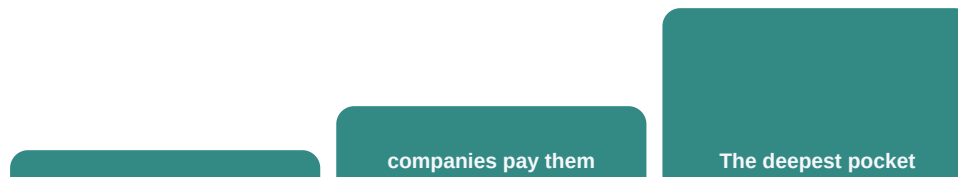
Experts

A few searchers know a subject better than the AIs do. They show it in public.†

STEP 3

AI labs

They buy the questions the AIs got wrong, plus the checked answer. **BUYER 2, THE BUSINESS**



* An answer costs us **85¢ per 10,000**. 4 tiers: free, \$5/mo, \$9/mo, \$19/mo.

† **We line the buyers up first, then pay.** Companies are phase 1, run by hand: ask 20 what they would pay before a line of code. If nobody says yes, we saved a year. We will ask companies what they want & design the bounties program around that before we engage experts on the platform & outreach to other experts. **That is bounties. Everywhere else both sides arrive together:** the games bring the people in while the outreach brings the companies, & neither waits for the other.

EARN MONEY · 1 OF 4

Answer for a stated price

EARN MONEY · 2 OF 4

Get hired directly

EARN MONEY · 3 OF 4

Win disputes

EARN MONEY · 4 OF 4

Weekly contests

GROW FOR FREE

Invite a friend

GROW FOR FREE

Creator badges

HABIT & STATUS

Streaks & leaderboards

HABIT & STATUS

Dispute Hunter challenge

EXTRA LOOPS DISCOVERED ALONG THE WAY

Innovative group features not available on any platform today

THE QUIET ONE · RUNS FIRST

Your notebook



1 Track record - moves on everything they do, including a 👍 or a 👎. It is standing, never cash. It is what gets them offered the paid work

2 Correction Earnings - moves only when a solution or a correction is accepted. Shown in \$
A bot that farms the first one gets nothing it can spend.

Our first real test.

A group with an active Facebook membership now has its own forum, & its members come over next. **The wider plan is to give group owners better tools at a price they can afford**, by partnering with the companies that host communities like theirs.



Why it compounds: about 3 searches in 5 turn up a disagreement, & every disagreement is free raw material for the next answer. **On our 91-question everyday bank, 55 did.** 9 loops feed each one back to the top of the flywheel.

The cheapest door in: Fact Check

Paste what another AI told you. We say where it's shaky. **No account, no email, & never a signup in front of the result.**

- **It works on people who have never heard of us.** That is what makes it distribution, not a feature.
- In people & persona testing, every one said they'd tell a friend about it, & every one also said they would not pay for it.
- **The email ask, if any, comes after the answer is already on screen.**

- It is also the cheapest proof we have that we are not just a wrapper around one AI: it exists specifically to catch the others.

17 / 22

Contributor, sign up, first search

Solver, first paid job · \$500-\$2K/yr

Senior Solver, a track record · \$5K-\$15K/yr

Master Solver, a strong year · \$50K-\$200K+/yr

Solver Laureate, a career capstone

The ~500 Topic & Safety Matrix. The board below shows 8 topics out of ~500 researched or tested variables, because 1 AI is better at stem cell science & another at robotics, & calling them both “science” hides that. The same record runs a 2nd time for safety, across 245 safety topics, & a 3rd time for the 7 questions with no right answer & 4th time to check out of date answers (all included in the ~500).

The leaders keep changing. We watched it move this month: the model we tested ourselves against was retired 6 days later, & a second one was discontinued the same week. Yesterday's best is not today's best.

So we keep the list, subject by subject. From launch, which AIs are best at which topic is rewritten by real corrections on our own platform, as answers come back & people correct them. Here is the board.

ENGINE	TIER	CRYPTO	BIOHACK	CURRENT EVENTS	INVESTING	CYBERSEC	MEDICAL	SCIENCE	LEGAL										
Claude Fable										Pro	78	85	70	88	90	92	95	91	
Claude Opus										Pro	63	82	68	86	88	90	93	89	
OpenAI o3										Pro	82	84	72	90	94	93	96	90	
GLM (non-USA)										Plus	70	72	78	75	76	78	82	79	
GPT-5.6 Sol										Pro	80	83	82	88	91	91	94	88	
DeepSeek R1 (non-USA)										Pro	80	77	60	88	88	81	94	79	
Claude Sonnet										Plus	75	80	68	82	86	86	90	85	
Claude Haiku										Plus	62	65	60	68	70	72	76	72	
GPT-4o										Plus	72	75	80	80	82	84	86	82	
Gemini Pro										Tested	44	78	90	0*	84	88	92	84	
Kimi (non-USA)										Plus	66	68	78	74	72	75	80	80	
Qwen Max (non-USA)										Plus	72	74	80	78	80	82	88	81	
DeepSeek (non-USA)										Plus	76	75	65	82	84	80	91	78	
GPT-OSS										Free	50	52	45	55	58	56	62	54	
Gemini Flash Lite										Free	48	60	82	45	62	68	72	60	
Llama 3.1 8B										Tested	54	55	48	58	62	60	68	58	
Grok										Pro	75	74	92	80	80	78	86	76	
Mistral										Pro	70	72	65	76	82	78	85	84	
Command R+										Tested	60	62	75	78	68	70	72	82	

91-100 expert

76-90 strong

51-75 competent

21-50 weak

0-20 unreliable

These board scores are our best current read of the landscape (public tests, July 2026), not our own lab measurement yet. The 169 AIs we tested ourselves are a separate count, on page 1. "Tested" = scored by us, not on a plan. *Gemini's investing score is a real result, not a data-entry error: it's built to refuse investment questions.

The safety half: 245 safety topics, 22 categories reallysolved.com/gtm/safety

We are not the only ones **saying this.**

Bill Gates published a nearly 6,000-word essay on August 26th 2026 warning that artificial intelligence is advancing far faster than governments and society are prepared to handle.

BILL GATES · AUG 2026

“AI is confident, not always correct.”

ETHAN MOLLICK, WHARTON

“Base model capability is commoditizing on schedule. **Trust isn't.**”

DR SETH DOBRIN, SILICON SANDS NEWS · AUG 2026

“Chat AIs are trained to be agreeable, and **agreeable advice is worthless the day you're about to make a real mistake.**”

HYPERAUTOMATION AI REPORT · AUG 2026

Fractl's survey found that people are not walking away from AI. **They are using it more & believing it less.**

FRACTL / SEARCH ENGINE LAND · JUN 2026

Dr Seth Dobrin has said that the AI that answered you last month is not the one answering today, & nothing tells you it changed.

DR SETH DOBRIN, SILICON SANDS NEWS · AUG 2026

→ *We watch for that.*

CustomerThink has said that the people who built it cannot say why it answered that.

CUSTOMERTHINK · AUG 2026

→ *Somebody outside has to check the answer itself.*

Dr Seth Dobrin has said that as AI starts **doing** things instead of just saying them, an unchecked answer stops being a nuisance **& becomes a liability.**

DR SETH DOBRIN, SILICON SANDS NEWS · AUG 2026

VentureBeat has reported that a new \$6B AI company graded its own accuracy against OpenAI, Anthropic & Google, **on a test it built itself.**

American Express has said it will now cover the bill when an AI agent buys the wrong thing.

AMERICAN EXPRESS · APR 2026

→ *Somebody still has to decide what “wrong” was.*

The industry says picking **the right model for each job** is how AI stays affordable.

NVIDIA, ANDREESSEN HOROWITZ, Y COMBINATOR, MICROSOFT, META & OTHERS · JUL 2026

→ *That is our engine.*

→ *We match each question to the right AI at the right price, & cheap independent AIs check the expensive ones.*

“The real opportunity is not in picking the best model but instead in **building a learning loop on top of models where human capital and token capital compound.**”

SATYA NADELLA, MICROSOFT · JUN 2026

→ *The winner is not the best model. **It is the loop of people correcting it.***

Ben Lorica has said that companies are learning to turn their own data into specialist AI they own.

BEN LORICA, GRADIENT FLOW · JUL 2026

→ *When everyone can train their own AI, **verified human correction becomes the scarce part.***

The Financial Times & Bloomberg have reported that the market is now pricing what **frontier AI costs to build.**

FINANCIAL TIMES / BLOOMBERG · JUL 2026

→ *We never take that cost on.*

Why we get better every day, & why a copycat can't catch up.

Published test, February 2024: the same AI got 93% right on facts that don't move, & 26% on facts that do. 125 questions in each group.

Knowing which AI is right **cuts our own bill, starting day 1**. If somebody pays us for that knowledge later, **treat it as a bonus & count it as zero**.

PAYS OFF FROM DAY 1

1. We stop using the AI that got worse

PAYS OFF FROM DAY 1

2. We tell people when their answer goes out of date

WE NEED 1 TO TRY IT

3. Companies that put an AI inside their own product

BEST IDEA, WORST ODDS

4. The AI companies themselves

SLOW, FEW SALES, BIG CONTRACTS

5. Reports for big companies choosing an AI

BIGGEST, & THE LONGEST WAIT

6. Insurance, one day

Read it this way: the first 2 pay for themselves. The other 4 are what those 2 make possible.

Specialist AIs too, not just the big ones.

We've mapped 200+ AIs built for 1 field: medicine, law, finance, science, math & dozens of languages. **Every one that makes sense will be in the lineup by the time the sites launch,**

tested against the big AIs on its own subject. Actuarial work has no specialist AI yet, **so that is the one we intend to own.**

Funds are buying small firms & handing the work to AI. **We check that work before a client sees it.**

Accountants, bookkeepers, tax preparers, IT help desks & property managers. Owners are retiring, funds are buying, & their plan already has a step where the AI's work gets checked. **We are built to be that step.**

STEP 1 · THEIR AI

Collects the paperwork

Emails, invoices, uploads & forms, sorted & filed.

STEP 2 · THEIR AI

Writes the first draft

The tax return, the ledger, the monthly report.

STEP 3 · US

Several AIs check the draft

AIs from different companies read it. Where they agree, it moves on.

Where they disagree, a person sees exactly which line & why.

STEP 4 · THEIR PERSON

A person signs off

Nothing reaches a client without it. Their time goes to the lines we flagged, not to rereading everything.

Every fix a person makes becomes a rule.

The correction is saved in the firm's notebook, so the same mistake gets caught next time, in every firm the fund owns. **Their checking gets better every week. The data stay with the company & we keep an anonymized record if we can sell the anonymized human corrected data.** It is the same human correction data the rest of this deck is built to collect.

Nothing new to build: the several-AI check, the expert crowd & the notebook already exist. This is 1 more buyer for the same machine.

Every number, every test, every rival.

CAP TABLE: Uncomplicated. Bob currently owns 100%.

TAM = everyone frustrated with wrong AI answers

Below are 6 more pages. They are not short summaries where we picked the good bits. In 2 of them we say, out loud, that we got a number wrong earlier & then fixed it, which most companies would quietly delete instead.

The machine, drawn as 1 gearbox live, tap any gear

The safety & online-harm matrix scored the same way as accuracy

Who else is doing this the full map, plus Microsoft & the 11 small rivals

The forecast, line by line every row, every assumption

GTM notes & the stack for the GTM engineer

Where every number came from including the ones we refused to use

Let's build the growth machine together →