

Why voice AI fails

We read what owners say about AI phone agents. It is not pretty.

Built on 10 million minutes of real calls and 85 real complaints

VoiceROI

September 2026

The short version

We collected 85 real complaints about 14 AI phone products from Reddit. We sorted them into ten groups. Here is what small businesses ran into, and what we do about it.

85 complaints on Reddit. 131 complaints across every source we checked. 41 separate threads and reviews. Counted one by one, and you can check our work.

These are the top ten complaints. If you run a dental office, a law firm, or a service business, every one of these has a price tag. Missed calls. Wrong bookings. Staff time. Your good name.

The top ten complaints

#1 You end up doing the work

16% of the complaints we counted (14 of 85)

They sold you a robot. Then they handed you a second job. Owners spent hours on prompts, rules, and knowledge files. Months later, keeping it right was still their job. One team said it cost them more than one hundred hours of work. Another was trying to hire help just to make it useful.

"burned over 100+ dev hours building on [vendor]"

— a Reddit user, 2025

"i spent hours tweaking prompts on [vendor]"

— a Reddit user, 2025

What it costs a small business

Every hour you spend tuning a phone bot is an hour you do not spend on patients, clients, or sales. If you stop tending it, it slides. If you keep tending it, you are the unpaid tech team.

What VoiceROI does differently

With VoiceROI, you do nothing. We build it. We run it. You pay for results, not software.

#2 You pay. They do not deliver.

12% of the complaints we counted (10 of 85)

You paid for calls that did not help. You paid for test calls. You paid a big top-up before you saw one result. Some owners said the promises were not kept, help never came, and no refund was in sight.

"NO REFUND is in sight"

— a Reddit user, 2024

"the credits system feels unnecessarily opaque"

— a Reddit user, 2026

What it costs a small business

Fees you cannot predict eat the money the bot was meant to earn. A month lost chasing a sales rep or a refund is a month of missed calls.

What VoiceROI does differently

VoiceROI is Results as a Service. Our promise is plain: 20% more sales in 6 months, or we work for free until you get there.

#3 It just did not pay off

12% of the complaints we counted (10 of 85)

The demo was great. The results were not. Owners said one call in five reached its goal. Others said the bot could mimic a voice and not much else. Some pulled the plug. Some said voicemail did the job better.

"The tech is AMAZING. But the results IMO were terrible"

— a Reddit user, 2024

"around 20% of calls achieved a goal"

— a Reddit user, 2024

What it costs a small business

A phone bot that does not book, sell, or help is worse than no bot at all. You paid for it, your callers put up with it, and you got nothing back.

What VoiceROI does differently

We pay attention to results because we are paid on results. 20% more sales in 6 months, or we work for free until you get there.

#4 You ask. Then you wait.

11% of the complaints we counted (9 of 85)

You ask a question. Then you wait. Owners said the pause before each answer was long enough to feel wrong. Some measured it at two seconds and more. One said the bot was slow on the very calls that mattered most, when a caller needed a real person.

"call latency is very very visible (~ >2s !!)"

— a Reddit user, 2025

"issues (awkward pauses, barge-in problems)"

— a Reddit user, 2026

What it costs a small business

A slow reply feels like a dead line. Callers hang up, or they talk over the bot, and the call falls apart. That is a lost lead who will call the next name on the list.

What VoiceROI does differently

Every Riley call is graded from the recording. A slow, awkward call gets a bad grade and a plain reason. It cannot hide.

#5 The booking did not happen

11% of the complaints we counted (9 of 85)

The bot booked the wrong staff. It told a caller a time was open when it was not. It set appointments for cars that were not for sale. It could not cancel or move a booking. Names and numbers went into a black hole. One owner said the bot never even ended the call after booking.

"it just kept booking the wrong staff on the calendar"

— a Reddit user, 2026

"ignores this and tells the caller the time is available anyway"

— a Reddit user, 2026

What it costs a small business

A wrong booking upsets two people. The caller who shows up. The staff member who was never told. Lost caller details are lost sales, and you will never know about them.

What VoiceROI does differently

Riley is graded on every call, from the recording. If a booking did not happen, the call is graded red, with a plain reason.

#6 It won't do what you tell it

9% of the complaints we counted (8 of 85)

Owners wrote the rules. The bot ignored them. It stopped following instructions past a certain length. It told callers whatever they wanted to hear. It contradicted the business. One owner tried to train it and gave up.

"often contradicts what we're trying to do"

— a Reddit user, 2026

"anything over ~6,000 characters ... it simply stops following instructions"

— a Reddit user, 2026

What it costs a small business

Your rules are there for a reason: prices, hours, what you will and will not do. A bot that ignores them makes promises you have to keep or break.

What VoiceROI does differently

Every Riley call is checked against how it should have gone, with a plain-English reason for the grade. We fix what drifts. You do nothing.

#7 It sounds like a robot and talks over people

9% of the complaints we counted (8 of 85)

Callers could tell. The voice was flat and robotic. It sounded pushy. It kept asking the same thing. And it cut people off before they finished their sentence.

"She doesn't let them finish answering"

— a Reddit user, 2025

"not at all professional sounding"

— a Reddit user, 2026

What it costs a small business

Your phone voice is your front door. A robot that talks over people makes your business sound cheap and rude, even when your team is neither.

What VoiceROI does differently

Riley's calls are graded from the real recording, not a transcript. Talk over a caller, and the grade drops. The reason is in plain words.

#8 Callers hang up on it

9% of the complaints we counted (8 of 85)

Callers got frustrated. They hung up. One owner said all of their callers hung up on it. Another said none of their clients would talk to a robot. Others said it confused customers and made the business look disorganized.

"All my callers get frustrated and hang up on it"

— a Reddit user, 2026

"It confuses customers and makes us look unorganized"

— a Reddit user, 2026

What it costs a small business

A caller who hangs up on your bot rarely calls back. That is a patient, a client, or a job gone to the next business on the list.

What VoiceROI does differently

Try Riley yourself first. Start the live demo, take the call, and see the grade within minutes.

#9 It makes things up

4% of the complaints we counted (3 of 85)

The bot made things up. It invented branches and services that do not exist. It offered services after hours and never asked. Its answers changed from one call to the next.

"it keeps hallucinating random things like fake branches or services"

— a Reddit user, 2025

"it offered our services after hours without checking with us"

— a Reddit user

What it costs a small business

A made-up price, service, or opening time is a promise your business now has to break. That costs trust, and it can cost you a client for good.

What VoiceROI does differently

Every Riley call gets a red, yellow, or green grade from the recording. The reason is in plain English. A made-up answer is caught in minutes, not months.

#10 It breaks and drops calls

4% of the complaints we counted (3 of 85)

It worked in the demo. Then real calls came in and it started to break. It dropped a call mid-sentence. The system went down. One reviewer said it never worked properly at all.

"drop a call mid-sentence"

— a Reddit user, 2025

"using it for real businesses ... it starts to break"

— a Reddit user, 2026

What it costs a small business

A dropped call is a caller who thinks you hung up on them. Downtime during business hours is a closed sign on your door.

What VoiceROI does differently

We check on Riley every five minutes, all day. Before we change anything, we test it on real calls.

What the testers found

The people who test voice agents for a living say the same thing. Hamming has watched over 10 million minutes of real calls. Here is what they found, in plain words.

They find out days after your callers do

"teams with transcript-only analytics discover issues 2-3 days after customers experience them. Teams with proper observability catch degradation in minutes."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 3

In plain words:

Most vendors only save a written copy of the call. When the agent starts failing, they find out two or three days after your callers did.

What it costs you:

Two or three days of bad calls before anyone notices. Each one could be a lost booking or an angry client.

It guesses wrong, sounds sure, and calls it done

"A voice agent that misclassifies intent confidently routes the caller to the wrong flow. Both calls might complete, but only one achieves the customer's goal."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 5

In plain words:

A human who does not understand asks a question. A voice agent guesses, then confidently sends the caller down the wrong path. The call still looks 'done' on paper.

What it costs you:

Callers leave without what they came for. The vendor's report still says the call went fine.

Callers who quit get counted as 'handled'

"High containment with low CSAT indicates 'false containment' - users giving up rather than getting helped."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 7

In plain words:

Vendors brag about how many calls the robot 'handled' alone. Many of those callers just gave up and hung up.

What it costs you:

A caller who gives up is a lost patient or client. The vendor counts it as a win.

‘Complete’ on paper, but they had to call back

“If a customer calls back within that window, the original call didn’t resolve their issue - even if it was marked “complete.””

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 8

In plain words:

If the caller has to phone again in the next few days, the first call did not work. It does not matter what the report says.

What it costs you:

Repeat calls waste your callers’ time and your staff’s time. They make you look sloppy.

Five times slower than a real conversation

“industry median TTFW is 1.4-1.7 seconds - 5x slower than the 300ms human conversational expectation. This explains why users report agents that “feel slow” or “keep getting interrupted.””

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 10

In plain words:

Most voice agents take about a second and a half to start talking. People expect a reply in a third of a second. That is five times slower. So callers feel the agent is slow, or they talk over it.

What it costs you:

Slow, awkward calls make callers hang up. Or they think your business is behind the times.

The average hides one bad call in ten

“A 300ms average can hide 10% of calls spiking to 1500ms. At 10,000 calls/day, that’s 1,000 terrible experiences that don’t appear in average metrics.”

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 15

In plain words:

The ‘average’ reply time can look great. Yet one call in ten can be painfully slow. Take 10,000 calls a day, and that is 1,000 bad calls the average never shows. (Hamming’s example, not a measurement.)

What it costs you:

The vendor’s average looks fine. Meanwhile one in ten of your callers gets a bad call.

They fix the wrong thing, so it breaks again

"Cascading failures are common. Audio degradation causes transcription errors, which cause intent misclassification, which causes task failure. Without multi-layer correlation, you'll see the task failure but chase the wrong root cause."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 18

In plain words:

One small problem, like bad audio, snowballs. The agent mishears. Then it misreads what you want. Then it fails the task. Most vendors see only the failed task and fix the wrong thing.

What it costs you:

Problems come back again and again. The real cause was never found.

It says 'booked' when nothing was booked

"Claimed completion without verification leads to false positive metrics."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 19

In plain words:

The agent can say 'you are booked' and nobody checks that it is true. The report counts it as a win.

What it costs you:

Callers show up for appointments that were never booked. That is lost income and an angry client.

It worked last month. Then something changed.

"Model updates, prompt revisions, and ASR provider changes trigger behavioral drift."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 21

In plain words:

The vendor changes the engine, the rules, or the voice. The agent starts to act differently. No one warns you.

What it costs you:

An agent that worked last month can quietly start failing this month.

It gets a little worse each week. No one sees it.

"Gradual degradation is harder to catch than sudden failures."

— Hamming, Post-Call Analytics for Voice Agents, Feb 2026, p. 28

In plain words:

A sudden crash gets noticed. A slow slide in quality, a little worse each week, usually does not.

What it costs you:

You lose calls a few at a time for months before anyone spots it.

What VoiceROI does differently

Riley is our AI receptionist. Every call is graded from the recording, red, yellow or green, with a plain reason. We build it. We run it. You do nothing. And you pay for results, not software: 20% more sales in 6 months, or we work for free until you get there.

THE GUARANTEE

20% more sales in 6 months, or we work for free until you get there.

How we counted

One row per complaint, straight from the source. We counted complaint statements, not customers. One person can post more than one complaint. We left out hearsay. We left out complaints about human staff. We left out posts by rival founders. Full list on request.

We started with 172 complaint excerpts.

115 came from Reddit.

We counted 85 of them and left out 30 that were secondhand, about human staff, or too weak to count.

They came from 41 threads and reviews about 14 vendors.

The percentages in this report are shares of the 85 counted Reddit excerpts.

The unit is a complaint excerpt, not a customer.

Method appendix

One row equals one excerpt. Group ids and the labels used across this report:

doing-the-work	You end up doing the work
pay-no-deliver	You pay. They do not deliver.
no-payoff	It just did not pay off
slow	You ask. Then you wait.
booking	The booking did not happen
ignores-rules	It won't do what you tell it
robot	It sounds like a robot and talks over people
hang-up	Callers hang up on it
makes-things-up	It makes things up
breaks	It breaks and drops calls
mishears	It mishears names and numbers
no-reason	Nobody can tell you why it failed

Appendix: the 85 counted Reddit quotes

Quote	Year	Group
"it keeps hallucinating random things like fake branches or services"	2025	It makes things up
"I've already tried setting strict_guardrails"	2025	It won't do what you tell it
"i spent hours tweaking prompts on [vendor]"	2025	You end up doing the work
"issues (awkward pauses, barge-in problems)"	2026	You ask. Then you wait.
"the voice still feels a bit robotic"	2026	It sounds like a robot and talks over people
"Role drift during verification"	2025	It won't do what you tell it
"Latency spikes on escalations"	2025	You ask. Then you wait.
"Memory contamination"	2025	It makes things up
"handling call quality and edge-case conversations was toughest"	2025	You end up doing the work
"almost no traction, people hung up"	2025	Callers hang up on it
"it still came across as very robotic"	2023	It sounds like a robot and talks over people
"frustrated by it constantly saying, 'is that available?'"	2023	It sounds like a robot and talks over people
"the latency is too big and the voice too robotic"	2024	You ask. Then you wait.
"Pathways. They're time consuming to setup and rigid"	2025	You end up doing the work
"using it for real businesses ... it starts to break"	2026	It breaks and drops calls
"anything over ~6,000 characters ... it simply stops following instructions"	2026	It won't do what you tell it
"the calendar integration [is] bad"	2026	The booking did not happen
"it just kept booking the wrong staff on the calendar"	2026	The booking did not happen
"ignores this and tells the caller the time is available anyway"	2026	The booking did not happen
"the booking is then attempted it fails"	2026	The booking did not happen
"no way for the Ai to reschedule or cancel"	2025	The booking did not happen
"[vendor] AI voice agent is a little bit slow"	2025	You ask. Then you wait.
"clients that kept dropping the service I offer"	2026	It just did not pay off
"the debugging experience is rough"	2026	Nobody can tell you why it failed
"burned over 100+ dev hours building on [vendor]"	2025	You end up doing the work
"complex state machine built on brutally brittle infrastructure"	2025	You end up doing the work
"endless error handling"	2025	You end up doing the work
"drop a call mid-sentence"	2025	It breaks and drops calls
"This wasn't a prompt engineering problem; it was a systems nightmare"	2025	You end up doing the work
"expect to spend 80% of your time wrestling with HTTP requests"	2025	You end up doing the work
"[vendor] is a component, not the whole car"	2025	You end up doing the work

"latency is not great yet"	2024	You ask. Then you wait.
"around 20% of calls achieved a goal"	2024	It just did not pay off
"lagging on real calls"	2025	You ask. Then you wait.
"struggles a bit with following the instructions intelligently"	2026	It won't do what you tell it
"trial and error to really refine the instruction set"	2026	You end up doing the work
"per-call pricing makes it hard to justify"	2026	You pay. They do not deliver.
"the credits system feels unnecessarily opaque"	2026	You pay. They do not deliver.
"All my callers get frustrated and hang up on it"	2026	Callers hang up on it
"I do not like the voice, which cannot be changed"	2026	It sounds like a robot and talks over people
"not at all professional sounding"	2026	It sounds like a robot and talks over people
"not natural or comes across as pushy"	2026	It sounds like a robot and talks over people
"none of my clients wants to talk to a robot they hang up"	2026	Callers hang up on it
"does not give a neat [record] of the name and contact information"	2025	The booking did not happen
"Too laborious. A VM is better"	2025	It just did not pay off
"She doesn't let them finish answering"	2025	It sounds like a robot and talks over people
"transcriptions to often be inaccurate"	2025	It mishears names and numbers
"This is a fatal flaw"	2025	The booking did not happen
"it's not reliable enough and causes too many problems"	2025	It just did not pay off
"There AI bot is horrible"	2026	It just did not pay off
"continue telling the customer whatever she wants"	2026	It won't do what you tell it
"often contradicts what we're trying to do"	2026	It won't do what you tell it
"It confuses customers and makes us look unorganized"	2026	Callers hang up on it
"not been able to find an IT team that can exterminate her"	2026	You end up doing the work
"set appointments on the Stop Sale units"	2026	The booking did not happen
"Was a little clunky"	2026	It sounds like a robot and talks over people
"it offered our services after hours without checking with us"	—	It makes things up
"it can be a little unpredictable"	—	It won't do what you tell it
"spend a lot of time with customer service to tweak the AI agent"	—	You end up doing the work
"getting my customers to trust [vendor] is going to be a little tricky"	2026	Callers hang up on it
"needs some tweaking"	2026	You end up doing the work
"still needs some tinkering"	2026	You end up doing the work
"could not get my call back number right, like 8 times"	2026	It mishears names and numbers

"I had to hang up"	2026	Callers hang up on it
"promising 300 guaranteed appointments at \$45 each"	2024	You pay. They do not deliver.
"claimed they couldn't fulfill their guarantees"	2024	You pay. They do not deliver.
"NO REFUND is in sight"	2024	You pay. They do not deliver.
"The tech is AMAZING. But the results IMO were terrible"	2024	It just did not pay off
"I couldnt ever get the ai agent to do what I wanted"	2025	It won't do what you tell it
"their system went down for a bit"	2025	It breaks and drops calls
"just wants to be transferred ... that costs \$4??"	2025	You pay. They do not deliver.
"calling 5 times. That'll cost you \$50 plus"	2025	You pay. They do not deliver.
"it was a fucking nightmare"	2025	It just did not pay off
"the only thing it did well was mimicking a human voice"	2025	It just did not pay off
"Outside of that it couldn't do shit"	2025	It just did not pay off
"customers hated it"	2026	Callers hang up on it
"the verbal portion ... does not translate well for our use"	2026	Callers hang up on it
"it occasionally double-books if two calls come in at the exact same time"	2026	The booking did not happen
"It took almost a full month just to get a basic meeting"	2025	You pay. They do not deliver.
"the quoted price was \$2,000 per month"	2025	You pay. They do not deliver.
"[vendor] really does not seem to care"	2025	You pay. They do not deliver.
"call latency is very very visible (~ >2s !!)"	2025	You ask. Then you wait.
"test call delay is upwards of 1 minute sometimes"	2025	You ask. Then you wait.
"the core product and latency is ridiculous"	2025	You ask. Then you wait.
"a few scheduled callbacks, ~9% conversion"	2025	It just did not pay off