VOICE AI IN PRODUCTION

No App, Just a Phone Call

Building LLM voice agents for people who don't trust AI

WHO'S TALKING

Pavel Shchekotov

Founding Engineer
Sophieconnects

Co-Founder
Denk Solutions

The Hypothesis

“What if onboarding was just answering a call?”

Instead of an app with a menu — a curated, concierge-like experience. The bet: this feels personal, increases retention, and keeps users involved.

Option A — voice AI inside the app the conventional path

Option B — voice AI over the phone **our bet ✓**

Why Phone?

Our users already use the phone every day. Nothing new to learn.

No app

Nothing to download or update

No sign-up

No account, no forms, no passwords

Just call

Dial the number — you're already inside

Service starts

We call back or text when needed

The Reality Check

Perception and trust hit us before any technology did.

People don't want to talk to robots
Many simply hang up on anything synthetic

The word “AI” creates resistance
Just knowing it's AI is enough to opt out

Unknown number = scam fear
A new service calling you feels suspicious

Accents break speech-to-text
Misheard users are frustrated users

Unexpected Human Behavior

Some users never fully realized they were talking to a bot. And then…

They went off script
Unrelated questions, out of nowhere

They shared personal stories
The bot became a confidant

They judged by the voice
Assumptions about who 'she' was

…and some men started flirting with the AI

Designing Trust: Meet Sophie

No face. No gimmicks. Every detail engineered to feel safe.

Calm voice

Neutral female, latest ElevenLabs — less robotic

Human name

“Sophie” humanizes the whole experience

Honest hello

We say it's AI up front — wording crafted carefully

No avatar

Zero visual reference — like Siri or Alexa

Caller ID

CNAM + “save our number” at onboarding

Memory

Each call builds on previous conversations

Scripts Are Product Design

ON SCREEN

“Hi! This is Sophie, your personal assistant. Quick heads-up before we start: I'm an AI — is that okay with you?”

~3 seconds to read

OUT LOUD

~10 seconds to listen

Every word matters when it is spoken aloud.

We replayed every phrase out loud, and built guardrails that pull off-script users back into the flow.

Phone Infrastructure Is Messy

Some of it you can engineer around. Some of it you just absorb.

Calls blocked by local providers
Sometimes there's nothing you can do

Landline callers
You can't text them back — dead end

Sudden drop-offs
Mid-call, unexplained, unrecoverable

Voicemail
Must be detected — every miss costs money

Latency, at least, has known fixes: closer infra, faster providers, local inference → ~500–700 ms end-to-end.

Voicemail Detection Was Not Solved

It looks like an easy problem. It isn't. We stacked three layers:

3 · LLM detection
the judgment layer

2 · Heuristics
“you've reached…”, long digit readouts

1 · Twilio AMD
built-in automated machine detection

Why it matters

Undetected voicemail = wasted money — and a broken experience for the humans waiting on the other side.

SMS: Tiny Messages, Big Problems

Hi, it's Sophie! Your match is ready 🎉 Tap to confirm your plan: pay.example.co/x8x2 — Reply STOP to opt out

159/160 · 1 segment · emoji counted as 2+

every character is money — punctuation included

Charged per segment, not per message

160 chars each — budget your punctuation

Payment links look scammy

Lowest delivery rate of everything we sent

Spam filters are ruthless

“STOP to opt out” required; dating content flags easily

The paywall is already the biggest drop-off point. Now imagine your payment message simply never arrives.

THE FINAL LESSON

Voice AI is not just about intelligence.

It's about trust, reliability, and cost.