



ISO 9001

NASSCOM


Clutch Top B2B

DPIIT

THE ENTERPRISE FIELD GUIDE · 12-18 MIN READ

Voice AI That Speaks Your Customer's Language. Natively.

A practical playbook for Indian enterprise leaders replacing legacy IVR with multilingual Voice AI agents. Hindi, Tamil, Telugu, Kannada, Marathi, Bengali, Gujarati, Punjabi, Malayalam — built by Swaran Soft, deployed in 3-4 weeks, owned by you.

 **Book a Discovery Call**

● 12-18 min read · 6 sections

- हिन्दी Hindi
- தமிழ் Tamil
- తెలుగు Telugu
- ಕನ್ನಡ Kannada
- मराठी Marathi
- বাংলা Bengali
- ગુજરાતી Gujarati
- ਪੰਜਾਬੀ Punjabi
- മലയാളം Malayalam
- + English

The Phone Line Is Still Your **Most Used** — and **Most Broken** — Channel.

Indian enterprises handle billions of customer calls each year. Most of them route through IVR trees and human agents that were designed for an English-speaking minority — while 88% of Indian internet users prefer a regional language. Here's what changed in the last 18 months.

67%

OF INBOUND LEADS LOST AFTER BUSINESS HOURS

High-consideration buyers — drones, equipment, financial services — call when they're ready. If you can't pick up, your competitor does.

40%

DROP-OFF ON ENGLISH-ONLY IVR IN TIER-2/3 CITIES

Regional buyers hang up before reaching menu option 3. Your CAC compounds with every abandoned call.

3.5 hrs

PER AGENT PER DAY ON REPETITIVE CALLS

12 questions account for 60-70% of all call volume. Your senior staff are working as FAQ readers.

Three Myths That Keep Boards From Moving

⚠ MYTH

"Voice AI sounds robotic — our customers will hate it."

✓ REALITY

Modern TTS (Sarvam, BharatGen, Dhvani) is human-quality in 9 Indian languages. Most callers can't tell.

⚠ MYTH

"AI can't handle thick Indian accents and regional dialects."

✓ REALITY

Accent normalisation + domain vocabulary tuning push STT accuracy above 95% — including code-mixed Hinglish.

⚠ MYTH

"Voice AI means cloud — we'd be sending customer audio offshore."

✓ REALITY

On-premise audio processing is now standard. DPDP-compliant. Zero audio leaves your VPC. **Swaran Soft deploys this by default.**

How We Got Here: The 25-Year Evolution of Enterprise Voice

2000s	2010s	2020	2023	2025+
Touch-tone IVR	Cloud Telephony	Voice Bots (English)	Multilingual Voice AI	Agentic Voice AI

Wondering where your contact centre fits on this curve? **Have a 20-min call with our team.**

How Voice AI Actually Works — In Plain English.

You don't need to be a machine learning engineer to make this decision. Here's the entire pipeline, broken down for a CXO.

From Phone Call to CRM Update — In Under 1.2 Seconds.

When a customer dials in, the call audio is first captured by your telephony provider — **Exotel, Twilio, or Knowlarity**. The audio stream is then handed to the **STT engine**, which transcribes it in real time.

Next, an **NLU layer** figures out the caller's intent. A small, fast LLM then composes the right response — pulling data from your CRM, ERP, or knowledge base where needed. Finally, **TTS** turns the response back into voice.

Every interaction is logged, sentiment-scored, and synced to your CRM. The AI escalates to a human agent with full context whenever confidence drops below your threshold.

THE ANALOGY

Think of it as a multilingual team member who never sleeps, never forgets context, speaks 9 Indian languages fluently, and never asks for a raise. Except this one is also available at 2 AM on a Sunday in Bhubaneswar.

Customer Calls



EXOTEL · TWILIO · KNOWLARITY



Speech → Text (9 Languages)



STT · ACCENT NORMALISATION · <300MS



Intent + Entity Extraction



NLU · ORCHESTRATOR · PLANNER



LLM + CRM/ERP Lookup



SARVAM · BHARATGEN · OLLAMA



Text → Voice Response



TTS · CONFIGURABLE TONE · SSML



CRM Updated · Outcome Logged



DPDP COMPLIANT · ON-PREMISE

WHAT THIS MEANS FOR YOU

Your **3.5 hours/day per agent** spent on repetitive calls collapses to under 30 minutes — handled by AI.

Customer audio **never leaves your VPC**. Full DPDP, RBI, and HIPAA-aligned compliance by default.

Every call now contributes **structured data** to your CRM — intent tags, sentiment scores, next-best-action.

See the pipeline live

Book a demo and hear a Voice AI agent handle a real call in Hindi, Tamil, or your language of choice.

[Book a Discovery Call →](#)

[Voice AI Overview](#)

5 Voice AI Use Cases **Already Paying For Themselves** in Indian Enterprises.

These are pulled directly from Swaran Soft deployments and adjacent industry patterns.

TELECOM · FIELD SERVICE

Voice AI Dispatcher for Field Engineers

● MEDIUM COMPLEXITY · 3-4 WEEKS

▲ PROBLEM

A telecom NOC detects an incident at 11 PM. Human dispatcher manually calls 7 engineers, in English, hoping someone picks up. Average dispatch time: 4+ minutes. Engineers from Tier-3 towns are skipped because they "prefer regional language".

⚙️ VOICE AI AGENT

Auto-dials engineers in Hindi or their preferred regional language. Reads job details verbally, confirms availability, logs acceptance, and updates the ticketing system — all without a human in the loop.

↑ BUSINESS IMPACT

73% MTTR reduction. 3 dispatcher FTEs redeployed to escalation handling. 100% language preference match. Engineers report a 22-point NPS jump because they're spoken to in their own language.

73%

MTTR REDUCTION

45s

AVG DISPATCH TIME

-3 FTE

DISPATCH HEADCOUNT

HEALTHCARE · PATIENT OPS

Outbound Appointment Reminder Agent

● QUICK WIN · 2-3 WEEKS

▲ PROBLEM

A multi-specialty chain runs 34% no-show rates. SMS reminders are ignored. Manual calling teams can't keep up with 1,200 daily appointments across 9 languages.

⚙️ VOICE AI AGENT

Calls patients 24h and 2h before appointment in their preferred language. Handles "I need to reschedule" conversationally, updates the HIS, and sends a WhatsApp confirmation. Escalates anything medical to a human.

↑ BUSINESS IMPACT

No-show rate cut from 34% to 8%. 100% of reschedules handled without human intervention. Patient satisfaction NPS up 22 points. ROI in under 8 weeks.

-76%

NO-SHOWS

+22

PATIENT NPS

100%

RESCHEDULES AUTOMATED

BFSI · COLLECTIONS

Empathetic EMI Reminder Agent

● HIGHER COMPLEXITY · 4-6 WEEKS

▲ PROBLEM

Collections calls are aggressive, repetitive, and damage brand trust. Human callers can't consistently maintain empathetic tone across 400 daily calls. RBI compliance is inconsistent.

⚙️ VOICE AI AGENT

Calls in customer's language with calibrated empathy. Offers EMI deferral options, handles dispute conversation, escalates complex cases with full context. Every call is RBI-compliant and audit-trail-ready.

↑ BUSINESS IMPACT

+18% collection rate improvement. Human escalation kept under 12%. 100% call compliance logging. NPS up 38 points because customers actually feel heard, in their language.

+18%

COLLECTION RATE

+38

NPS IMPROVEMENT

<12%

ESCALATION RATE

RETAIL & D2C · INBOUND

Multilingual Order & Returns Agent

● QUICK WIN · 2-3 WEEKS

▲ PROBLEM

D2C brand fields 600+ daily calls about order status, returns, and pricing — 80% in Hindi or regional languages. Wait times exceed 6 minutes. CSAT drops every weekend when staffing thins.

⚙️ VOICE AI AGENT

Authenticates caller by phone number, pulls order from Shopify/Magento, handles status, returns, and refunds end-to-end. Switches between Hindi-English mid-call without breaking conversation flow.

↑ BUSINESS IMPACT

Wait time dropped from 6m to under 30s. 78% of calls resolved without human handoff. CSAT up 19 points. Weekend staffing reduced to skeleton crew. **ILLUSTRATIVE EXAMPLE**

-95%

WAIT TIME

78%

SELF-RESOLVED

+19

CSAT LIFT

MANUFACTURING · AFTER-SALES

AMC Renewal & Service Booking Agent

● MEDIUM COMPLEXITY · 3-5 WEEKS

▲ PROBLEM

Equipment AMC renewals slip because no one calls customers in time. Service bookings come through 17 channels and get lost. Field engineer scheduling is chaos.

⚙️ VOICE AI AGENT

Outbound calls 30 days before AMC expiry in the customer's language. Books service slots conversationally, dispatches engineers via the field app, and pushes everything into SAP or your ERP.

↑ BUSINESS IMPACT

AMC renewal rate up 31%. Service booking lead-to-visit time down 40%. Field engineer utilisation up 22%. **ILLUSTRATIVE EXAMPLE**

+31%

AMC RENEWALS

-40%

BOOKING TIME

+22%

ENGINEER UTILISATION

25+ years. 350+ global clients. India's leading Agentic AI partner.

Swaran Soft has been building enterprise systems since 2001 — long before "AI" was a buzzword. Our voice agents run in production for telecom, healthcare, BFSI, and manufacturing leaders across India, the UAE, Europe, and the US. Open-source stack. Zero vendor lock-in. Yours to own.

Which use case fits your contact centre?

Tell us your call volume and languages. We'll tell you what's feasible in 4 weeks.

[Submit Your Use Case →](#)

[View Case Studies](#)

05 · AVOID THESE TRAPS

5 Voice AI Pitfalls That Kill Deployments — And How We Prevent Them.

We've seen these failures repeatedly across telecom, BFSI, and healthcare deployments. None are technical limitations — all are planning failures. Here's the playbook.

⚠ Treating Voice AI as a Cost-Cutting Project Instead of a CX Project

WHAT GOES WRONG

Procurement scopes voice AI as "agent headcount reduction" → minimum-viable language coverage → robotic voice → angry customers → board questions the ROI by month 4.

WHY IT HAPPENS

Finance teams want the cost-saving line item. CX teams aren't in the buying conversation.

HOW SWARAN SOFT PREVENTS IT

Our 45-min discovery starts with the CX outcome — NPS, CSAT, FCR — not the FTE saving. Cost reduction is a downstream effect, not the headline metric.

🌐 Shipping a Generic English Voice Bot and Calling It "Multilingual"

WHAT GOES WRONG

Vendor adds a Google Translate layer on an English bot. Translations are stilted, accents fail, dialect words are missed. Customers hang up faster than they did on IVR.

WHY IT HAPPENS

Most global voice AI vendors retrofit Indian languages onto English-first models. Voice quality, intent detection, and domain vocabulary are all afterthoughts.

HOW SWARAN SOFT PREVENTS IT

We build on India-first models — Sarvam AI, BharatGen, Bhashini, Dhvani — trained on Indian language data natively. Domain vocabulary tuning is included for your industry.

📡 Sending Customer Audio to Foreign Cloud APIs

WHAT GOES WRONG

Voice AI pipeline routes raw call audio to US-based STT/TTS APIs. The DPO finds out in a quarterly audit. The project is paused. The vendor disappears.

WHY IT HAPPENS

Cheap voice AI vendors don't have on-premise pipelines. Cost optimisation overrides compliance until someone in legal asks the right question.

HOW SWARAN SOFT PREVENTS IT

All audio processing happens on your VPC or our DPDP-aligned India data centres. Configurable retention. Audit-trail-ready by default. Private AI Lab available for full on-premise.

No Clean Human-Handoff Strategy

WHAT GOES WRONG

AI confidently answers questions outside its scope, or transfers calls to humans with zero context — so the human re-asks everything. Customer thinks they're talking to two stupid systems.

WHY IT HAPPENS

Confidence thresholds aren't tuned. Warm transfer with context isn't built. Vendor demoed the happy path only.

HOW SWARAN SOFT PREVENTS IT

Every deployment includes calibrated confidence thresholds, warm transfer with full conversation history, and weekly performance reviews for the first 30 days post-launch.

Voice Data Never Flows Back to CRM, BI, or Strategy

WHAT GOES WRONG

Voice AI runs, calls get handled, but nothing structured ever reaches the CRM. Sales and marketing remain blind to what customers are actually saying on calls.

WHY IT HAPPENS

Vendor scoped voice AI as a standalone product, not a data flywheel. CRM integration was a "phase 2" item that never came.

HOW SWARAN SOFT PREVENTS IT

CRM/ERP integration is non-negotiable in week 1. Every call generates intent tags, sentiment scores, entity extractions, and next-best-actions — landing in Salesforce, Zoho, Freshdesk on day one.

Avoid these pitfalls from day one

Our structured 3-4 week deployment prevents every one of these traps.

[Start Your Voice AI Pilot →](#)

[Voice AI Services](#)

THE SWARAN SOFT PLAYBOOK: 5 BEST PRACTICES

Start With One Use Case That Has Measurable CX Impact

THE PRACTICE

Pick one high-volume, narrow-scope use case — appointment reminders, AMC renewals, EMI follow-up — and instrument it with CX metrics from day zero.

WHY IT WORKS

Narrow scope = clean training data = high accuracy = fast wins = board buys the next phase.

Pick Indian-First Models, Not Translated English Bots

THE PRACTICE

Use Sarvam, BharatGen, Dhvani, Bhashini as the foundation. Avoid retrofitted translations from English LLMs.

WHY IT WORKS

Native models handle code-mixed Hinglish, dialect variation, and cultural context that English-first systems miss completely.

Make Data Residency a Day-One Architecture Decision

THE PRACTICE

Architect for on-premise or VPC-based audio processing from week 1. Configure DPDP-aligned retention. Build the audit trail before you build the bot.

WHY IT WORKS

Retrofitting compliance is 10x more expensive than building for it. Bank, hospital, and government deployments all live or die on this single decision.

Design Human Handoff Like a First-Class Feature

THE PRACTICE

Define confidence thresholds per intent. Build warm transfer with full transcript. Have human agents trained on AI-handoff etiquette.

WHY IT WORKS

Customers don't mind talking to AI — they mind repeating themselves. Context-preserving handoffs lift CSAT more than perfect AI ever could.

Treat Voice Data as Strategic Asset, Not a Compliance Cost

THE PRACTICE

Pipe intent tags, sentiment scores, and entity extractions into your CRM and BI from day one. Build dashboards before launch.

WHY IT WORKS

Voice data tells you what customers want — and what they're about to churn over — before they ever email you. It's the cheapest VOC research you'll ever buy.

THE SWARAN SOFT VOICE AI METHODOLOGY

01

CX-First Discovery

45-min discovery focused on the CX outcome — NPS, FCR, CSAT — not on FTE reduction. We define success in your customer's words.

02

India-Native Build

Built on Sarvam, BharatGen, Bhashini, Ollama. On-premise audio processing. DPDP-compliant by default. Open-source stack you own.

03

Production From Day 1

No POC graveyards. We ship a production-grade voice agent in 3-4 weeks with full CRM integration, monitoring, and 30-day SLA.

From legacy IVR to conversational AI

Our methodology has delivered voice agents for telecom, healthcare, BFSI and manufacturing leaders across India.

[Book a Discovery Call →](#)

[Voice AI Overview](#)

From IVR to Conversational AI in 3-4 Weeks.

A structured deployment path — no surprises, no scope creep.

WEEK 01

Telephony Integration & Language Config

Connect Exotel/Twilio/Knowlarity. Configure language models. Map call flows.

WEEK 02

FAQ Training & Intent Mapping

Train on your catalogue, FAQs, pricing, escalation triggers. Map intents to CRM fields.

WEEK 03

UAT With Your Team

Real-call testing with sales and support. Tune confidence thresholds. Validate CRM flow.

WEEK 04

Go-Live & Monitoring

Production deployment. Real-time dashboard. 30-day SLA-backed support. Weekly reports.

"We replaced our English-only IVR with Swaran Soft's voice agent in 4 weeks. Within 60 days, our after-hours conversion rate doubled and our dispatcher team moved to higher-value escalation work. The fact that it runs on our own infrastructure was the deciding factor for compliance."

VP, Customer Operations · Tier-1 Telecom Operator · India

COMPOSITE EXAMPLE · DRAWN FROM MULTIPLE PRODUCTION DEPLOYMENTS

Ready to Replace Your IVR? Let's Talk.

Book a Discovery Call

45-min, no-obligation. We'll map your call volume, languages, and integration points, and tell you exactly what's feasible in 4 weeks.

[Schedule Call](#)

Submit Your Use Case

Tell us your call volume, languages, and current pain. We'll send a tailored pricing sheet and a sample voice AI call in your target language — within 24 hours.

[Get Pricing Sheet](#)

Start your Voice AI journey today

From IVR to conversational AI in 3-4 weeks — 350+ enterprise transformations across 25+ years.

[Book Free Consultation →](#)

[Explore Voice AI](#)