The Readiness Gap

Why Voice AI pilots in India stall

EkStep | Voice AI



By: Santosh Kevlani, Manmath Goel,
Pranava Hiremath, Ananya Shrivastav

September 2026

Contents

1. Executive summary..... 03
2. Barriers to Voice AI adoption..... 05
3. Know-When: When to start a Voice AI pilot..... 07
4. Know-How: The six capabilities and how to build them.... 09
    4.1 Outcome definition.....
    4.2 Vendor evaluation.....
    4.3 Conversation design.....
    4.4 Testing and QA.....
    4.5 Workflow integration.....
    4.6 Change management .....
5. Voice AI Readiness Self-Assessment..... 22
6. What a successful pilot design could look like..... 24
7. Conclusion..... 25

1 Executive Summary

India's rapid digitisation story receives frequent praise. Getting to 1.4 billion digital identities, 750 million active smartphone users and 25 billion real-time UPI transactions per month is not a small feat for a country at India's development stage. However, the headline print glosses over an uncomfortable reality.

The adoption of technology is still not universal. Around 25 percent of the population cannot operate a smartphone, which excludes them from services delivered through an app or a screen. 20 percent of Indians converse in regional languages and dialects that are not officially recognised - Bhili, for instance, spoken by around 10 million people, is not a scheduled language and is supported by almost no digital service.

A technology like Voice AI can ensure that the benefits of digitisation truly reach the last mile. It does not require access to smartphone or savviness with digital applications, only a phone to make and receive calls. By also supporting low-resource languages,Voice AI can extend distribution of technology-driven service to speakers who are excluded from existing interfaces.

Voice AI makes particular sense for the impact sector. Non-profits, social enterprises and government bodies can deliver information and gather feedback faster with Voice AI than with a resource-constrained field team. A citizen-feedback helpline, for instance, can capture open-ended answers instantly that would have taken multiple SOPs and several weeks for a human call centre. For impact organisations, this changes the economics of last-mile delivery.

The technology is proven at national scale. MahaVISTAAR, Maharashtra's state farm-advisory line, answers real-time questions on weather, crop stress and market prices for more than 2.5 million farmers across Maharashtra in Marathi & Bhili, available on both telephony and app.

Sarlaben, Amul's voice assistant, supports 4 million farmers in Gujarat on cattle health and breeding, many of them women who have never owned a smartphone.

But Voice AI adoption is still lagging its true potential. While awareness is yet to catch up, even organisations who are actively interested in adoption are struggling to scale Voice AI deployments. This is despite a thriving ecosystem of more than a hundred Voice AI startups and a competitive cap on pricing.

For instance, EkStep Foundation ran two pan-India programmes enabling Voice AI adoption, giving over 50 impact organisations access to subsidised minutes. Even with that pricing support, less than 30% could scale beyond pilots — a reminder that low pricing alone isn't the constraint; scaling demands far more than cost relief, the harder constraints lie elsewhere.

Through in-depth interviews with 40 participating organisations and other adopters/ builders, organisational readiness was revealed as an equally important barrier to adoption as technology and pricing.

Organisation readiness emerged as a continuum of six capabilities, most of them non-technical:

1. Outcome definition
2. Vendor evaluation
3. Conversation design
4. Testing and QA
5. Workflow integration
6. Change management

The report explains each of these through real world examples of piloting organisations and inputs from builders, investors and the widerVoice AI community.

The report is chiefly intended for adopters - especially those in the impact sector - who may be exploring their own Voice AI journey. By leveraging the lived experience of similar organisations and a self-assessment scorecard contained in this report, organisations can build and bolster their own readiness to deploy Voice AI.

2 Barriers to Voice AI adoption

Voice AI bridges a natural gap in a country like India where only 38 percent of households are digitally literate and an estimated 350 million Indians still use basic feature phones. The impetus to leverage Voice AI is particularly strong for impact organisations that target precisely this segment and are resource constrained to scale distribution.

The supply side in the Voice AI ecosystem is evolving rapidly. More than 100 Indian startups now operate across the stack, from speech recognition and language models through orchestration and telephony to finished applications. Per-minute pricing has compressed to as low as ₹1-3, and average round-trip latency is <800 milliseconds. Use cases in financial services and lead qualification are seeing active interest forVoice AI solutions.

However, large-scale adoption has not followed. In conversations with builders and investors across the ecosystem, most estimate that fewer than 10 percent of enterprise call volumes currently go through Voice AI — with adoption even more uneven in the impact/MSME sector.

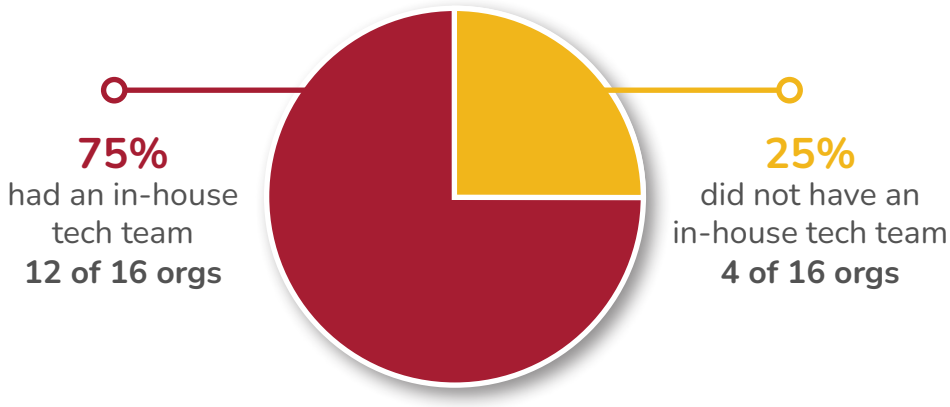
While product and pricing challenges are talked about routinely – higher pricing for complex use cases or product failures especially for code-mixed, regional dialect speech – non-tech barriers are less so. Through over 40 structured interviews with adopters across six sectors, three webinars and one survey of 16 organisations, organisation readiness emerged as an equally important barrier to Voice AI deployments. The bottleneck to adoption is relatively non-technical - despite having internal tech teams, a majority of adopters struggle with deployment.

Not every barrier to Voice AI adoption is within an adopter's control — but the research shows that a significant share of it is, concentrated in two areas: the know-when — recognising when to considerVoice AI in their journey — and the know-how — building internal readiness alongside the

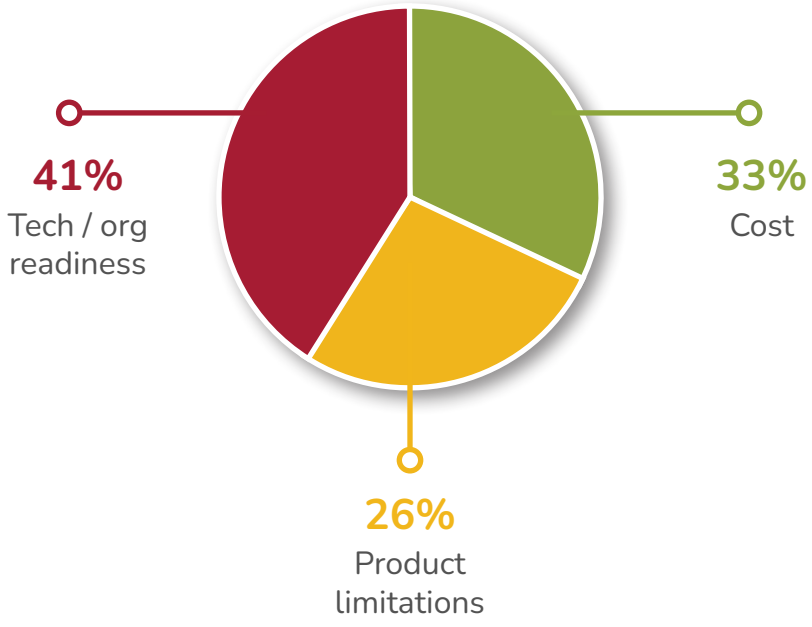
technology itself. These are the two components this paper focuses on, since they're the ones adopters can actually act on. It closes with a self-assessment scorecard adopters can use to assess their current and required readiness to deployVoice AI in their workflows.

Despite having in-house tech teams, organisations cited tech/org readiness as the biggest barrier.

IN-HOUSE TECH TEAM



BARRIERS CITED



3 Know-When

When to start a Voice AI pilot

Voice AI does not suit every use case, and matching the technology to the problem is the first decision an organisation must make. An early-childhood education NGO in the sample joined a Voice AI programme before deciding what problem it wanted to solve. Its use case, prompting field workers to upload attendance data on time, could have been achieved with a simple, scheduled SMS rather than a complex Voice AI build. This was a common experience across the sector – where the excitement about new technology exceeded the know-how to imbibe it.

The following ladder helps place a use case before vendor shortlisting. Build effort increases from top to bottom.

THE USE-CASE READINESS LADDER

		BUILD EFFORT ↓
L0 Static broadcast	One-way pre-recorded messages, no interaction	Voter-awareness calls, vaccination alerts, exam-date reminders
L1 Basic feedback	Simple outbound; Yes/No or numeric capture, no verification	"Press 1 if ration received"
L2 Verified feedback	Inbound and outbound; quantitative plus short qualitative, verified data	Scheme-eligibility checks, lead qualification
L3 Conversational intelligence	Context-aware; sentiment and trend analysis, multi-channel	Citizen helpline, customer service
L4 Integrated workflow	Calls trigger real actions and connect to live systems	Route a grievance, check application status mid-call, trigger escalation

Levels 0 and 1 can often be met without Voice AI. Where a broadcast or a form-fill resolves the requirement, an IVR or an SMS may deliver it faster and at lower cost.Voice AI earns its cost from Level 2 upward, where a two-way exchange is the objective.

WHAT A PILOT COSTS



Team

1 technical/ IT staff member and 1 non-technical programme manager[1]



10–20K

minutes to target for most pilots in the impact sector



₹20–60K

for a 3-month pilot at ₹2-3 per minute, excluding team costs and bandwidth

1. Applies to L2 pilots (verified feedback). L3 and L4 use cases — conversational intelligence and integrated workflows — typically demand more technical resourcing beyond this baseline team.

A Voice AI deployment also requires upfront investment and dedicated team bandwidth. At a minimum, this would include a team of at least 1 technical/ IT staff member and 1 non-technical programme manager to drive technical integration and programme rollout respectively. Most pilots in the impact sector should target anywhere between 10,000 and 20,000 minutes with an estimated cost of ₹2 to ₹3 per minute. This implies a cost of anywhere between ₹20,000 and ₹60,000 for a 3-month pilot excluding team costs and bandwidth.

4 Know-How

The six capabilities and how to build them

Across the organisations studied, organisation readiness can be thought of as a continuum of six capabilities to be built and bolstered during and after a deployment:

01	02	03	04	05	06
Outcome definition	Vendor evaluation	Conversation design	Testing and Q&A	Workflow integration	Change management

Most organisations across the 40+ sample had a gap in one or more capability that prevented Voice AI pilots reaching their intended scale and outcome.



4.1 OUTCOME DEFINITION

The first capability is to define what success looks like and the metrics to track it. Without this critical first step, a pilot will lack evidence without a clear determinant for a go/ no go decision. Organisations in the sample underestimated this step, joining the programme and launching a pilot without defining success outcomes.

Case · Agritech platform, Biomass sourcing

The agritech platform working on biomass sourcing is the clearest counterexample. Their technical lead – a quantitative researcher - defined clear success metrics (connect rate, talk-time distribution and lead conversion) at the outset of the pilot. This was used to design an A/B test to compare calls that included price and quantity information against those that did not and eventually scale deployment to 5000–10000 farmers.

Across 1,000 cold outbound unique calls placed, the results were:

35%	Connect rate: 35% — meaningful for cold rural outreach with no prior relationship, no opt-in, and a single attempt
8–10%	8–10% of answered calls ran past 90 seconds —not simple yes/no responses but farmers speaking unprompted about crop cycles, storage constraints, and pricing expectations, providing qualitative intelligence that field staff would have taken weeks to collect per cluster
1.5%	Lead conversion: 1.5% — against a prior baseline of zero, since the organisation had never been able to reach farmers at this volume before. At 5,000–10,000 calls planned for the next round, even a 1.5% conversion rate produces a viable supply pipeline
50 days → 2 wks	Outreach cycle: reduced from 50 days to 2 weeks — the most direct like-for-like comparison with the previous model



The contrast case is a large education NGO running an internship verification campaign across 49,000 students. The technology performed — 74 percent of targeted students received a call. But no data quality threshold had been defined in advance, and the conversation flow had not been sequenced to reach its purpose efficiently. The result: only 3 percent of students produced fully usable data. As the programme director reflected:

> The confirmatory questions were too many. We went on to ask everything and then went on to the real survey. I think that's where people lost interest — because otherwise the call pickup and response was already there.
>
> **PROGRAMME DIRECTOR, EDUCATION NGO**

The difference between these two organisations was a definition of clear success and therefore, a Voice AI workflow designed to move a specific and measurable metric.

4.2 VENDOR EVALUATION

A vendor should be tested against the organisation's own language, dialect, and use case, rather than a demonstration or a referral. Evaluation failures recurred due to similar oversight.

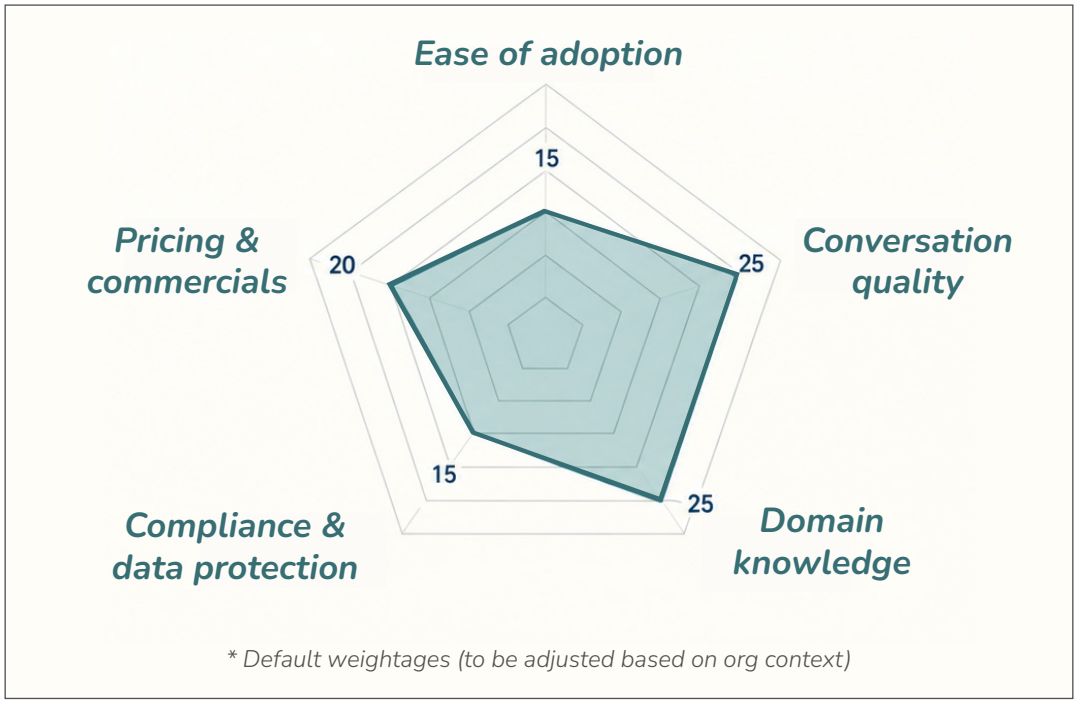
Case · Worker-Matching Platform & Social-Welfare Access NGO

A blue-collar worker-matching platform compared two vendors and selected the one with the stronger analytics dashboard. Its core users spoke in dialects from Uttar Pradesh, Bihar, and Jharkhand. However, the organisation failed to include dialect coverage as part of their vendor selection. Similarly, a social-welfare access NGO chose a vendor primarily on the responsiveness of its sales team, without indexing on the accuracy of the Hindi-English-Marathi code-switching. Both processes set the pilot back by a few weeks that could have been avoided by an upfront, methodical process to select the right vendor.

> It's not about the tool. It's about whether the vendor actually sits with you and figures out how to make things work. We had enough back-and-forth with our Voice AI vendor to figure out how certain things could work.
>
> **TECH LEAD, CIVIC TECH NGO**

A scored evaluation against a fixed set of criteria addresses this by making every factor explicit and weighted:

A SCORED VENDOR EVALUATION, FIVE CRITERIA



15% Ease of adoption and integration

- What is the deployment time/effort?
- What must we have ready on our side?

25% Conversation quality

- Which LLM models are used?
- How is interruption handling and latency?

25% Domain knowledge

- What does post-launch support look like?
- Does the vendor have a track record in a similar sector

15% Compliance and data protection

- What encryption standards does the vendor follow?
- Who owns the data, and for how long?
- Is the data stored on-premise or cloud?

20% Pricing and commercials

- Is pricing per minute or per task completion?
- Any upfront/ maintenance cost or recurring tech-support cost?

The weighting matters because the strongest demo does not always indicate the best fit. In a worked comparison from the research, scoring two vendors equally across the five criteria, the vendor with weaker conversation quality scored higher overall on the strength of its ease of adoption and operational support.

CRITERION (EQUAL 20% WEIGHT)	VENDOR A	VENDOR B
Ease of adoption	2	4
Conversation quality	5	2
Domain experience and ops support	2	5
Compliance and data protection	4	4
Pricing and commercials	3	3
WEIGHTED TOTAL (OUT OF 5)	**3.2**	**3.6**

Operational fit outweighed a stronger demonstration, which is the more common outcome for organisations with limited engineering capacity to absorb a harder integration.

4.3 CONVERSATION DESIGN

Many organisations handed over a broad script to the vendors vs. thinking through conversation design, the decision branches, the fallbacks and the handoff to a human. This resulted in avoidable frustration and a pause in momentum.

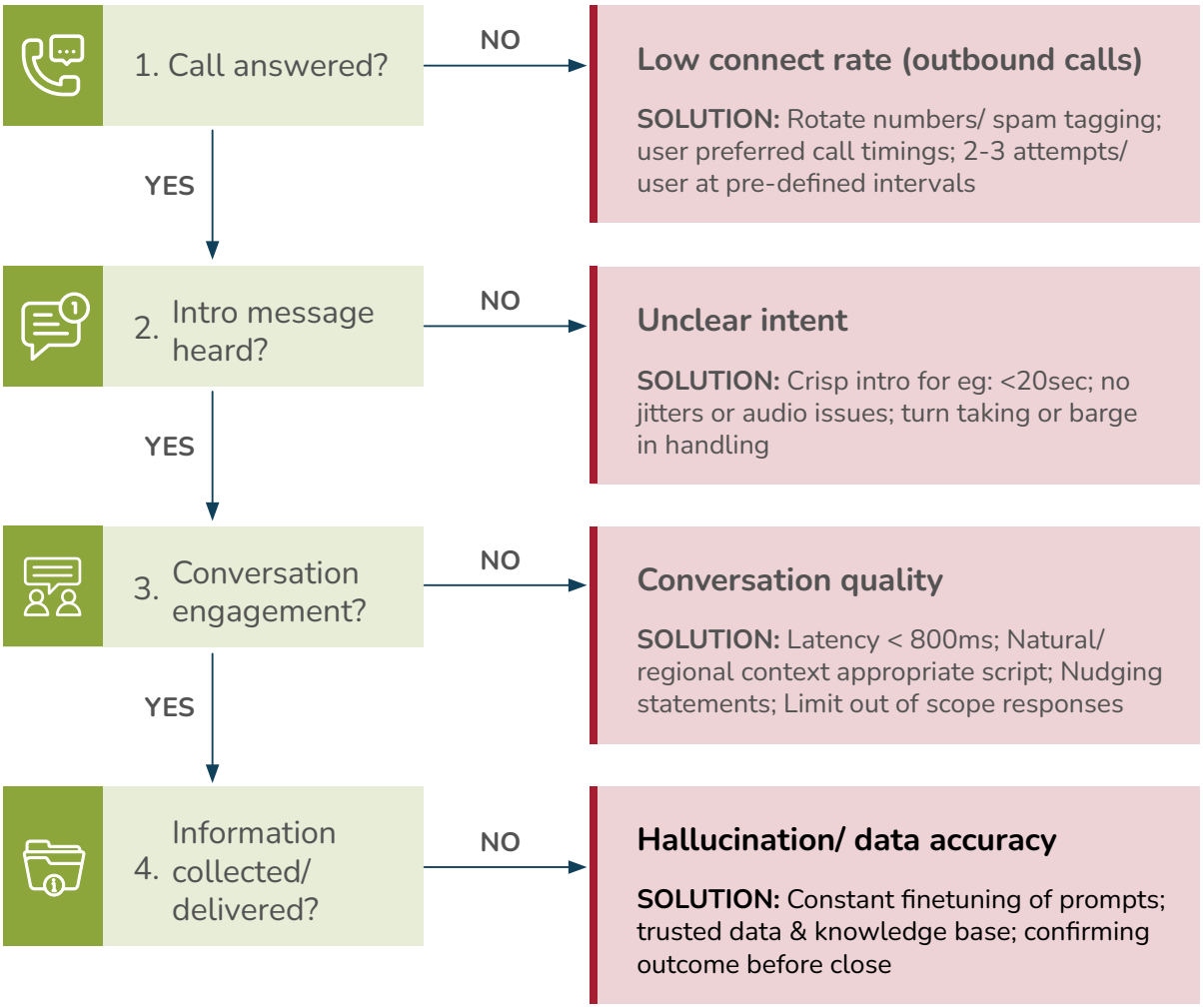
Case · Fintech Home-Loan Marketplace

A fintech home-loan marketplace learned this by rebuilding. Its first bot tried to qualify every kind of lead the business generated. It misqualified 80 percent of them as interested, and calls ran only 20 to 25 seconds, too short to finish a qualification script.

The team scrapped it and built four single-purpose bots instead, one each for bulk qualification, a housing-partnership channel, feedback and hiring. Narrowing each bot to a single job raised accuracy. The clearest gain came when the bot called warm, opted-in housing leads in Marathi, where average talk time rose from around 40 seconds on cold leads to six minutes and two loan cases closed within 10 to 12 days.

> You are forced to say yes. That's what we call in the creative business a lean-in approach. We have principles called the yes ladder. If the person says yes three times, they have an 84% higher chance of listening to why you called.
>
> FOUNDER, TECH IMPLEMENTATION ORG

WHERE CALLS DROP OFF, AND WHAT FIXES EACH STAGE



A bot with a long decision tree performs worse than a bot with a short one, so a bot that does one thing well beats a bot that does everything adequately.

Case · Citizen-Feedback Platform

A citizen-feedback platform approached conversation design systematically. First, they asked callers upfront when they were free to talk thereby raising call connect rates from 35 to 66 percent. Second, they were able to lock in a caller's preferred language on first contact that prevented code-switching glitches. Finally, splitting the conversation flow in two – a short structured path for general citizens and an open-ended one for policy experts, gave each group a suited length, 3 to 4 minutes against 10 or more. Cumulatively these changes helped the platform reach 18 times as many people as it had in the seven prior years combined, 400,000 to 500,000 calls in a single cycle.



35 → 66%

call connect rates, after asking callers upfront when they were free to talk



18x

as many people reached as in the seven prior years combined



400-500K

calls in a single cycle

SIX CONVERSATION DESIGN PATTERNS

PATTERN	WHAT IT DOES	EXAMPLE SCRIPT LINE
Opening	Establish who is calling and secure a first yes.	"Hi — am I speaking to XYZ? This is a two-minute call about your ration entitlement."
Clarification	Recover a partial or code-mixed answer without restarting.	"Sorry, did you mean the September instalment or the October one?"
Confirmation	Read the captured value back before moving on.	"So that is 4 quintals, collected on Tuesday. Correct?"

PATTERN	WHAT IT DOES	EXAMPLE SCRIPT LINE
Handoff	Move to a human at the point the bot cannot proceed.	"I will connect you to an officer now. Please stay on the line."
Error recovery	Exit cleanly after two failed attempts instead of looping.	"I am not able to hear you clearly. We will call back tomorrow morning."
Closing	State what happens next and by when.	"Your request is registered. You will get an SMS within two days."

4.4 TESTING AND QA

Building a robust QA process emerged as a non-negotiable in most conversations. The organisations that reached scale treated QA as a continuous, iterative function vs a one-off exercise.

Case · Mortgage Finance Marketplace & Staffing Firm

A mortgage finance marketplace put one dedicated person full-time on call QA, benchmarking and prompt iteration, running several cohort experiments a week. As a result, its false-positive rate fell from 80 to 35 percent across several months of that work, not through a single fix. A staffing firm ran a lighter version, a small team spending one to two hours a day on manual call audits, catching flow failures before they compounded across thousands of calls.

> You need one dedicated person monitoring calls every day. Making the calls can be automated. The harder part is checking whether the bot handled them well. Someone still needs to review the calls and verify what happened.
>
> PRODUCT MANAGER, FINTECH HOME LOAN PLATFORM

A two-phase process can keep this process manageable at scale. The first phase is a quick scan, the second a deep audit. The deep audit produces a quality score that is trackable across model updates, languages and configurations.

	QUICK SCAN	DEEP AUDIT
PURPOSE	Sort calls into fine and needs-review	Diagnose what failed, where and how seriously
SAMPLE	A large batch, 50 to 100 calls	The 10 to 20 percent flagged, plus a few good calls to compare
TIME PER CALL	One to two minutes	As long as diagnosis takes
LOOKS FOR	Missing consent, policy or bias issues, dropouts, breakdowns, repeat callers	Failures across speech recognition, model response and spoken output
SCORING	Pass or fail	Each issue rated critical, major or minor, call scored 1 to 5
AUTOMATION	Largely automatable	Partly automatable

Without QA, a small error ends up being compounded across the whole deployment. A pilot that looks like it is working can generate unusable data across its entire call volume.

4.5 WORKFLOW INTEGRATION

Unlike traditional software, a voice agent rarely operates in a silo but is in fact part of the overall workflow of an organisation. For true efficiency gains, Voice AI must be connected to the channels, databases and follow-up systems an organisation already runs and failure to do so can make the entire pilot redundant.

Case · Skilling Enterprise, Rural Water Organisation & Agritech Platform

A skilling and mentorship enterprise built a bot that asked students questions rather than supplying answers. The vendor had promised a browser-based web SDK that was not available in production, and the enterprise lost roughly 40 days rebuilding the deployment around a different architecture that routed students through Telegram and IVR. Similarly, a rural water organisation piloting a voice assistant could not reach a go/no-go decision at all, because the WhatsApp follow-up integration was never completed and the team could not test the full user journey end to end.

On the other hand, an agritech platform matched each channel to the task it suited. Voice handled the initial outreach call, which is personal and harder for a farmer to ignore. WhatsApp handled the operations that followed, including weight confirmations, payment alerts and dispute resolution. The organisation planned the entire workflow at the outset which was a key determinant of success.

> As soon as the call is done, the data comes to our backend and is passed to the CRM. Simultaneously we are building a WhatsApp flow where this data is also shared with the WhatsApp vendor. End to end — all of this was set up in one and a half months.
>
> **GROWTH PM, HEALTH INSURANCE PLATFORM**

A Voice AI agent rarely delivers standalone outcomes and adopters should spend time thinking through and embedding it into their overall workflow.

4.6 CHANGE MANAGEMENT

The cases available suggest that a voice deployment depends on the people expected to use it. This includes both the internal staff as well as external users of the product.

Case · State Government Department

A state government department was an example of the former. A senior official championed the Voice AI pilot internally, and a dedicated consultant team handled day-to-day problem definition, campaign design, and agent scripts without waiting on the procurement cycle. The resulting campaign rolled out to Maize farmers produced crop data that informed a state policy report and a larger budgetary allocation to the programme. The department attributed this to the survey output. At its peak, the programme made 3 lakh calls per day. What made this possible was internal ownership rather than any feature of the technology.

Case · Edtech Organisation

An EdTech organisation working with schools worked on resistance from users of the technology. Teachers who were a part of the programme viewed AI as a threat to their job security. The organisation decided to position Voice AI as a diagnostic aid to identify the concepts individual students find difficult – a genuine discovery tool for teachers vs. a teacher replacement tool. The different positioning is what made the tool stick with teachers resulting in higher adoption.

> “When voice AI was first introduced for candidate screening, staff raised concerns about job security and whether the technology would replace their roles. Those concerns largely subsided once the system was integrated into day-to-day workflow and people could see that the AI was handling the parts of the job nobody wanted — repeat calls to unresponsive candidates — while freeing them to focus on placed candidates and relationships.”
>
> **GROWTH PM, HEALTH INSURANCE PLATFORM**

Top down leadership buy-in and effective communication to the wider team are both equally critical in ensuring org wide adoption.

5 Voice AI Readiness Self-Assessment

The six capabilities can be set out as a short diagnostic for an organisation to complete before it commits budget or selects a vendor. Each capability is scored from 1 to 5, where 1 indicates the capability is absent and 5 indicates it is an established practice. The scores are summed to give a total out of 30.

#	CAPABILITY	THE QUESTION TO ASK	FIRST ACTION IF THE SCORE IS LOW	SCORE
1	Outcome definition	Has the organisation documented the problem, the caller, and the metrics that would confirm the pilot succeeded?	Prepare a one-page brief stating the problem, the call type, the target population, and three success metrics. If a broadcast or a form would address the problem, reconsider whether voice is warranted (Section 2).	xx/5
2	Vendor evaluation	Have shortlisted vendors been tested on the organisation's own language, dialect, and use case, rather than a demonstration or referral?	Run a paid or trial call batch with each shortlisted vendor on the organisation's actual audience, and score the results against the five criteria in Section 3.2.	xx/5
3	Conversation design	Has the organisation scripted the flow, the branches, the fallbacks, and the human handoff itself?	Map the call as a funnel (Section 3.3), identify the stage at which callers currently drop off, and script that stage first.	xx/5

#	CAPABILITY	THE QUESTION TO ASK	FIRST ACTION IF THE SCORE IS LOW	SCORE
4	Testing and QA	Is one named person responsible for reviewing calls and iterating the bot as a continuing responsibility, not a launch-day task?	Assign one person to the two-phase QA process (Section 3.4) and validate on a small clean batch before scaling.	xx/5
5	Workflow integration	Has every vendor-promised integration been confirmed to run in production before signing?	Require a live demonstration of each integration on the organisation's own systems as a condition of contract.	xx/5
6	Change management	Have the affected staff and users been informed what the tool is for, and what it is not?	Identify who is affected, address the concern directly, and state the tool's purpose to them before deployment.	xx/5
TOTAL OUT OF 30				xx / 30

READING THE TOTAL



6–14

NOT READY

Build the two weakest capabilities before committing budget or approaching vendors.



15–22

VIABLE WITH CONDITIONS

A pilot can start, provided the low-scoring capability is assigned an owner in week one.



23–30

ESTABLISHED

Ready to pilot at the intended scale, with QA and integration running from the first call.

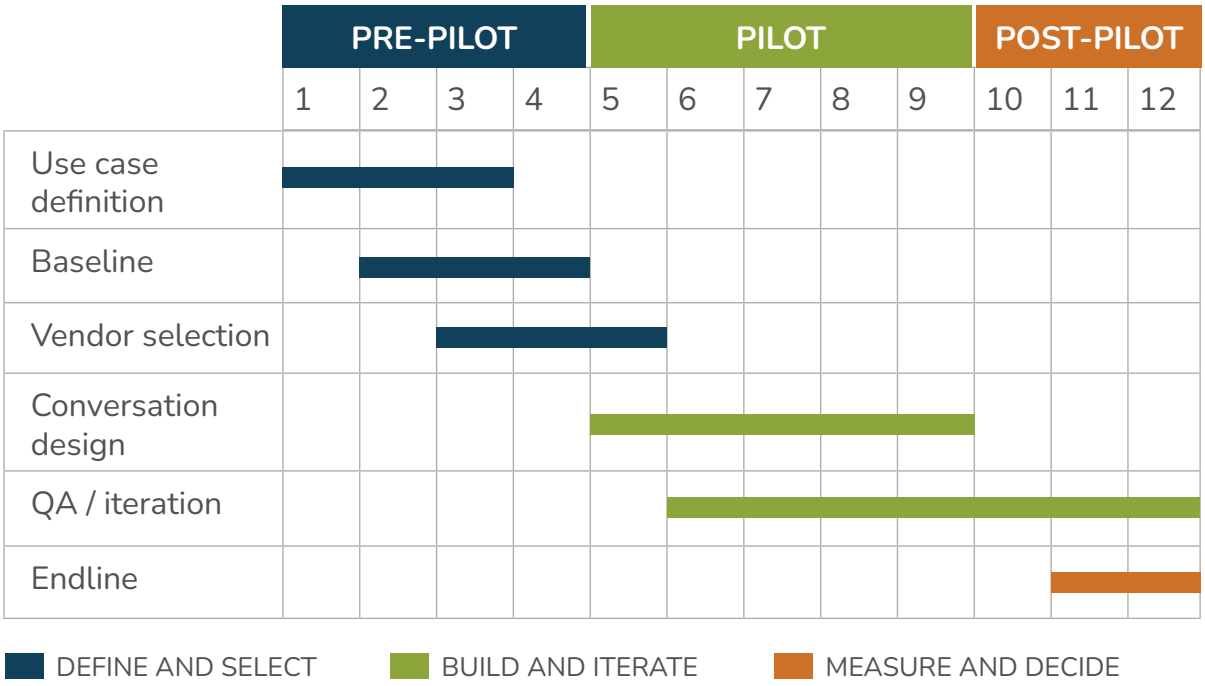
The six capabilities can be set out as a short diagnostic for an organisation to complete before it commits budget or selects a vendor. Each capability is scored from 1 to 5, where 1 indicates the capability is absent and 5 indicates it is an established practice. The scores are summed to give a total out of 30.

6 What a successful pilot design could look like

The capabilities that take a Voice AI pilot from pilot to scale need more internal deliberation than access to technology. From EkStep's conversations with 40-odd adopters, a genuinely usable Voice bot typically takes a few months to build and integrate, and capabilities like QA and conversation tooling for code-switching need continual tuning even after the pilot is green-lit.

An organisation needs both the know-when to choose the right problem and the know-how to build internal readiness. Successful organisations approached this in a structured three-month cycle, combining a pilot budget with top-down leadership buy-in and dedicated team members.

THE THREE-MONTH CYCLE



PRE-PILOT
Weeks 1-4

Organisations defined outcomes, ran a baseline, and evaluated vendors on criteria beyond conversation quality.

PILOT
Weeks 5-9

Teams refined conversation flow with regular QA, feeding flags to vendors until quality improved.

POST-PILOT
Weeks 10-12

Results were measured against baseline on cost, resolution and satisfaction to enable a go/no-go call.

6 Conclusion

Unlike traditional software that is often plug-and-play and deterministic,Voice AI (and the larger agentic AI universe) has no standard workflows. Instead, a successful implementation can vary significantly by the use-case, end-users and internal readiness of an organisation - factors extrinsic to the technology itself.

The report takes the lens of adopters - especially those in the impact sector - who may be exploring or initiating their own Voice AI journey. Through examples from similar real-world organisations, readiness stood out as a significant barrier - particularly in impact organisations - and equally important as the product or pricing itself that can make or break a pilot. This organisation readiness and the six distinct elements it comprises is intended to be a basic checklist and along with the templates and anecdotes can serve as useful guiding principles in designing a Voice AI pilot.

Ultimately, each organisation's pathway to scaling Voice AI will be different but organisation readiness is a pre-requisite that must be built and bolstered continuously and collaboratively by adopters, system integrators and vendors so it can truly be inclusive and transformative.