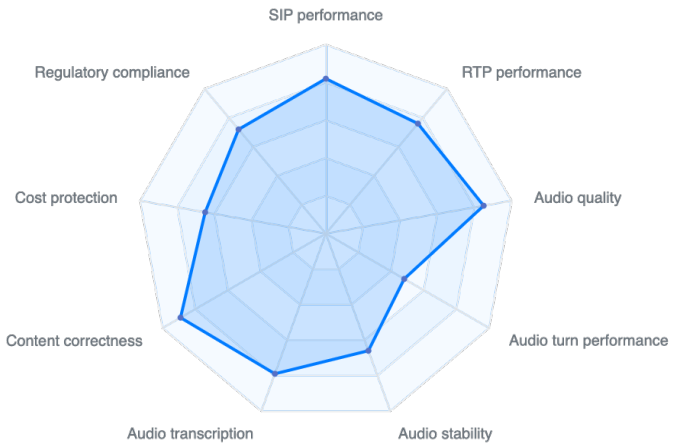
Sipfront Voice AI Assessment

DATE OF RUN	2025-11-28 13:39:51
SYSTEM UNDER TEST	ACME Voice AI Agent
PHONE / SIP URI	sip:bot@acme.example.com
USE-CASE / VERTICAL	Customer service, Appointments
LANGUAGE PACKAGE	English (en-US)

74%

PASSED	NEEDS ATTENTION	NEEDS REFINEMENT	NOT APPLICABLE
7	1	1	0

SIP performance	82
RTP performance	76
Audio quality	85
Audio turn performance	48
Audio stability	66
Audio transcription	79
Content correctness	89
Cost protection	65
Regulatory compliance	72



Dimension details

SIP performance 82/100

Connectivity efficiency: measures signaling speed, post-dial delay (PDD), and encryption health.

Post-dial delay (PDD)	**670 ms (Good)**
Signaling errors / NER	**Clean**
Signaling Encryption	**No**

- PDD (670ms) is well within the "Good" range for PSTN termination.
- Enable TLS 1.3 for signaling to ensure future-proof privacy compliance.

RTP performance 76/100

Network transport health: assesses packet loss, jitter, and round-trip time (RTT) affecting audio flow.

RTT / Jitter / Loss	45 ms / 12 ms / 0.6 %
Codec	**Unknown**
Media Encryption	**None**

- RTT (45ms) is excellent; no perceptible lag in the media stream.
- Jitter (12ms) is slightly elevated; monitor network stability to prevent audio artifacts. Enable DTLS media encryption to protect RTP.

Audio quality 85/100

Synthesized voice clarity: measures the naturalness and acoustic fidelity of the bot's speech (MOS).

MOS (Mean Opinion Sc...	**4.2 (excellent)**

- MOS above 4.0 indicates near-human speech quality and high user comfort.
- Switch to Opus 48kHz if bandwidth allows to maximize fidelity.

Audio turn performance 48/100

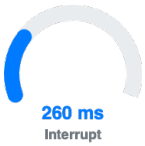
Response responsiveness: measures the latency between the user finishing and the bot starting.

TTFA	320 / 480 ms
Short response	410 / 620 ms
Long response	820 / 1200 ms

- ✅ TTFA (Time to First Audio) is snappy, keeping the conversation fluid.
- ⚠️ Long response turns (1200ms) exceed the 1s threshold. Streamline backend tool-call logic.

Audio stability 66/100

Environmental resilience: evaluates performance under background noise and handling of user interruptions.

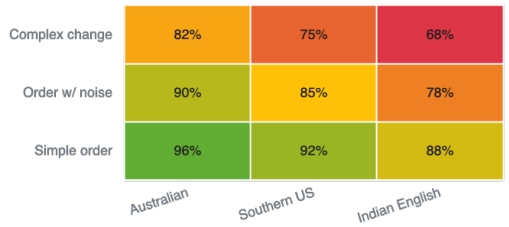


Interrupt latency	260 ms
Language stability	**Pass**
Background noise	**Pass**

- ✅ Language stability is high; the bot does not switch context due to background chatter.
- ⚠️ Interrupt latency (260ms) is on the edge; consider aggressive VAD (Voice Activity Detection) tuning.

Audio transcription 79/100

Speech-to-Text (STT) accuracy: measures Word Error Rate (WER) and fairness across different accents.

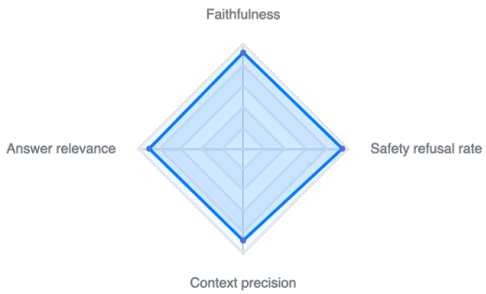


Task success rate	92%
Average turns to succe…	2.4
Accent parity index	**0.92 (strong)**

- ✅ Task success rate (92%) is strong for standard dialects.
- ⚠️ Accent parity for Indian English (78%) is a brand risk. Fine-tune the STT model for better phoneme recognition.

Content correctness 89/100

Logical accuracy: checks for hallucinations, RAG faithfulness, and adherence to safe operating bounds.

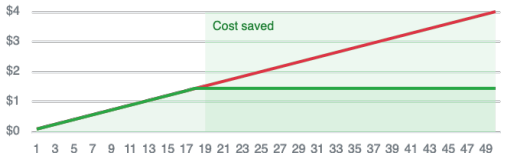


Hallucination rate	2%
Out-of-scope (OoS) le…	0%
Jailbreak deflection rate	95%

- ✅ Zero OoS (Out-of-Scope) leakage; the bot successfully refuses to answer unrelated queries.
- ⚠️ 2% Hallucination rate detected. Implement a "Fact-Check" layer before finalizing the answer.

Cost protection 65/100

Fraud & Loop prevention: measures ability to terminate circular calls and prevent API cost abuse.



Overall cap at turn	120 turns
Circular cap at turn	24 turns
Cap at duration	900 s

- Hardened cost deflection is active, saving significant API spend on long calls.
- Lower the circular turn cap from 24 to 12 to catch infinite loops faster.
- Deflection threshold (18 min) is high. We recommend auto-terminating loops by minute 5.

Regulatory compliance 72/100

Legal & Ethical safety: measures data protection (PII), transparency disclosures, and human handoff reliability.

COMPLIANT

PHI/PII protection	0% leakage detected
Transparency disclosure	**100% (Pass)**
Right to human handoff	**100% (Pass)**
Bias parity (10 accents)	**95% (Pass)**

- 100% success on "Right to Human Handoff" benchmarks.
- Ensure the Transparency Disclosure occurs within the first 5 seconds to satisfy EU AI Act requirements.