

How often do people realise they are talking to an AI?

A transcript-level census of 102,154 connected human phone conversations

Callable AI · Research Note CA-2026-01 · Published 17 August 2026

0.25%

of human conversations
included
an explicit AI challenge

253

conversations with a
detection utterance

102,154

qualifying human
conversations analysed

1 in 404

conversations where the
caller said it out loud

Summary

Across 30 days of live outbound calling (19 July – 17 August 2026), 0.25% of connected human conversations contained an explicit challenge to the agent's humanity — a caller asking whether they were speaking to a real person, naming the agent as AI, or calling it a robot, machine or recording. That is roughly one conversation in every 404.

This is a census, not an estimate: every qualifying conversation in the period was scanned, so the figure carries no sampling error. A separate 500-conversation random sample was drawn as a cross-check and returned 1 detection in 493 conversations with parseable caller turns (0.20%), consistent with the population figure.

The rate is stable week to week (0.21%–0.32%) and rises with conversation length: 0.17% on conversations of 25–59 seconds, 0.83% on conversations of two to five minutes.

What was measured

The measured quantity is voiced detection: the caller said, out loud and on the recording, that they believed they were not speaking to a human. It is not a measure of private suspicion, and it is not a Turing test. Its strength is that every counted event is an auditable utterance in a stored transcript; its limit is stated plainly in Limitations below.

Data and method

Source

All rows in the production **call_logs** table created between 19 July and 17 August 2026 (Australia/Brisbane), spanning 11 distinct business tenants running outbound voice agents. Transcripts are speaker-labelled, with caller speech on **User:** turns and agent speech on assistant turns.

Inclusion criteria

Filter	Rule	Rationale
Transcript present	non-null, length > 50 characters	a conversation must be readable to be scored
Duration	≥ 25 seconds	excludes immediate hang-ups with no exchange
Caller turns	≥ 3 User: turns	requires a genuine two-way exchange

Exclusions

132,276 conversations met the inclusion criteria. 28,787 (21.8%) were then removed as machine answerers rather than humans: voicemail greetings, carrier and Google Call-Assist screening bots, and “this call is being recorded” announcements. Those systems reliably produce false detections — the robot in the transcript is the network's, not the caller's. That left 103,489 human conversations, of which 102,154 had machine-parseable caller turns and form the denominator throughout this note.

Detection rule

Only caller turns were scanned; agent speech was discarded before matching, so the agent's own wording can never trigger a detection. A conversation counts once, however many times the caller raises it. The case-insensitive pattern covers five families: real person / human challenges, explicit naming of AI or artificial intelligence, robot or bot, machine or computer, and recording or prerecorded.

Results

Headline rate

Measure	Value
Conversations meeting inclusion criteria	132,276
Machine answerers excluded (voicemail, screening bots)	28,787
Human conversations	103,489
Human conversations with parseable caller turns (denominator)	102,154
Conversations containing a detection utterance	253
Detection rate	0.25% (1 in 404)
95% confidence interval (Wilson)	0.22% – 0.28%

The interval is reported for completeness only. Because every qualifying conversation in the period was scanned, 0.25% is the observed population value; the interval describes the rate one would expect if the period were treated as a draw from ongoing operations.

Stability over time

Five calendar weeks, Brisbane time. The final week is partial (three days) and its three detections sit well inside normal variation for that volume.

Week commencing	Conversations	Detections	Rate	Tenants
20 July 2026	17,838	49	0.27%	6
27 July 2026	19,696	50	0.25%	9
3 August 2026	25,709	81	0.32%	11
10 August 2026	33,182	70	0.21%	9
17 August 2026 (partial)	5,731	3	0.05%	8
Total	102,156	253	0.25%	11

Weekly totals differ from the headline denominator by two conversations that straddle a week boundary in local time.

Detection rises with conversation length

25–59 seconds		0.17%	128 of 75,262
60–119 seconds		0.37%	79 of 21,611
2–5 minutes		0.83%	43 of 5,207
Over 5 minutes		3.95%	3 of 76

The gradient is the clearest signal in the data: the longer someone stays on the line, the more likely they are to voice the question. Note the confound — the small number of very long conversations is enriched with callers who stayed on precisely because they were probing the agent, so the 3.95% cell (3 conversations of 76) should be read as directional, not as a stable rate.

How the question gets asked

Category	Conversations	Share of detections
Human / real-person challenge (“are you a real person?”)	187	73.9%
Called a machine or computer	32	12.6%
Asked if it was a recording or prerecorded message	22	8.7%
Named explicitly as AI	7	2.8%
Called a robot or bot	6	2.4%

Three quarters of detections are the generic “are you a real person?” — a question people also ask human offshore call centres. Only 7 conversations in 102,154 (0.007%) named the agent as AI unprompted.

Verbatim examples

“Are you a real person?” — 14 conversations

“Are you real?” — 6 conversations

“Is this a recording?” — 4 conversations

“Alex, is this a prerecorded message, or are you a real person?”

“Are you a human or are you a machine?”

“Am I talking to a real person or AI?”

Cross-check: 500-conversation random sample

A deterministic random sample of 500 conversations was drawn from the same eligible population (ordered by **md5(id)**, a stable pseudo-random ordering). 493 had parseable caller turns; 1 contained a detection (0.20%). That single conversation contained both “Are you real?” and, moments later, “You sound like a real person, mate.”

At a base rate this low, a 500-conversation sample is a weak instrument — a single conversation moves the result by 0.2 percentage points, and the expected count is only about one. The sample is reported because it was run, and because it agrees with the census; the census figure is the one to cite.

Limitations

Voiced detection only

Anyone who suspected the agent and simply ended the call without saying so is invisible to this method. The true rate of private suspicion is necessarily higher than 0.25%, and this study places no bound on it.

Short calls excluded by design

Conversations under 25 seconds or with fewer than three caller turns were excluded, because a hang-up with no exchange cannot be scored. If silent suspicion concentrates anywhere, it concentrates in that band.

Transcription error

Detections rest on automatic speech recognition. Spot-checking found occasional garbles that inflate counts (for example “you and I are bot” in an unrelated context) and near-misses that deflate them. The errors are small relative to the effect and run in both directions.

Keyword matching, not comprehension

The rule matches phrasing, not intent. A caller asking “are you a real person?” out of ordinary call-centre scepticism is counted identically to one who has concluded the agent is synthetic. This biases the reported rate upward, making 0.25% a conservative ceiling on confident detection.

Single-vendor, single-market

All conversations were placed by Callable AI voice agents to Australian numbers, largely in solar, finance and services outbound. The figure should not be read as a property of conversational AI in general.

Reproducibility

Every number in this note derives from read-only SQL against the production transcript store. No records were modified. The query pattern is: filter **call_logs** by window, transcript presence, duration and caller-turn count; remove machine answerers by transcript pattern; split each transcript into lines; retain only caller-labelled lines; and count distinct

conversations whose caller lines match the detection pattern. Re-running it over the same window reproduces the counts exactly.

Suggested citation

Callable AI (2026). How often do people realise they are talking to an AI? A transcript-level census of 102,154 connected human phone conversations. Research Note CA-2026-01, 17 August 2026.

Prepared by Callable AI from first-party production call data. Figures cover 19 July – 17 August 2026 (Australia/Brisbane) across 11 business tenants. No personally identifying information is reproduced in this document; verbatim quotations are caller utterances stripped of all identifying context.