

Evaluation Transcript

Voice-Based Native Language Government Scheme Agent (Telugu)

This section presents multiple evaluation scenarios demonstrating successful interaction, failure handling, and edge-case recovery, as required.

Scenario 1: Successful Interaction

Objective:

Evaluate whether the agent can correctly identify eligible government schemes using voice input in Telugu.

Turn 1

User (Voice Input – Telugu):

నాకు ప్రభుత్వ పథకాల వివరాలు కావాలి

System Action:

- Speech-to-text converts audio to Telugu text
- Agent enters ASK_AGE state

Agent (Voice Output – Telugu):

మీ వయస్సు ఎంత?

Turn 2

User (Voice Input – Telugu):

నా వయస్సు అరవై ఐదు సంవత్సరాలు

System Action:

- Age stored in memory
- Agent transitions to ASK_INCOME

Agent (Voice Output – Telugu):

మీ వార్షిక ఆదాయం ఎంత?

Turn 3

User (Voice Input – Telugu):

నా వార్షిక ఆదాయం ఒక లక్ష రూపాయలు

System Action:

- Income stored in memory
- Agent transitions to ASK_OCCUPATION

Agent (Voice Output – Telugu):

మీ వృత్తి ఏమిటి?

Turn 4

User (Voice Input – Telugu):

నేను రైతును

System Action:

- Occupation stored
 - Agent transitions to EVALUATE
 - Eligibility Tool invoked
 - Scheme Retrieval Tool invoked
-

Final Response

Agent (Voice Output – Telugu):

మీ వివరాలను పరిశీలించాను.

మీ వయస్సు అరవై ఐదు సంవత్సరాలు.

మీ వార్షిక ఆదాయం ఒక లక్ష రూపాయలు.

మీ వృత్తి రైతు.

ఈ వివరాల ఆధారంగా మీరు ఈ ప్రభుత్వ పథకాలకి అర్హులు:

- వృద్ధాప్య పెన్షన్
- ఆరోగ్య శ్రీ
- రైతు భరోసా

Result: Successful completion

Scenario 2: Failure Handling (Unclear Speech)

Objective:

Test agent response to unclear or noisy audio input.

Turn 2 (Failure Case)

User (Voice Input):

(Low volume / unclear audio)

STT Output:

మారిల్స్ ల్ల...

System Action:

- Hallucination detected
- Input rejected

Agent (Voice Output – Telugu):

సరిగా వినిపించలేదు. దయచేసి స్పష్టంగా మాట్లాడండి.

Result:

Failure detected and handled gracefully without corrupting memory.

Scenario 3: Incomplete Information

Objective:

Verify that the agent does not proceed without required details.

Turn 3

User (Voice Input – Telugu):

నాకు తెలియదు

System Action:

- Income missing
- State remains ASK_INCOME

Agent (Voice Output – Telugu):

మీ అర్హతను నిర్ణయించడానికి మీ వార్షిక ఆదాయం అవసరం. దయచేసి చెప్పండి.

Result:

Agent correctly waits for required information.

Scenario 4: Contradiction Handling (Edge Case)

Objective:

Test memory consistency and contradiction detection.

Turn 2

User:

నా వయస్సు అరవై ఐదు సంవత్సరాలు

Turn 4 (Contradiction)

User:

నా వయస్సు ముప్పై సంవత్సరాలు

System Action:

- Conflict detected with stored memory

Agent (Voice Output – Telugu):

మీరు వయస్సు గురించి భిన్నమైన సమాచారం ఇచ్చారు. దయచేసి సరైన వయస్సును నిర్ధారించండి.

Result:

Contradiction detected and resolved via clarification.

Evaluation Summary

Criterion	Result
Voice-first interaction	✓
Native Telugu pipeline	✓
Agentic reasoning	✓
Tool usage	✓
Conversation memory	✓
Failure handling	✓
Edge-case recovery	✓

Conclusion

The evaluation demonstrates that the system:

- Operates fully in voice-first Telugu
- Maintains conversation memory
- Uses multiple tools
- Handles failures and contradictions
- Implements a true agentic workflow