

## Evaluation Transcript and Evaluation Criteria

This section documents representative interactions with the system, including successful cases, failure cases, and edge cases, to demonstrate correctness, robustness, and compliance with requirements.

### 1. Successful Interaction (Eligible Scheme)

#### Scenario:

User provides valid age, income, and state information.

#### Transcript:

Agent:

నమస్కారం. నేను సహాయవాణి. ప్రథమత్వ పథకాల కోసం మీకు సహాయం చేస్తాను.

User:

నా పేరు జాహాని

Agent:

ధన్యవాదాలు అండి. దయచేసి మీ వయస్సు, ఆదాయం, మరియు రాష్ట్రం చెప్పండి.

User:

నా వయస్సు 45

Agent:

దయచేసి మీ ఆదాయం మరియు రాష్ట్రం వివరాలు చెప్పండి.

User:

ఆదాయం రెండు లక్షలు

Agent:

దయచేసి మీ రాష్ట్రం వివరాలు చెప్పండి.

User:

ఆంధ్రప్రదేశ్

Agent:

మీరు YSR Cheyutha పథకానికి అర్థులు.

అవసరమైన పత్రాలు ఆధార్ కార్డు మరియు బ్యాంక్ పాస్‌బుక్.

Agent:

ఇంకొక వ్యక్తి కోసం చెక్ చేయాలా? అవను లేదా కాదు చెప్పండి.

User:

కాదు

Agent:  
ధన్యవాదములు

**Result:**

The system successfully identified eligibility, retrieved scheme details, and ended the conversation gracefully.

---

## 2. Failed Interaction (No Eligible Schemes)

**Scenario:**

User does not satisfy eligibility conditions.

**Transcript:**

User:  
నా వయసు 30  
ఆదాయం ఐదు లక్షలు  
తెలంగాణ

Agent:  
మీ వివరాలకు సరిపడే ప్రభుత్వ పథకాలు ప్రస్తుతం లేవు.

Agent:  
జంకోక వ్యక్తి కోసం చెక్ చేయాలా? అవును లేదా కాదు చెప్పండి.

User:  
కాదు

Agent:  
ధన్యవాదములు

**Result:**

The system correctly rejected eligibility based on defined rules and handled the case without errors.

---

## 3. Edge Case: Missing Information

**Scenario:**  
User provides incomplete details.

**Transcript:**

User:  
నా వయసు 50

Agent:

దయచేసి మీ ఆదాయం మరియు రాష్ట్రం వివరాలు చెప్పండి.

User:

అంద్రప్రదేశ్

Agent:

దయచేసి మీ ఆదాయం వివరాలు చెప్పండి.

**Result:**

The system retained previously provided information and prompted only for missing details.

---

#### 4. Edge Case: Telugu Number Words

**Scenario:**

User provides income using Telugu number words.

**Transcript:**

User:

ఆదాయం రెండు లక్షలు

**System behavior:**

Income correctly interpreted as 200000.

**Result:**

Local numeric parsing handled Telugu number expressions without requiring additional LLM calls.

---

#### 5. Failure Handling: Noise / Silence

**Scenario:**

User speech is unclear or absent.

**Transcript:**

Agent:

▶ READY — SPEAK AFTER THE BEEP

(No speech detected)

Agent:

క్షమించండి, మీ మాటలు సరిగ్గా వినిపించలేదు. దయచేసి మళ్ళీ చెప్పండి.

**Result:**

The system handled the failure gracefully and prompted the user to retry.

---

 **Evaluation Criteria**

The system was evaluated against the following criteria:

**Voice-First Interaction**

All interactions are conducted using speech input and speech output.

**Status:** Satisfied

**Native Language Support**

Telugu is used consistently across speech recognition, reasoning, and response generation.

**Status:** Satisfied

**Agentic Workflow**

The system follows a planner–executor–memory pattern and adapts responses based on conversation state.

**Status:** Satisfied

**Tool Usage**

The system uses multiple tools, including an eligibility engine and a scheme knowledge base.

**Status:** Satisfied

**Conversation Memory**

User information is retained across turns and reset appropriately when starting a new session.

**Status:** Satisfied

**Failure and Edge-Case Handling**

Noise, silence, missing information, and Telugu numeric expressions are handled correctly.

**Status:** Satisfied