

2. AI Processing Strategy

1. Audio Transcription + Analysis ✓

Workflow:

1. Audio Upload:

- User uploads voice note via mobile app.
- The file is stored in Amazon S3 and triggers a message in AWS SQS queue.

2. Transcription:

- A background Whisper worker (hosted in ECS/Fargate or Lambda) fetches the audio.
- Transcribes speech to text using OpenAI Whisper or faster-whisper (if self-hosted).

3. Text Processing:

- Transcription is passed to GPT-4o Mini for:
 - Segmenting notes room-wise.
 - Extracting decisions, issues, actions.
 - Normalizing time, names, or ambiguous terms.

4. Metadata Enrichment:

- Adds metadata like timestamp, speaker confidence, detected room names, etc.

Why this design?

- Whisper is robust to accents and noisy environments (ideal for construction sites).
- GPT-4o Mini balances reasoning ability and cost for post-processing and insight generation.

2. Cross-Referencing with Floor Plans ✓

Approach:

- Each floor plan PDF is uploaded alongside the visit.
- OCR is run (Tesseract or AWS Textract) to extract:
 - Room names
 - Labels or annotations

Cross-Matching Logic:

- Extracted room names from the transcription are matched against OCR'd names from the floor plan.
- Fuzzy matching (e.g., fuzzywuzzy or RapidFuzz) handles discrepancies like:
 - "Living Room" ↔ "L.R."
 - "Bed 1" ↔ "Bedroom 1"

Output:

- Annotated mapping of issues and decisions to specific rooms in the floor plan (e.g., bounding boxes or footnotes).
- Enables visual correlation for reviewers on the dashboard.

3. Key Prompts for Data Extraction ✓

System Prompt to GPT-4o Mini:

You are an AI assistant helping to summarize architecture site visit notes.

Your task: Given transcribed audio, extract the following as structured JSON:

- Project details (name, date, engineer, location)
- Room-wise observations
- Decisions made during the visit
- Action items, including assigned stakeholders
- Mentions of any vendors, contractors, or design issues

Output only valid JSON. Omit any parts that are unclear or ambiguous.

Optional Prompts for Revisions / Additions:

- “Regenerate only the action items”
- “Highlight all safety-related concerns mentioned”

4. Error Handling When AI Fails ✓

Failure Modes:

Failure Type	Detection	Mitigation
Whisper fails	Empty transcription	Notify user, allow re-upload
Transcription gibberish	Confidence < threshold	Send to human QA queue
LLM response malformed JSON	Regex / parser check	Retry, fallback to older prompt
Missing room names	No match in OCR set	Highlight in red for review
API failure / timeout	Exception logging	Retry queue + notify dashboard

Fallback Mechanisms:

- Auto-versioning: Store previous versions of insights.

- Manual override: Allow users to edit output on dashboard.
- Logs captured via Prometheus + Grafana for system-level insights.