info@appinfoedge.com
www.appinfoedge.com

Interviewer AI Agent

(High level feature details)

PREPARED BY

Appinfoedge Solutions Pvt. Ltd.

Version: 1.0

Interviewer AI Agent - Requirement Overview

1) Project Overview

The Interviewer AI Agent is a web platform that runs realistic interviews with candidates. It can:

- Create or accept a Job Description (JD).
- Read and understand a candidate's CV.
- Generate and ask tailored interview questions.
- Run text or video interviews where the AI speaks and listens live.
- Judge technical and behavioral skills, plus body language and face expression during video.
- Produce a final score out of 10 and a detailed report that helps with hiring decisions.
- Provide a web dashboard for recruiters to manage the full process.

2) Key Objectives

- Automate mock interviews to reduce manual effort and bias.
- Generate JDs from a simple job title (e.g., "Java Developer").
- Parse CVs into structured data (skills, experience, education).
- Create role- and profile-based questions (technical + behavioral + scenario).
- Control interview difficulty manually or let AI set it from the CV.
- Provide a web dashboard to manage JDs, CVs, interviews, schedules, and reports.
- Run live video interviews where the AI speaks and listens like a human interviewer.
- Analyze body language and facial expressions (video only) to add soft-skill signals.
- Deliver a final result: per-parameter scores (0–10), overall score (0–10), and a clear recommendation.

3) Scope

In scope: JD generation, CV parsing, question generation, chat interview, video AI interview (live voice), scheduling, scoring, reports, dashboard, ATS/calendar/video integrations, analytics.
Out of scope (initial): Background checks, offer management, payroll, full ATS replacement.

4) Roles

Recruiter/Admin

- Upload CVs, create JDs, launch interviews, set/auto difficulty, schedule video sessions, review reports.

Candidate

- Receives invite, joins chat or video interview, answers AI, sees optional feedback if enabled.

AI Agent (System)

- Reads JD + CV, generates questions, conducts the interview (chat/video), probes deeper, scores answers, analyzes non-verbal cues (video), produces the report.

5) Functional Requirements

5.1 JD Management

- **Feature**: Enter a job title or paste a JD. AI can create a complete JD (responsibilities, must-have skills, good-to-have).
- **Story**: Recruiter types “Java Developer”. AI suggests a JD with bullet points. Recruiter edits and saves.
- **Done when**: Saved JD appears in job list; version history kept.

5.2 CV Parsing & Analysis

- **Feature**: Upload PDF/DOCX/TXT CVs. Extract name, contact, skills, education, role history, years, keywords.
- **Story**: Recruiter uploads CV; system shows “Skills: Java, Spring Boot, Docker, AWS | Exp: 5 yrs”.
- **Done when**: Parsed profile is viewable and searchable; errors are surfaced for manual fix.

5.3 Interview Question Generation

- **Feature**: AI generates technical, behavioral, and scenario questions using JD + CV.
- **Story**: For a Spring Boot profile, AI creates beginner/intermediate/advanced questions; recruiter can reorder/remove/add.
- **Done when**: A question set is saved to the interview plan.

5.4 Interview Difficulty

- **Feature**: Difficulty = Manual (Beginner/Intermediate/Advanced) or AI Auto (reads CV seniority, project depth).
- **Story**: A 7-year CV triggers Advanced difficulty and more system-design questions.
- **Done when**: Selected difficulty is visible and locked into the plan.

5.5 Chat Interview (Text)

- **Feature**: AI interviews via chat, asks follow-ups based on answers, and scores each response.
- **Story**: Candidate answers; AI probes: “How is DI different from Service Locator?” then scores the depth and correctness.
- **Done when**: Transcript and per-question scores are saved.

5.6 Video AI Interview (Live Voice)

- **Feature**: AI interviewer speaks and listens in real time (two-way audio/video).
  - AI Voice (TTS): The AI reads out questions naturally.
  - AI Listening (STT): The AI converts candidate speech to text and understands in real time.
  - Real-time follow-ups: AI adapts questions based on the spoken answer.
  - Live captions: Show captions for accessibility and clarity.
  - Network resilience: Auto switch to audio-only or chat if bandwidth is poor.
- **Story**: At 4:00 PM, a candidate joins the link. AI greets them with voice, verifies audio/video, confirms consent for analysis, then starts questions. The AI listens, follows up, and keeps the flow natural, like a human interviewer.
- **Done when**: Full recording, transcript, and scores are saved; recruiter can replay key moments.

5.7 Body Language & Facial Expression Analysis (Video)

- **Feature**: During video, the system analyzes non-verbal signals to support soft-skills assessment.
  - Face expression trends (positive/neutral/negative, confusion moments).
  - Attention/eye-gaze approximation (screen focus), head nods, posture openness.
  - Speaking pace, pauses, filler words (“um”, “uh”), clarity and confidence signals.

- **Notes**: This is assistive only (not a final verdict). Recruiters can turn it on/off at job level. The UI clearly shows candidate consent and data use.
- **Story**: Candidate maintains steady engagement, smiles occasionally, answers clearly with low filler rate; system adds positive soft-skill signals to the report.
- **Done when**: Non-verbal insights appear in the report with timestamps and confidence ranges.

5.8 Scheduling & Calendar/Video Integration

- **Feature**: Book time slots, send invites via Google/Outlook; links for Zoom/Teams or in-app WebRTC room.

- **Story**: Recruiter picks 25-Aug 4 PM; candidate gets email/SMS with “Join Interview” link and reminders.
- **Done when**: Event created; reminders sent; join link works.

5.9 Recruiter/Admin Dashboard

- **Feature**: Jobs, CVs, interviews, schedules, reports, analytics; filters and search.
- **Story**: Recruiter opens candidate profile → sees scores, transcript, key clips, strengths/areas to improve, and recommendation.
- **Done when**: All core actions are managed from one UI.

5.10 Reports & Feedback

- **Feature**:
  - Full transcript (chat or STT from video).
  - Per-question scores + comments.
  - Per-parameter scores (0–10) + overall score (0–10) with weights.
  - Recommendation (e.g., Strong Hire / Hire / Borderline / No Hire).
  - Key moments (clips) and non-verbal insights (video).
- **Story**: Report shows “Technical 8, Problem-Solving 7, Communication 7, Behavioral 7, Presence 6; Overall 7.3; Recommendation: Proceed to panel.”
- **Done when**: Report can be viewed, exported (PDF), and shared.

6) Scoring Model

Default parameters (0–10 each, role-based weights adjustable):

1. Technical Depth (e.g., Java, Spring, system design) – default weight: 40%
2. Problem Solving & Reasoning – 20%
3. Communication Clarity (structure, articulation) – 15%
4. Behavioral & Teamwork (ownership, collaboration, examples) – 10%
5. Role Fit / Experience Alignment – 10%
6. Presence & Delivery (Video) (confidence, pace, low filler, posture; assistive) – 5%

Overall Score (0–10) = weighted average of the above.
Recommendation logic (editable):

- ≥ 8.5 → Strong Hire
- 7.0–8.4 → Hire
- 6.0–6.9 → Borderline / Consider
- < 6.0 → No Hire

7) UX Flows

Recruiter

1. Create Job → Generate/Edit JD → Save.
2. Upload CV(s) → Review parsed profiles.
3. Create Interview → Pick Chat or Video (AI talks + listens) → Difficulty (manual/auto).
4. If video: pick slot → send invite.
5. After interview: open report → share/export.

Candidate

1. Open invite link → device check (mic/cam) → consent for video analysis → join.
2. AI introduces itself, asks questions, probes deeper.
3. End screen.
4. Thank you email.

Video Room UI

- AI avatar nameplate, question panel, captions on/off, timer, progress, reconnect, switch to audio/chat fallback.

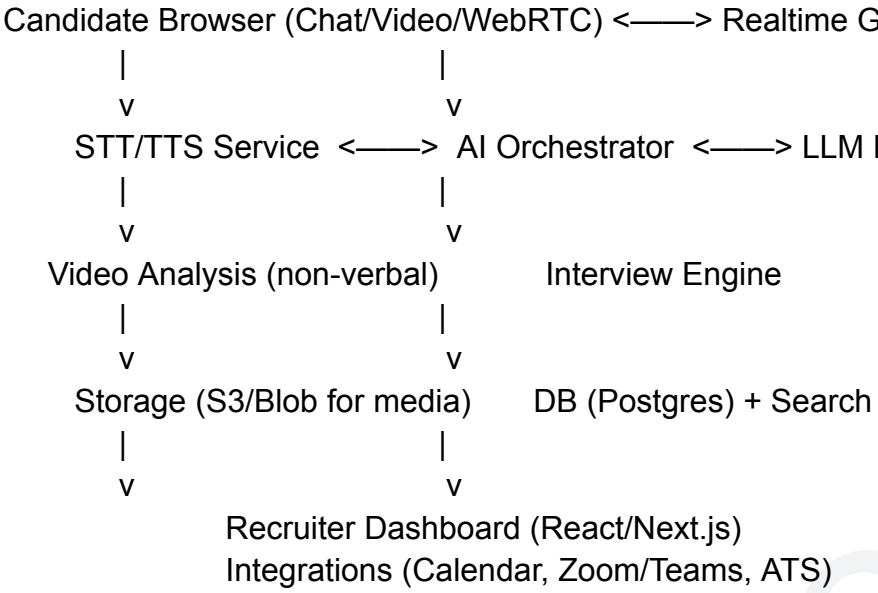
8) Integrations

- **LLMs**: OpenAI GPT-4.5 / GPT-4-Turbo, Anthropic Claude, Cohere.
- **Speech**: STT (e.g., Whisper/Azure/Google) and TTS (e.g., Azure/Amazon Polly).
- **Video**: WebRTC in-app; optional Zoom/Teams links.
- **Calendar**: Google/Outlook.

9) Technology Stack

- **Frontend (Web App)** → React.js / Next.js, WebRTC for video; WebSockets for live events.
- **Backend** → Python (FastAPI/Django)
- **Database** → PostgreSQL / MongoDB
- **AI Layer:** LLM router (OpenAI/Claude/Cohere), Prompt templates, Guardrails.
- **AI Models** → OpenAI GPT-4.5, Claude, Cohere
- **Video** → WebRTC (real-time video)
- **Vision (video analysis):** MediaPipe/OpenCV for posture/facial signals; optional AWS Rekognition/Azure/Google Video AI.
- **Speech** → STT/TTS APIs (e.g., Whisper, Azure, Google Speech)
- **Vision** → OpenCV, MediaPipe, or APIs for emotion detection
- **Infra:** AWS/Azure/GCP; Docker + CI/CD; Monitoring (Prometheus/Grafana).

10) High-Level Architecture



11) Data Model (key entities)

- Job (id, title, JD, skills, difficulty_mode, created_by)
- Candidate (id, name, email, phone, cv_url, parsed_profile_json)
- Interview (id, job_id, candidate_id, mode=chat|video, status, scheduled_at, room_link)
- Question (id, interview_id, text, topic, difficulty)
- Answer (id, question_id, raw_text/transcript, score, comments)
- NonVerbalSample (interview_id, timestamp, features_json)
- Report (interview_id, scores_json, overall_score, recommendation, pdf_url)

12) Acceptance Criteria

JD Generation

- Given a job title, AI returns JD with ≥ 5 bullets (responsibilities + skills).
- Recruiter can edit and save; version stored.

CV Parsing

- Upload succeeds for PDF/DOCX/TXT; at least name + top 5 skills extracted or clear error shown.

Question Generation

- For JD+CV, system creates ≥ 12 questions across technical, behavioral, scenario; recruiter can edit.

Chat Interview

- Transcript and per-question scores saved; follow-ups are contextual to answers.

Video AI Interview (voice)

- AI asks questions via TTS; candidate answers via mic; STT transcript captured; follow-ups based on speech.
- Live captions toggle works; network fallback to audio/chat triggers automatically.
- Recording saved with timestamped transcript.

Body Language & Facial Expression

- When enabled and consented, report shows non-verbal summary + timestamps.
- Feature can be disabled per job; consent screen must be accepted to proceed.

Scoring & Report

- Report shows all parameter scores (0–10) and overall score (0–10) with weights and a recommendation.
- Export to PDF works; share link (internal) works.

Scheduling

- Calendar invite created; reminder sent; join link valid.

Dashboard

- Recruiter can filter by job/status; open any interview; view transcript, scores, clips.

13) Phases & Deliverables

Phase 1 (MVP)

- JD (manual + AI), CV parsing, question gen, chat interview, basic scoring, dashboard, basic reports.

Phase 2

- Video AI interview (AI speaks + listens), live captions, non-verbal analysis (assistive), scheduling, recordings, advanced reports, exports.

14) Final Notes

- The Video AI Agent speaks and listens live during interviews.
- Body language and expression analysis is included (video), clearly consented, and assistive only.
- The final report includes all key parameters and an overall score out of 10 with a hiring recommendation.

15) Key Terminology

1. AI Agent / Interviewer AI Agent

An intelligent software program powered by LLMs (Large Language Models).
It acts as a virtual interviewer, capable of:

- Asking questions.
- Listening to answers.
- Analyzing responses.
- Judging body language & facial expressions.
- Giving a final score and interview report.

2. ATS (Applicant Tracking System)

A recruitment software that organizes and tracks candidates.

- Stores resumes/CVs.
- Tracks interview status.
- Manages scheduling.
- Provides a searchable candidate database.
  Our system will integrate ATS-like features to store all candidates, resumes, scores, and reports.

3. LLM (Large Language Model)

AI models trained on massive datasets (like GPT-4.5, GPT-4 Turbo, Claude, Cohere).

- Used here to:
  - Generate Job Descriptions (JDs).
  - Parse resumes.
  - Create interview questions.
  - Conduct live conversation.

4. JD (Job Description)

A written summary of the job role, responsibilities, and required skills.

- Input Methods:
  - Manual (entered by HR/Admin).
  - AI-generated (only role/title given, e.g., “Java Developer”).

5. Resume / CV Parsing

The process of extracting key details (skills, education, work history) from a candidate’s CV.

- Helps AI to:

- Match JD vs candidate skills.
- Set difficulty level.
- Personalize questions.

6. Difficulty Level (Easy/Medium/Hard)

The level of complexity of interview questions.

- Manual mode → recruiter chooses difficulty.
- AI mode → system sets difficulty based on candidate resume.

7. STT (Speech-to-Text)

Technology that converts spoken words into text.

- Used to transcribe candidate's answers during live interviews.

8. TTS (Text-to-Speech)

Technology that converts text into human-like speech.

- Used for the AI Agent to speak questions live during interviews.

9. Conversational AI

The combination of STT + LLM + TTS that allows the AI Agent to:

- Speak like a human.
- Listen to answers.
- Respond naturally in a video interview.

10. Video Interview (Live)

An online session where the candidate interacts with the AI Agent face-to-face (like Zoom/Meet).

- Features:
  - Two-way audio & video.
  - Recording.
  - AI-driven real-time evaluation.

11. Body Language & Facial Expression Analysis (Computer Vision)

AI-based video analysis that evaluates candidate's soft skills:

- Confidence (eye contact, gestures).
- Nervousness (fidgeting, lack of focus).
- Engagement (smile, attentiveness).
  Helps judge non-verbal communication.

12. Interview Scoring / Evaluation Report

At the end of the interview, the AI generates a result card.

- Scored out of 10 for each parameter:
  - Technical Knowledge.
  - Problem-Solving.
  - Communication.
  - Confidence & Body Language.
  - Cultural/Behavioral Fit.
- Final overall score helps recruiters decide hire/no-hire.

13. Admin / Recruiter Dashboard

A web panel for HR/managers to:

- Upload CVs.
- Manage JDs.
- Schedule interviews.
- Set difficulty levels.
- Track progress of candidates.
- View/download interview reports.

14. Scheduling System

Feature that lets HR or AI Agent schedule interviews:

- Candidate receives calendar invites & email notifications.
- Video meeting link is generated automatically.