The Shadow Instructor

Master your next interview with your own AI Shadow

Gemini 3 Hackathon Submission

Team: The Shadow Instructor Team

Date: February 9, 2026

Contents

1 Abstract **2**

2 Introduction **2**
2.1 Problem Statement 2
2.2 Proposed Solution 2

3 System Architecture **2**
3.1 High-Level Design 3
3.2 Data Flow 3

4 Key Features **3**
4.1 Resume-Based Personalization 3
4.2 Dual-Agent "Good Cop, Bad Cop" 4
4.3 Real-Time Visual Feedback ("Heads Up Display") 4

5 Implementation Details **4**
5.1 Technology Stack 4
5.2 Prompt Engineering Strategy 4

6 Challenges & Solutions **5**
6.1 Latency Management 5
6.2 Hallucination Control 5

7 Future Scope **5**

8 Conclusion **5**

9 Links & References **5**

1 Abstract

Technical interviews are a high-pressure environment where candidates often fail not due to a lack of knowledge, but due to poor communication and anxiety. **The Shadow Instructor** is a revolutionary dual-agent AI platform designed to bridge this gap. Powered by Google's **Gemini 3.0** model family, it provides a realistic interview simulation where an adversarial "Interviewer Agent" conducts the session, while a benevolent "Shadow Agent" analyzes performance in real-time. This report details the architecture, implementation, and impact of the system, showcasing how multimodal AI can transform professional development.

2 Introduction

2.1 Problem Statement

Job seekers, particularly students and early-career professionals, face significant challenges in interview preparation:

1. **Performance Anxiety**: A lack of realistic environments to practice verbal communication under pressure.
2. **Generic Feedback**: Existing mock interview solutions often provide generic responses that lack technical depth or specificity to the user's background.
3. **Context Awareness**: Standard AI chatbots fail to tailor questions to the candidate's actual resume or specific job description.
4. **Latency Issues**: Text-based interactions disrupt the natural flow required for verbal interview practice.

2.2 Proposed Solution

The Shadow Instructor addresses these issues through a web-based platform that offers:

- **Real-time Voice Interaction**: Sub-second latency conversation using Gemini 3.0 Flash.
- **Resume Contextualization**: Automated parsing of PDF resumes to seed the interview context.
- **Dual-Agent Feedback Loop**:
  - The *Interviewer* conducts the session, challenging the user on technical concepts.
  - The *Instructor* ("Shadow") analyzes the transcript to generate a comprehensive scorecard on technical accuracy and communication clarity.

3 System Architecture

The system follows a Hybrid Client-Server-Cloud architecture designed for minimal latency and maximum reliability.

3.1 High-Level Design

- **Frontend (Next.js 16)**: Handles the user interface, audio capture (PCM 16kHz), and direct WebSocket communication. It manages local audio buffering and implements Voice Activity Detection (VAD) for interruptibility.
- **Backend (FastAPI)**: Serves as the trusted orchestration layer. It handles resume uploads, parses PDF content, and manages the session state machine.
- **AI Engine (Google Gemini)**:
  - **Gemini 3.0 Flash**: Powers the low-latency live interview session (Interviewer Agent).
  - **Gemini 3.0 Pro**: Powers the high-level reasoning and detailed feedback generation (Shadow Agent).

3.2 Data Flow

1. **Initialization**: User uploads resume → Backend parses text → System Prompt constructed with resume context. 2. **Session Start**: Frontend establishes WebSocket connection to Backend. 3. **Interaction Loop**:

- User speaks → Audio Streamed to Backend.
- Gemini 3.0 Flash processes audio → Generates Audio Response.
- Backend streams response to Frontend → User hears Interviewer.

4. **Shadow Loop (Parallel)**:

- Transcript + User Audio Metadata sent to Gemini 3.0 Pro.
- Shadow Agent analyzes "Vibe" metrics (pacing, hesitation) and Technical Accuracy.
- Real-time feedback pushes to Frontend via secondary WebSocket channel.

4 Key Features

4.1 Resume-Based Personalization

Unlike generic interview bots, The Shadow Instructor ingests your actual PDF resume. The system extracts:

- **Skills**: Programming languages, tools, and frameworks.
- **Experience**: Past projects and roles.

This allows the AI to ask questions like: *"I see you used React at your last internship. Can you explain how you handled state management in that project?"*

4.2 Dual-Agent "Good Cop, Bad Cop"

The separation of concerns is critical. The **Interviewer Agent** is prompted to be challenging, skeptical, and focused on drilling down into technical details. The **Shadow Agent** is prompted to be supportive, observant, and focused on coaching. This mirrors the real-world dynamic of having a mentor review your performance.

4.3 Real-Time Visual Feedback ("Heads Up Display")

Using a specialized React hook (`useShadowObserver`), the frontend analyzes the video stream locally to detect:

- Absence (User left frame)
- Poor Lighting
- Lack of Eye Contact

This provides immediate, ephemeral feedback toasts to correct behavior *during* the interview.

5 Implementation Details

5.1 Technology Stack

Component	Technology
Frontend	Next.js 16, TypeScript, Tailwind CSS, Framer Motion
Backend	Python 3.12, FastAPI, PyPDF
Real-time	WebSockets, WebRTC (Video Preview)
AI Models	Google Gemini 3.0 Flash, Gemini 3.0 Pro
Deployment	Vercel (Frontend), Render (Backend)

Table 1: Technology Stack

5.2 Prompt Engineering Strategy

We utilized distinct system prompts for each agent.

- **Interviewer Prompt**: "You are a tough technical interviewer. Do not accept surface-level answers. Probe for understanding. Keep responses concise (under 30 seconds)."
- **Shadow Prompt**: "You are an expert communication coach. Analyze the user's last response. Did they answer the question directly? Was their tone confident? Output JSON only."

6 Challenges & Solutions

6.1 Latency Management

Challenge: Audio-to-audio latency with standard LLMs is too high (3-5 seconds) for a natural conversation.

Solution: We adopted Gemini 3.0 Flash's native audio streaming capabilities, eschewing intermediate transcribers (STT) and synthesizers (TTS). This reduced end-to-end latency to sub-500ms.

6.2 Hallucination Control

Challenge: The AI would sometimes invent details about the user's resume.

Solution: We implemented a "Strict Context Injection" mechanism where the parsed resume text is injected into the system prompt with clear delimiters and instructions to *only* reference provided information.

7 Future Scope

- **Video Analysis Integration**: Deeper integration of Gemini 3.0 Pro Vision to analyze body language (posture, hand gestures) coupled with the audio stream.
- **Multiplayer Mock Mode**: Allowing a human friend to take over the "Interviewer" role while the AI remains as the "Shadow Instructor" to grade both parties.
- **Code Collaboration**: A shared whiteboard environment where the user can write code that the AI reviews in real-time.

8 Conclusion

The Shadow Instructor represents a significant leap forward in automated educational tools. By combining the speed of Gemini 3.0 Flash with the reasoning depth of Gemini 3.0 Pro, we created a system that feels less like a chatbot and more like a human mentor. It democratizes access to high-quality interview coaching, helping candidates everywhere land their dream jobs.

9 Links & References

- **Public Project Link**: `https://the-shadow-instructor.vercel.app`
- **GitHub Repository**: `https://github.com/aryan-dani/The_Shadow_Instructor`
- **Demo Video**: `https://youtu.be/example-video-link`