

# HARSHAWARDHAN SHRIVASTAVA

connect.harshawardhan@gmail.com | github.com/HarrisWarner04 | linkedin.com/in/connect-harshawardhanshrivastava | Bhopal, India

## Summary

AI native builder focused on shipping real AI systems end to end, from real-time voice agent and hybrid RAG pipelines to MCP servers and automated evaluation harnesses. Worked on debugging asynchronous systems, benchmarking retrieval and latency and turning prototype failure modes into measurable engineering improvements. 3× national hackathon winner with a bias toward building and iterating on real problems under time constraint. Eager to explore and learn more about capabilities of AI systems.

## PROJECTS

### miniVoxSetu : Real-Time Multimodal Voice AI Agent for Banking

2026 | [github.com/HarrisWarner04/miniVoxSetu](https://github.com/HarrisWarner04/miniVoxSetu)

(Python, FastAPI, WebSockets, Groq, Deepgram, Qdrant, PyTorch)

Available at: [minivoxsetu.bugbiceps.in/](https://minivoxsetu.bugbiceps.in/)

- Engineered a real-time multimodal voice AI agent for Indian banking, combining streaming STT/TTS, semantic analysis, acoustic emotion detection, hybrid RAG, and asynchronous processing; reduced end-to-end conversation latency from **1,250ms to ~390ms**.
- Built and evaluated a **hybrid RAG engine** combining BM25 and vector search with Reciprocal Rank Fusion; automated benchmarks improved **Hit@1 from 68.4% to 93.8%** and reached **100% Hit@3**.
- Designed a **7-layer barge-in system** using AudioWorklet VAD, client queue flushing, WebSocket interruption signals, generation-ID gating, and stale-task pruning; reduced interrupt-to-silence latency to **<280ms** and resolved an asynchronous audio race condition.
- Built an automated evaluation harness covering **RAG retrieval, PII precision/recall/F1, acoustic inference, STT/LLM/TTS latency, end-to-end latency, barge-in, and regression testing**; PII redaction achieved **1.000 F1** on baseline synthetic banking transcripts.

### GitHub Analyzer: MCP Server & Agentic Code Review Tool

2026 | [github.com/HarrisWarner04/githubAnalyser](https://github.com/HarrisWarner04/githubAnalyser)

(Python, MCP SDK v2, Gemini 2.5 Flash, Groq LLaMA 3, GitHub REST API, Docker)

Available at: [githubAnalyser.bugbiceps.in/](https://githubAnalyser.bugbiceps.in/)

- Built a **Model Context Protocol server from scratch using the official MCP Python SDK v2**, exposing a typed `analyze_repo` tool over JSON-RPC/stdio for Claude Desktop, Cursor, Kiro, and MCP Inspector.
- Designed a two-model analysis pipeline using **Gemini 2.5 Flash for deep repository analysis** and **Groq LLaMA 3 8B for fast structured scoring**, evaluating repositories across **8 weighted engineering categories**.
- Engineered asynchronous, rate-limit-aware GitHub ingestion with **httpx and Pydantic**, producing structured scores, severity-ranked issues, and prioritized recommendations.
- Added **Streamable HTTP transport and Docker deployment** for remote hosting, extending the same MCP tool from local subprocess execution to a deployable service.

### HireReady : Multi-Modal AI Interview & Placement Platform

2026 | [github.com/HarrisWarner04/HireReady](https://github.com/HarrisWarner04/HireReady)

(Team Project | FastAPI, React, Firebase, ChromaDB, Groq, Gemini, Whisper, MediaPipe)

Available at: [hireready.bugbiceps.in/](https://hireready.bugbiceps.in/)

- Built an end-to-end AI placement platform combining **resume ATS analysis, GitHub codebase verification, semantic job matching, and multimodal AI mock interviews** into a composite PlaceScore.
- Designed a deterministic + neural ATS engine combining **keyword matching, semantic similarity, format analysis, skills coverage, experience relevance, and education matching** into a reproducible 0–100 score.
- Implemented **ChromaDB-based semantic job matching**, Whisper transcription with hallucination filtering, and adaptive interview generation grounded in resume and GitHub context.
- Built GitHub codebase intelligence using repository metadata and source analysis to **cross-check claimed skills against actual code evidence** and generate interview talking points.

## TECHNICAL SKILLS

**Languages:** Python

**Generative AI:** LLMs, Agentic Pipelines, RAG, Hybrid Search, MCP, Prompt Engineering, Evaluation Harnesses

**Backend:** FastAPI, Firebase, WebRTC, WebSockets

**Database:** Qdrant, ChromaDB, Firebase, MySQL

**Tools & Skills:** Git, GitHub, Prompt Engineering, Responsible AI, n8n

**ML models:** DistilHuBert, Librosa, Ollama

## ACHIEVEMENTS

- 3x National Hackathon Winner: NASA Space Apps Challenge, Qubix 2025, KV Sangathan.
- Finalist in 8 National Hackathons including Smart India Hackathon 2025.
- President, Student Activity Council OIST Bhopal (2025–Present).

## INTERNSHIPS

### Freelance AI Automation

Dec 2025 – Present

- Delivered 5+ AI-powered workflow automations (client pipelines, lead capture, document processing) for small businesses.

### Technology Intern: HaronIndia Pvt. Ltd. · Remote

Nov 2025 – Jan 2026

- Implemented AI features for farmers alongside the founding team.

## EDUCATION

**B.Tech. Computer Science & Engineering** - Oriental Institute of Science and Technology, Bhopal

2023 -2027

CGPA: 8.02