Interactive Multilingual AI Career Platform – Build Plan

# 1) Data You’ll Need

Seed knowledge (skills ontology, role templates, resources, modules).  
Runtime user data (resume text, JD, extracted skills, session state, audio snippets).  
Small labeled data (ATS calibration, interview rubric examples, intent router).

# 2) Models to Use / Train

Off-the-shelf: MiniLM embeddings, LLM for generation, Whisper STT, Cloud TTS, Translation API.  
Light training: Logistic Regression for role fit scorer, intent router (optional small classifier).  
Rules: ATS scoring formula, skill-path synthesis (topo sort).

# 3) Tools & Libraries

Backend: Flask, faiss-cpu, sentence-transformers, scikit-learn, pypdf, pdfminer, docx, flask-socketio.  
Frontend: React, Vite, Tailwind, shadcn/ui, lucide-react, recharts, react-dropzone, WebRTC, Web Audio API.  
Infra: SQLite/Postgres, Render/Fly/Cloud Run, structlog, environment variables for secrets.

# 4) Web Requirements

Frontend UX: Tabs (ATS Checker, Career Advisor, Skill Builder, Dashboard), language selector, chat dock.  
APIs: /ats/check, /ats/rewrite, /advisor/analyze\_profile, /advisor/match\_roles, /advisor/roadmap, /skills/path, /interview/\*, /chat (SSE/WS), /ws/voice.  
Realtime: SSE for text, WebRTC for voice/video.  
Security: transient audio storage, PII stripping, CORS rules, content safety filters.

# 5) Component Pipelines

ATS Checker: parse resume/JD → score → rewrites via LLM.  
Career Advisor: extract skills → role fit → RAG resources → roadmap generation.  
Skill Builder: gap → modules → topo sort → calendar packing → explanations.  
Voice/Video: STT → translate → LLM → translate → TTS → avatar (WebRTC).

# 6) Acceptance Criteria

ATS: consistent scores, keyword-rich rewrites, ATS-friendly PDF export.  
Advisor: valid JSON roadmaps with citations.  
Skill Builder: prerequisite consistency, hours/day respected.  
Voice/Video: <3.5s round-trip, language switching works, synced avatar audio/video.

# 7) 2-Day Build Plan

T0-4h: repo scaffold, seeds, analyze\_profile + match\_roles API.  
T4-8h: ATS check + scoring + UI gauge, roadmap endpoint with RAG + JSON guardrails.  
T8-12h: Skill path synthesis + UI calendar, README/resume bullets.  
T12-16h: SSE /chat stream, intent router, ATS rewrite endpoint.  
T16-20h: Voice MVP (mic → STT → reply).  
T20-24h: Video avatar, TTS, polish + demo prep.

# 8) Config (Environment Variables)

LLM\_PROVIDER=...  
EMBED\_MODEL=all-MiniLM-L6-v2  
DB\_URL=sqlite:///app.db  
STT\_PROVIDER=whisper  
TTS\_PROVIDER=...  
TURN\_URL=...

# 9) Prompt Templates

ATS bullet rewrite: concise STAR with JD keywords.  
Plan generation: ≤90-min tasks, citations, why\_this.  
Interview feedback: JSON {score, what\_went\_well, improve\_next\_time, follow\_up, hint}.  
Translation wrapper: translate to <LANG>, preserve tech terms.