UtopiaHire – Technical Specifications Book

1. Project Overview

UtopiaHire is an AI-powered employability platform specialized in the **aviation sector** (flight attendants, cabin crew, pilots).
It assists users in improving their career readiness through CV analysis, interview simulation, job matching, and AI-guided mentorship — with a focus on fairness, transparency, and domain relevance.

---

2. Proposed Tech Stack

Layer	Technology / Framework	Purpose / Notes
Frontend	**React + TypeScript (MERN)**	User interface, dashboard, file upload, visualization, and chat interface.
Backend API	**Node.js / Express**	Manages RESTful routes, handles authentication, and connects frontend to AI microservices.
AI & NLP Microservices	**Flask + Python**	Runs modular AI logic: CV parsing, NLP analysis, job matching, and interview evaluation. Communicates via JSON APIs.
Database	**MongoDB**	Stores user profiles, resumes (metadata only), job postings, and feedback reports in JSON format.
Communication Protocol	**RESTful API (JSON)**	Ensures lightweight communication between services and client apps.
Authentication & Security	**Google OAuth + JWT**	User identity verification and secure token-based sessions.
OCR / CV Parsing	**YOLOv8 + Tesseract (or PaddleOCR)**	Extracts text and structure from PDF/Scanned CVs for preprocessing.

NLP Models & APIs	**spaCy + Sentence-BERT + Gemini API**	Handles language understanding, entity extraction, skill matching, and improvement feedback.
AI Interview Voice	**Eleven Labs / Annyang**	Voice interaction and audio output during interview simulation.
Hosting & DNS	**Dockerized Deployment + Cloudflare DNS**	Containerized services with security and performance optimizations.

---

3. Core Features and Technologies

Feature	Description & Functionality	Core Technologies & AI Components
1. CV Reader, Reviewer & Career Roadmap Generator	- Parses CVs (PDF/image/text) and extracts structured information (skills, certifications, experience, languages).- Uses NLP to evaluate employability based on airline job profiles.- Generates improvement suggestions and career roadmap.	- **OCR:** YOLOv8 or Tesseract.- **NLP:** spaCy (entity extraction), Sentence-BERT (semantic similarity), Gemini API (text feedback generation).- **Backend:** Flask microservice returning JSON output.- **Integration:** REST API to Node.js backend.
2. AI Interview Simulator & Language Coach	- Generates interview questions and analyzes candidate responses.- Provides linguistic and behavioral feedback, optionally using voice interaction.	- **LLM:** Gemini API for Q&A generation.- **Speech:** Eleven Labs or Annyang.- **Text Analysis:** TextBlob or spaCy for sentiment and tone detection.- **Deployment:** Flask microservice.
3. Job Matcher + Airline Ranking System	- Uses NLP to match extracted CV data with aviation job postings.- Computes semantic similarity between candidate skills and job descriptions.- Ranks airlines by compatibility and highlights improvement areas.	- **NLP:** spaCy (keyword/entity parsing), Sentence-BERT (semantic embeddings), Gemini API (natural language feedback).- **Data:** Job listing JSON dataset.- **Output:** REST API returning match percentages and improvement advice.
4. Networking Platform for Users & Companies	- Enables connections between aviation job seekers, mentors, and recruiters.- Supports private messaging and company pages.	- **MERN Stack:** React + Node.js + MongoDB.- **Matching Algorithm:** Simple AI-based keyword or role matching (expandable).

5. AI Chatbot for Guidance & Support	- Acts as an intelligent assistant providing career insights, interview tips, and progress summaries.	- **Gemini API:** Prompt-engineered conversational model.- **Integration:** Flask service for natural responses.- **Frontend:** React chat interface with streaming responses.

---

4. Security & Compliance

- **Authentication:** Google OAuth 2.0 with JWT for token-based authorization.
- **Encryption:** HTTPS/TLS for all communication; AES-256 for stored CVs.
- **Data Privacy:** Anonymized user data during AI analysis; GDPR and Tunisian Law n°2004-63 compliance.
- **Hosting Security:** Cloudflare DNS for DDoS protection and secure SSL termination.
- **Access Control:** Role-based access for users, recruiters, and admins.

---

5. Deployment & Architecture Overview

1. **Frontend (React + TypeScript)** — user interface for CV upload, feedback, job recommendations, and interviews.
2. **Backend (Node.js / Express)** — central API hub connecting frontend, database, and Flask AI microservices.
3. **AI Microservices (Flask + Python)** — modular services for:
    - **CV Parsing & NLP Evaluation**
    - **Job Matching & Ranking**
    - **Interview Simulation**
    - **Chatbot Guidance**
4. **Database (MongoDB)** — JSON documents for users, resumes, feedback, and job listings.

5. **Security Layer** — Google Auth, JWT validation, HTTPS, Cloudflare DNS.

All data flows use **RESTful JSON** for uniformity and portability.

---

6. Suggested Future Enhancements

Area	Suggestion	Value Added
AI Pipeline	Add **Redis caching** for NLP embedding reuse.	Boosts performance and reduces recomputation time.
Microservices Orchestration	Deploy using **Kubernetes** for scalable service management.	Ensures reliability and load balancing.
Domain-Specific AI	Fine-tune a **Sentence-BERT model** on aviation job datasets.	Improves precision in matching and scoring.
Multilingual Support	Integrate **spaCy multilingual pipelines** or Gemini multilingual API.	Supports English, French, and Arabic CVs.