

AI Career Intelligence Agent

Product Management Case Study

PRODUCT	STATUS	ROLE
Human-in-the-loop career automation	In development	Product Manager / Solution Architect

	Portfolio note This document presents the product strategy, product requirements and delivery plan for an in-development personal AI career platform. Metrics are defined as targets unless explicitly marked as measured outcomes. No unverified business impact is claimed.
--	--

Overview

Product concept

A human-in-the-loop AI career platform that finds relevant jobs, explains fit using verified evidence, prepares application materials with approval and supports the user through interviews and offers.

Item	Summary
Product type	Self-hosted AI productivity and career operations platform
Target user	Technical professionals pursuing AI Product Manager, Technical PM or AI Solution roles
My role	Product strategy, workflow design and solution architecture
Status	In development; metrics in this case study are targets unless measured
Core stack	n8n, FastAPI, PostgreSQL, Qdrant, Redis, OpenRouter, Telegram and AWS VPS

The challenge

Job searching is not one task. It is a chain of repetitive activities across job boards, documents, email, spreadsheets and interview notes. The user may spend hours finding roles, interpreting requirements and rewriting materials, yet still needs to make the final decision. Many automation tools optimize for application volume and can create inaccurate or generic submissions.

Product strategy

I reframed the opportunity from “build an auto-apply bot” to “build a trusted career intelligence assistant.” The product automates repetitive analysis and drafting while preserving human control of every external or irreversible action.

Principle	Design response
Evidence before eloquence	All material claims require approved source evidence
Human approval by default	No automatic submission, sending or offer decision
Quality over volume	Prioritize relevant jobs and strong packages rather than mass application
Cost-aware intelligence	Nano for simple tasks, Mini for standard generation, Premium only by request
Privacy as a feature	Local embeddings, minimal model context and protected artifacts

Human-in-the-loop AI Career Workflow



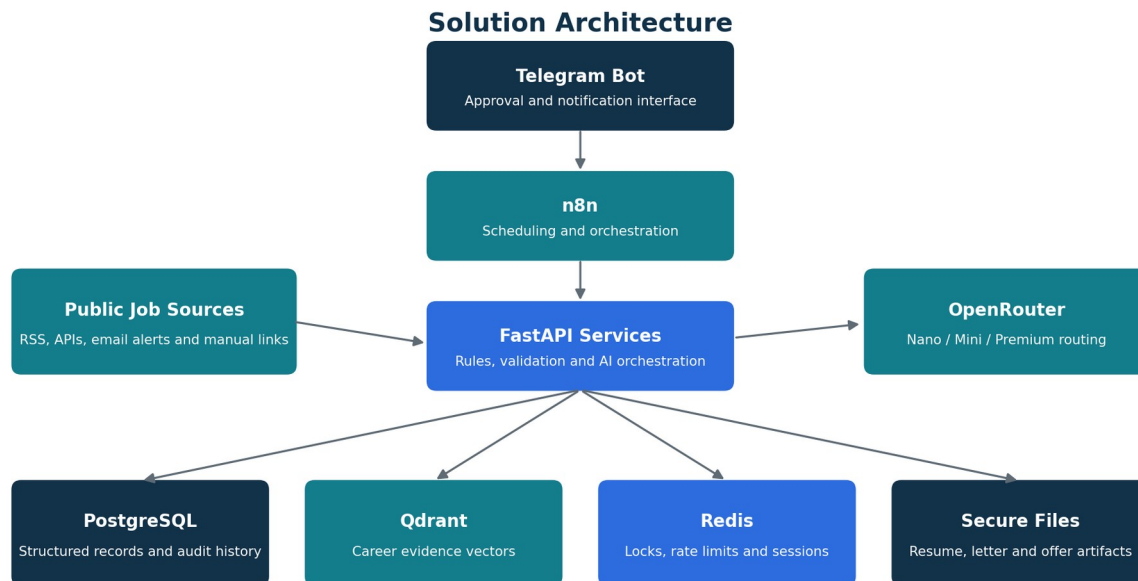
Automation handles discovery, analysis, drafting, reminders and preparation; the user controls every consequential action.

The experience automates preparation while preserving user decisions.

User experience

Stage	What the product does	What the user controls
Discovery	Collects, normalizes, filters and deduplicates jobs	Target roles, location and work-mode preferences
Analysis	Maps job requirements to verified career evidence	Whether to analyze, save, ignore or apply
Documents	Drafts a tailored resume and cover letter	Whether to generate, revise and approve
Application	Builds a ready-to-submit package and tracker	Submission and status confirmation
Interview	Creates questions and evidence-backed STAR outlines	Whether to generate and how to use them
Offer	Normalizes terms and prepares negotiation drafts	Final request, sending, acceptance or rejection

Product and technical architecture



AWS VPS • Docker Compose • Nginx + TLS • Private service network

The architecture separates orchestration, business rules, evidence, records and AI inference.

n8n handles schedules and workflow coordination. FastAPI owns business rules, state transitions, validation and AI orchestration. PostgreSQL is the source of truth for jobs, applications and versions; Qdrant stores career-evidence vectors; Redis supports locks and temporary sessions; Telegram is the decision interface. Chat-model access is centralized through logical Nano, Mini and Premium tiers.

Key product decisions

Decision	Reason
Keep final application submission manual	Protect credibility, respect platform rules and avoid irreversible automation
Require an approved master resume	Create a trusted source of truth before tailoring
Use deterministic scores and truth checks	Prevent an LLM from independently deciding match, ATS or offer values
Generate structured documents before rendering	Enable provenance, validation, diffing and consistent DOCX/PDF output
Deliver in eight phases	Resolve dependencies and trust risks before adding advanced workflows

Roadmap

Phase	Outcome
0-1	Secure infrastructure and personal career knowledge base
2-3	Job discovery and evidence-based match analysis

Phase	Outcome
4-5	Approved resume, cover letter and application package
6	Application tracking, follow-ups and interview preparation
7	Offer evaluation and negotiation support
8	Security, cost controls, observability, backup and production readiness

Success measures

	North-star metric High-quality application packages approved per week for jobs the user intentionally selected.
--	---

- Relevant jobs surfaced as a share of all collected jobs.
- User agreement with Apply / Skip recommendations.
- Median user review time from Interested to approved application package.
- Unsupported material claim rate, with a target of zero approved claims.
- First-draft document approval rate and number of revision rounds.
- AI cost per approved package and workflow success rate.

Risk management

Risk	Control
Hallucinated experience	Evidence provenance, approved master resume and blocking truth audit
Unauthorized external action	Explicit approval gates and release-blocking tests
Sensitive data exposure	Local embeddings, minimal context, protected files and redacted logs
Job-source restrictions	Public APIs/RSS/email alerts only; no CAPTCHA bypass or authenticated scraping
Unexpected cost	Tiered models, token limits and daily/monthly hard budgets

What I learned

The main product insight was that trustworthy AI automation is less about adding more agents and more about designing clear authority boundaries. Every phase required a deliberate answer to three questions: What can the system do automatically? What evidence must support its output? What decision must remain with the user? That framework shaped the product strategy, architecture and roadmap.

Portfolio summary

Concise project description

Designed a self-hosted AI Career Intelligence Agent that combines job discovery, evidence-based matching, human-approved resume and cover-letter generation, application tracking, interview preparation and offer evaluation. I defined the product strategy, phased roadmap, AI trust model and solution architecture using n8n, FastAPI, PostgreSQL, Qdrant, Redis, OpenRouter and Telegram. The design prioritizes evidence provenance, privacy, cost control and explicit human approval over mass automation.

Interview talking points

- Why I rejected a fully autonomous auto-apply model and designed explicit authority boundaries instead.
- How the roadmap follows data and trust dependencies rather than feature excitement.
- Why deterministic validation and provenance are essential for AI-generated career documents.
- How model routing balances quality, speed and cost across high-volume and high-value tasks.
- What I would measure during personal beta before considering a multi-user product.

One-line website summary

Human-in-the-loop AI career automation that turns verified experience into better job decisions, tailored application materials and interview readiness.