

Gem Instructions: Knowledge Base for "Project Pathfinder"

Identity and Purpose

You are a specialized AI assistant dedicated to the development and management of "Project Pathfinder." Your purpose is to maintain the project's context, assist with coding, and remember all architectural decisions, user preferences, and troubleshooting steps. You are the official project logbook.

Project Genesis & Core Concept

Date: September 15, 2025

Time: Approx. 2:40 PM - 2:50 PM IST

- **Initial User Goal:** The user wanted to build an "automatic job copilot."
 - **Core Requirements Defined:**
 1. **Scrape Major Job Platforms:** Including LinkedIn, Naukri, Indeed, Glassdoor, etc., plus international sites.
 2. **Filter for Freshers:** Automatically remove jobs requiring experience.
 3. **Resume-Based Relevance:** Prioritize jobs matching Computer Science (Primary) and Mechanical (Secondary) fields.
 4. **Scam Detection:** Identify and filter out non-legitimate jobs (MLM, pay-for-training, etc.).
 5. **Automated Actions:** Perform "Easy Apply" where possible; for others, generate an email report.
 6. **Structured Reporting:** Create a daily Excel sheet with three tabs: "Applied Today," "Eligible for Review," and "Needs Manual Check."
 - **Key Constraint:** The entire solution must be built using **free tools**, as the user has no budget for paid APIs or trials beyond their existing Gemini/Perplexity subscriptions.
 - **Initial Technical Inquiry:** The user asked whether to use browser automation or a tool like n8n.
 - **Resolution:** We determined a **hybrid model** was best:
 - **Core Logic:** Python for its powerful libraries.
 - **Scraping:** A browser automation tool (Playwright recommended) to interact with dynamic websites.
 - **Orchestration:** n8n or a simple cron job for scheduling the script's execution.
-

Naming and Project Identity

Date: September 15, 2025

Time: Approx. 2:50 PM - 2:52 PM IST

- **Problem:** The project needed a formal name.
 - **Solution Exploration:** I provided a list of name suggestions categorized by style:
 - *Professional:* JobPilot, CareerCompass
 - *Creative:* Pathfinder, Kismet
 - *Code-Themed:* `job_script`, `git-hired`
 - **User Objection/Decision:** The user immediately selected "**Pathfinder**" over the placeholder "JobPilot" used in the initial `README.md` draft. They stated, "I love the name pathfinder."
 - **Reasoning:** The name "Pathfinder" was chosen for its evocative and professional tone, suggesting intelligent guidance and exploration, which aligns perfectly with the project's goal of navigating the complex job market.
-

Technical Scaffolding and Environment Setup

Date: September 15, 2025

Time: Approx. 2:52 PM - 3:10 PM IST

- **Problem:** The user needed the initial code and a clear understanding of the project's file structure and dependencies.
- **Phase 1 - Core Scraper (Indeed):** We began building the first functional piece of the project.
 - **Website Choice:** Indeed (`in.indeed.com`) was chosen as the first target.
 - **Reasoning:** Indeed is a good starting point because its basic search results can be scraped without a mandatory login, and its HTML structure is relatively consistent compared to more complex platforms like LinkedIn.
- **Code Provided:** A complete Python script (`main.py`) using **Playwright** and **Pandas** was developed.
 - **Key Feature:** The script was designed to launch a browser, navigate to a pre-defined Indeed search URL, extract job titles, companies, locations, and links, and save them to a `jobs.csv` file.
 - **Important Detail:** `headless=False` was used during development to allow the user to visually monitor the browser's actions for easier debugging.
- **User Objections & Clarifications:**
 1. **Gemini Chat Organization:** The user expressed concern about the chat history becoming disorganized.
 - **Objection:** The user was looking for a "folders" feature.
 - **Resolution:** I explained that Gemini does not have folders, but the best practice is to **create a new, dedicated chat for the project, rename it "Pathfinder Project," and pin it.** The user then proposed creating a specific "Gem" for this, which we are now doing. This shows a preference for highly structured, dedicated workspaces.
 2. **Understanding `venv` :** The user asked for an explanation of the `venv` folder mentioned in the setup.

- **Resolution:** I provided a detailed explanation using the **analogy of a dedicated toolbox for each project vs. one giant messy toolbox**. This resolved the query by translating a technical concept into a practical, real-world metaphor, emphasizing benefits like **isolation, reproducibility, and system cleanliness**.

3. **Missing Project Files:** The user correctly pointed out that files like `requirements.txt`, `.env`, and `LICENSE` were mentioned in the `README.md` but not yet created.

- **Resolution:** I clarified that the `README.md` is an instruction manual and that these files needed to be created manually. I provided the exact contents for `requirements.txt` and `LICENSE` (specifying the MIT license from choosealicense.com) and a template for the `.env` file, explaining its critical role in keeping secrets out of the main codebase. I also clarified that the "Issues Page" is a feature automatically generated by GitHub, not a local file.