

Project Plan: WebGenius

Problem Statement

In the age of information overload, users visit dozens of webpages daily but often struggle to:

- Extract relevant info quickly
- Retain or organize useful content
- Ask deeper questions about what they just read

Solution Overview

WebGenius is a Chrome Extension powered by LangChain, Perplexity AI, and optionally Vector DBs to:

- Automatically scrape the current pages content
- Store & preprocess it in a backend
- Allow users to chat with that content using Perplexity
- Deliver a real-time RAG (Retrieval-Augmented Generation) experience directly in the browser

Core Features

Chrome Extension:

- Fetch current webpage URL
- Provide a popup UI with Scrape, Chat, and Summarize options

LangChain-Powered Backend:

- Scrape with WebBaseLoader
- Chunk, clean, and embed content
- Store in Firebase / MongoDB

Perplexity Integration:

- Use Perplexity Premium for answers
- Send context + query to get accurate responses

Optional Add-ons:

- Export chat/summary as PDF

Project Plan: WebGenius

- Categorize pages
- History tracking
- Auto-summary on page load

Tech Stack

Extension: HTML, JS, Manifest V3

Backend: FastAPI (Python)

Scraping: LangChain (WebBaseLoader)

Chatbot: Perplexity Premium API

Storage: Firebase / MongoDB / Supabase

Deployment: GitHub Pages, Render, Vercel

Architecture Overview

User -> Chrome Extension -> FastAPI Backend

Chat UI <- Answer from Perplexity <- LangChain + Storage

Use Cases

- Research Assistant
- News Summarizer
- Product Page Analyzer
- Learning Tool
- Legal/Policy Reader

Scalability & Future Vision

- Add OpenAI/Gemini fallback
- Add Google login & sync
- Vector DBs for memory (FAISS/Chroma)
- Analytics & usage insights
- Serverless deployment for scale

Project Plan: WebGenius

Action Plan (MVP Roadmap)

Week 1: Chrome extension + URL capture

Week 2: FastAPI + LangChain integration

Week 3: Firebase setup for storage

Week 4: Chatbot UI + Perplexity integration

Week 5: Page summary, export, chat history

Week 6: Testing, UI polish, deployment

Target Audience

- Students, Researchers
- Analysts, Journalists
- Product reviewers
- General Web Users

Innovation Factor

- Chrome extension + LLM + RAG combo
- LangChain scraping and chatbot workflow
- Instant answers from any webpage