Assessment Task: Chrome Extension + AI Data Platform

We want to build a platform similar to ChatGPT, but powered by user-scraped data from websites. The project will have two main parts:

Part 1: Chrome Extension (Scraping Tool)

- Create a Chrome extension with a simple UI (popup or side panel).
- The extension should have a "Scrape" button.
- When clicked, it should scrape the visible content of the current webpage (articles, blogs, text, etc.).
- Store the scraped content along with its source URL in a backend/database.

Part 2: AI-powered Q&A; Platform

- Build a simple web platform where users can ask questions in natural language.
- The system should search through the previously scraped content and return:
1. A relevant answer generated using AI (summarized/contextualized).
2. The reference links (original URLs) where the information was scraped from.

Technology Options

You can use any AI model/technology for answering questions, for example:
- Gemini Flash (free usage)
- OpenAI GPT models
- LLaMA / Ollama
- LangChain for orchestration
- Vector databases (Pinecone, Weaviate, Milvus, FAISS, etc.)

Expectations

- The solution doesn't need to be production-ready, but should demonstrate:
- Knowledge of Chrome Extension development.
- Ability to scrape and store data with metadata (URL, timestamp, etc.).
- Integration of AI for Q&A; using the scraped data.
- Returning both answers + reference links.

Bonus (Optional)

- Add a bookmark manager-style dashboard where scraped articles can be viewed, searched, and managed.
- Support scraping structured data (e.g., tables, lists).
- Implement a basic user authentication system for managing personal scraped content.

Final Deliverables:

1. Source code (GitHub link).
2. Short demo (video/screenshare or live).
3. A README explaining:
- Tech stack
- Setup instructions
- Approach/architecture