

Hosting Ollama

Bradley Williams

How to Run Your Own Local Ai Without Restrictions

This guide explains how to set up and run Ai locally on your computer for free, without limits.

1. Download and Install Ollama

- **Step:** Download the Ollama tool from their [official website](#).
Ollama simplifies running and customizing AI models locally.
 - **Error Check:** Ensure you download the version compatible with your operating system (Windows/Mac/Linux).
-

2. Install Llama Model

1. After installing Ollama, ensure it runs in the background. You can verify this by opening the terminal and typing a relevant command (`ollama status`).
 2. Visit the Ollama website to download the AI model you want. For this guide, we'll use **Llama 3.2**.
 - **Recommendation:** Choose the **3 billion parameter model** for compatibility with most laptops.
 3. Copy the model's download command from the website.(`llama run ollama3.2`)
 4. Open your terminal and paste the command.
 5. Hit **Enter** to begin the download.
- **Error Explanation:**
 - If the download fails, ensure you're connected to the internet.
 - Check if there's enough disk space for the model.
-

3. Running the AI Model

- Once the download is complete, the model will load, and you can start typing in the chat window.
 - **Important Note:** This setup doesn't include a user-friendly interface by default. It runs as a terminal-based application, resembling ChatGPT.
-

4. Adding a User Interface (UI)

1. Install Docker:

- Download and install Docker from [Docker's official website](#).
- After installation, verify that Docker is running in the background (check the taskbar for the Docker icon).

2. Install Open Web UI:

- Open Web UI is a third-party tool available on GitHub. It provides a simple UI for interacting with AI models.
- For this guide, we'll use Docker to set it up.

3. Run the Docker Command:

- Find the Docker installation command for Open Web UI on its GitHub page.
- Open your terminal, paste the command, and hit **Enter** to run it.

4. Access the Web UI:

- After installation, Docker will create a container for the Web UI.
 - Check Docker to find the **port number** (commonly **8080**).
 - Open your web browser and navigate to **http://localhost:<port>** (e.g., **http://localhost:8080**).
 - **Error Explanation:**
 - If the UI doesn't load, verify Docker is running and the container is active.
 - Check if the port is already in use by another application.
-

5. Additional Features

- Open Web UI includes a **file upload system** (e.g., for business privacy concerns).
 - Some models support **uncapped responses**, meaning they bypass restrictions seen in tools like ChatGPT.
 - Example:
 - ChatGPT might block code generation for sensitive tasks like RAT scripts.
 - Locally running models may provide unrestricted outputs. However, be aware that generated code might not always run successfully.
-

6. Hardware Requirements

- Larger models require a **decent GPU**. For laptops without a dedicated GPU, stick to smaller parameter models to avoid performance issues.
-

By following these steps, you can set up your own customizable AI environment, enjoy more freedom.
