

An aerial photograph of a vast field with alternating rows of red and green plants, likely a tulip field. A small tractor is visible in the upper right corner. A large, white, rounded rectangular overlay is centered on the image, containing the text.

**Imagine having your own
personalized search
assistant for Wellness and
Relaxation Media**

Project Overview

3 stages:

- Fetching Youtube data and create a structured dataset (.pdf)
- Cleaning and embedding the data
- Streamlit app





1. Fetching the data

- *Fetch YouTube Data* function to retrieve metadata: title, description, link
- 10 videos/ query
- Getting an API Key
- Queries grouped by category
- Key words for each category
- Function to fetch videos for all categories – storing in a dictionary and converting to PDF

Total of 6 categories:

- Mindfulness and Relaxation
- CBT
- Somatic Practices
- Productivity and Focus
- Sleep and Rest
- General Well-Being

2. Cleaning and Processing

- Data Loading and Split into Chunks (28)
- Sanitizing and Filtering
- Embeddings (LangChain) + Vector Database (Chroma)
 - Each chunk embedded using text-embedding-3-large from OpenAI
 - Vectors capture semantic meaning of the text

3. Building the App

- Streamlit Setup
- Loading Chroma database allowing for fast semantic searches using OpenAI embeddings
- Handling User Queries
 - Retrieve_chunks function searches the database for relevant chunks
 - Generating Responses
 - *Generate_response function constructs a detailed prompt and queries GPT-4, including:*
 - *Context: Relevant Videos from the database*
 - *Guidelines: Instructions for generating actionable and clear responses*
- Displaying Results

The Query Processing Pipeline:

User Query:

- The app collects user input
 - *"I want to feel less stressed"*.

Similarity Search:

- Using vector similarity, the query is compared against the pre-organized video database.
- Videos across all categories are searched, but those with descriptions matching stress-related terms are more likely to surface.

Recommended Videos:

- For a query like *"I want to feel less stressed"*, the app might return:
 - "10-Minute Guided Meditation" (from "Mindfulness and Relaxation").
 - "Yoga for Stress Relief" (from "Physical and Somatic Practices").
 - "Progressive Muscle Relaxation" (from "Physical and Somatic Practices").

So how would this be better than a simple Google/ YouTube Search?

Imagine this:

- You're feeling overwhelmed with work, can't sleep, and decide to search for relaxation techniques on YouTube.
- You type "how to relax" or "reduce stress," and YouTube suggests generic videos like "Relaxing Music" or unrelated content.
- **Lack of Personalization:** YouTube's algorithm may recommend content unrelated to your specific need, as it's based on past behavior, not your current query.
- **Welcome to this app!** Where you will not just find videos—you will find videos that are **actionable, tailored, and immediately useful.**



What to do Next?

Improve and Expand Video Selection:

- Revise the fetching keywords to ensure they cover diverse and relevant topics.
- Allow the app to fetch new videos dynamically via YouTube API.

Making the App more Dynamic:

- With YouTube API Integration and NLP processing to score the relevance of video descriptions



The background is a light beige color. It features several organic, rounded shapes with white outlines. These shapes contain different textures: a large purple one on the top left, a brown one with small dark specks in the center, a dark brown one at the top right, a green one with fine fibers at the bottom left, and a brown one with fine fibers at the top right. The text "Thank You" is centered on the right side of the image.

Thank You