Populate a Pinecone vector database from a website

# Populate a Pinecone vector database from a website#

Use n8n to scrape a website, load the data into Pinecone, then query it using a chat workflow. This workflow uses the HTTP node to get website data, extracts the relevant content using the HTML node, then uses the Pinecone Vector Store node to send it to Pinecone.

## Key features#

This workflow uses:

• HTTP node: fetches website data.

• HTML node: simplifies the data by extracting the main content from the page.

• Pinecone Vector Store node and Embeddings OpenAI: transform the data into vectors and store it in Pinecone.

• Chat Trigger and Question and Answer Chain to query the vector database.

## Using the example#

To load the template into your n8n instance:

• Download the workflow JSON file.

• Open a new workflow in your n8n instance.

• Copy in the JSON, or select Workflow menu > Import from file....

The example workflows use Sticky Notes to guide you:

• Yellow: notes and information.

• Green: instructions to run the workflow.

• Orange: you need to change something to make the workflow work.

• Blue: draws attention to a key feature of the example.