



FM

FLOWMATRIX.DEV

Premium Workflow

A practical guide – step by step

Knowledge Base Ingestion

Turn docs and FAQs into a searchable knowledge base: chunk content, generate embeddings and power semantic retrieval.

AI & SUPPORT

ADVANCED

ABOUT 3 HOURS SETUP

OVERVIEW

Knowledge Base Ingestion

Turn docs and FAQs into a searchable knowledge base: chunk content, generate embeddings and power semantic retrieval.

CATEGORY	DIFFICULTY	DURATION
AI & Support	Advanced	About 3 hours setup

What you use it for

Support bots, internal search and RAG assistants.

Tools in this workflow

Cohere Embed API · Pinecone Vector API · Supabase · n8n

How to implement it

1

Collect source documents

Gather all relevant documents that should be included in the knowledge database, such as FAQs, manuals, and informative texts.

WHAT YOU NEED TO DO

1. Identify the documents that are relevant for the knowledge database.
2. Upload the documents to a central repository.
3. Verify the completeness and currency of the documents.

CHECK AT THE END

- ☐ All relevant documents have been identified.
- ☐ Documents are uploaded to the central repository.

RESULT

All relevant documents are collected and ready for processing.

LINKS

Cohere Embed API: <https://docs.cohere.ai/embed>

2

Chunk and clean content

Split collected documents into smaller sections and remove irrelevant information to improve data quality.

WHAT YOU NEED TO DO

1. Use a script to split documents into smaller chunks.
2. Clean the chunks of irrelevant information.
3. Save the cleaned chunks in a structured format.

CHECK AT THE END

- ☐ Documents are successfully split into chunks.
- ☐ Irrelevant information is removed.

RESULT

Documents are split into clean chunks and ready for embedding generation.

LINKS

Pinecone Vector API: <https://docs.pinecone.io/docs>

3

Generate embeddings

Generate vectors for cleaned chunks that capture the semantic meaning of the content.

WHAT YOU NEED TO DO

1. Set up the Cohere Embed API and generate API keys.
2. Use the API to generate embeddings for each chunk.
3. Save the generated embeddings locally or in a database.

CHECK AT THE END

- ☐ API key for Cohere Embed API is set up.
- ☐ Embeddings are generated for all chunks.

RESULT

Embeddings are generated and ready for storage with metadata.

LINKS

Cohere Embed API: <https://docs.cohere.ai/embed>

4

Store vectors with metadata

Save generated vectors along with their metadata in a database for efficient querying.

WHAT YOU NEED TO DO

1. Set up the Pinecone Vector API and generate API keys.
2. Create a database table for storing vectors and metadata.
3. Save the vectors and metadata to the database.

CHECK AT THE END

- [] API key for Pinecone Vector API is set up.
- [] Vectors and metadata are stored in the database.

RESULT

Vectors and metadata are stored in the database and ready for querying.

LINKS

Supabase: <https://supabase.com/docs>

5

Expose a retrieval endpoint

Set up an endpoint that allows users to access the knowledge database and retrieve relevant information.

WHAT YOU NEED TO DO

1. Set up a server that provides query endpoints.
2. Implement logic for semantic search using stored vectors.
3. Test the endpoint with various search queries.

CHECK AT THE END

- [] Server is set up and running.
- [] Endpoint has been tested and functions correctly.

RESULT

A functioning endpoint is set up to access the knowledge database.

LINKS

n8n: <https://n8n.io/docs>

RECOMMENDED TOOLS & RESOURCES

Recommended Tools & Resources

All tools in this workflow with their official website.

Cohere Embed API

<https://cohere.com/>

Pinecone Vector API

<https://www.pinecone.io/>

Supabase

<https://supabase.com/?ref=flowmatrix>

Official URL

n8n

<https://n8n.io/?ref=flowmatrix>

Official URL

Full version as Premium PDF

All steps, tools and resources as a PDF – €4.99.

[Buy now – €4.99](#)