Zep Vector Store node documentation

# Zep Vector Store node#

Use the Zep Vector Store to interact with Zep vector databases. You can insert documents into a vector database, get documents from a vector database, retrieve documents to provide them to a retriever connected to a chain, or connect it directly to an agent to use as a tool.

On this page, you'll find the node parameters for the Zep Vector Store node, and links to more resources.

Credentials

You can find authentication information for this node here.

Examples and templates

For usage examples and templates to help you get started, refer to n8n's Zep Vector Store integrations page.

Parameter resolution in sub-nodes

Sub-nodes behave differently to other nodes when processing multiple items using an expression.

Most nodes, including root nodes, take any number of items as input, process these items, and output the results. You can use expressions to refer to input items, and the node resolves the expression for each item in turn. For example, given an input of five name values, the expression {{ $json.name }} resolves to each name in turn.

name

{{ $json.name }}

In sub-nodes, the expression always resolves to the first item. For example, given an input of five name values, the expression {{ $json.name }} always resolves to the first name.

name

{{ $json.name }}

## Node usage patterns#

You can use the Zep Vector Store node in the following patterns.

### Use as a regular node to insert, update, and retrieve documents#

You can use the Zep Vector Store as a regular node to insert or get documents. This pattern places the Zep Vector Store in the regular connection flow without using an agent.

You can see an example of this in scenario 1 of this template (the example uses Supabase, but the pattern is the same).

### Connect directly to an AI agent as a tool#

You can connect the Zep Vector Store node directly to the tool connector of an AI agent to use vector store as a resource when answering queries.

Here, the connection would be: AI agent (tools connector) -> Zep Vector Store node.

### Use a retriever to fetch documents#

You can use the Vector Store Retriever node with the Zep Vector Store node to fetch documents from the Zep Vector Store node. This is often used with the Question and Answer Chain node to fetch documents from the vector store that match the given chat input.

An example of the connection flow (the example uses Pinecone, but the pattern in the same) would be: Question and Answer Chain (Retriever connector) -> Vector Store Retriever (Vector Store connector) -> Zep Vector Store.

### Use the Vector Store Question Answer Tool to answer questions#

Another pattern uses the Vector Store Question Answer Tool to summarize results and answer questions from the Zep Vector Store node. Rather than connecting the Zep Vector Store directly as a tool, this pattern uses a tool specifically designed to summarizes data in the vector store.

The connections flow (this example uses Supabase, but the pattern is the same) in this case would look like this: AI agent (tools connector) -> Vector Store Question Answer Tool (Vector Store connector) -> Zep Vector store.

## Node parameters#

### Operation Mode#

This Vector Store node has four modes: Get Many, Insert Documents, Retrieve Documents (As Vector Store for Chain/Tool), and Retrieve Documents (As Tool for AI Agent). The mode you select determines the operations you can perform with the node and what inputs and outputs are available.

In this mode, you can retrieve multiple documents from your vector database by providing a prompt. The prompt will be embedded and used for similarity search. The node will return the documents that are most similar to the prompt with their similarity score. This is useful if you want to retrieve a list of similar documents and pass them to an agent as additional context.

Use Insert Documents mode to insert new documents into your vector database.

Use Retrieve Documents (As Vector Store for Chain/Tool) mode with a vector-store retriever to retrieve documents from a vector database and provide them to the retriever connected to a chain. In this mode you must connect the node to a retriever node or root node.

Use Retrieve Documents (As Tool for AI Agent) mode to use the vector store as a tool resource when answering queries. When formulating responses, the agent uses the vector store when the vector store name and description match the question details.

### Insert Documents parameters#

• Collection Name: Enter the collection name to store the data in.

### Get Many parameters#

• Collection Name: Enter the collection name to retrieve the data from.

• Prompt: Enter the search query.

• Limit: Enter how many results to retrieve from the vector store. For example, set this to 10 to get the ten best results.

10

### Retrieve Documents (As Vector Store for Chain/Tool) parameters#

• Collection Name: Enter the collection name to retrieve the data from.

### Retrieve Documents (As Tool for AI Agent) parameters#

• Name: The name of the vector store.

• Description: Explain to the LLM what this tool does. A good, specific description allows LLMs to produce expected results more often.

• Collection Name: Enter the collection name to retrieve the data from.

• Limit: Enter how many results to retrieve from the vector store. For example, set this to 10 to get the ten best results.

10

## Node options#

### Embedding Dimensions#

Must be the same when embedding the data and when querying it.

This sets the size of the array of floats used to represent the semantic meaning of a text document.

Read more about Zep embeddings in Zep's embeddings documentation.

### Is Auto Embedded#

Available in the Insert Documents Operation Mode, enabled by default.

Disable this to configure your embeddings in Zep instead of in n8n.

### Metadata Filter#

Available in Get Many mode. When searching for data, use this to match with metadata associated with the document.

This is an AND query. If you specify more than one metadata filter field, all of them must match.

AND

When inserting data, the metadata is set using the document loader. Refer to Default Data Loader for more information on loading documents.

## Templates and examples#

by Jimleuk

by David Roberts

by David Roberts

## Related resources#

Refer to LangChain's Zep documentation for more information about the service.

View n8n's Advanced AI documentation.

## AI glossary#

• completion: Completions are the responses generated by a model like GPT.

• hallucinations: Hallucination in AI is when an LLM (large language model) mistakenly perceives patterns or objects that don't exist.

• vector database: A vector database stores mathematical representations of information. Use with embeddings and retrievers to create a database that your AI can access when answering questions.

• vector store: A vector store, or vector database, stores mathematical representations of information. Use with embeddings and retrievers to create a database that your AI can access when answering questions.