
Drupal and OpenAI

— Unlocking New Capabilities in
Drupal with OpenAI Integration —

Vincenzo Gambino

Drupal Developer from London, UK.

15 years experience with Drupal.

Speaker at DrupalCamp London, DrupalCamp Munich, DrupalCamp Poland, DrupalCamp Switzerland, DrupalJam, Drupal Dev Days Vienna, DrupalCamp Florida, Chicago and Stanford.

Co-author of the Book: Jumpstart Jamstack Development: Build and deploy modern websites and web apps using Gatsby, Sanity.io, and Netlify.

Agenda

- AI Terminology
- Drupal AI Suite
 - AI Core
 - AI CKEditor
 - AI Explorer
 - AI Search
 - AI Translate
 - AI Assistant
- Demo
- Q&A

AI Terminology

— Prompts, Token, Embeddings
and more.. —

LLM - Large Language Model

LLMs are artificial intelligence models trained on vast amounts of text data. They can generate human-like responses, complete sentences, and assist in tasks such as text summarization, translation, and coding. Some multimodal models can also generate images and videos.

Examples of LLMs:

- GPT-4 (OpenAI)
- LLaMA (Meta)
- Mistral (Mistral AI)
- Claude (Anthropic)
- PaLM 2 (Google)

Prompts

Prompt: A prompt is essentially the way you "program" the model, typically by providing instructions or examples to guide its response.

Simple Prompt: "List three benefits of AI in education."

Prompt Engineering: "Explain three key benefits of AI in education. For each benefit, provide a real-world example of an AI-powered tool (e.g., adaptive learning platforms, automated grading systems, or personalized tutoring assistants) and describe how it enhances the learning experience for students and supports teachers in their tasks."

Temperature

Temperature controls the predictability of the model's responses. Lower values (e.g., 0.2) make the answers more predictable and factual, while higher values (e.g., 0.8) introduce more variability and creativity.

Temperature example

Prompt: "Write an article title about AI and healthcare."

Response with Temperature (0.2):

"AI Revolutionizes Healthcare by Improving Diagnostics."

Response with Temperature (0.8):

"The Evolving Role of AI in Medicine: Shaping the Future of Patient Care."

Response with Temperature (2):

"Cybernetic Doctors and AI Healers: Is the Future of Medicine a Sci-Fi Dream?"

Hallucination

Hallucination in AI occurs when the model generates information that appears plausible but is actually incorrect or irrelevant.

Prompt: "Who won the Nobel Prize in Physics in 2026?"

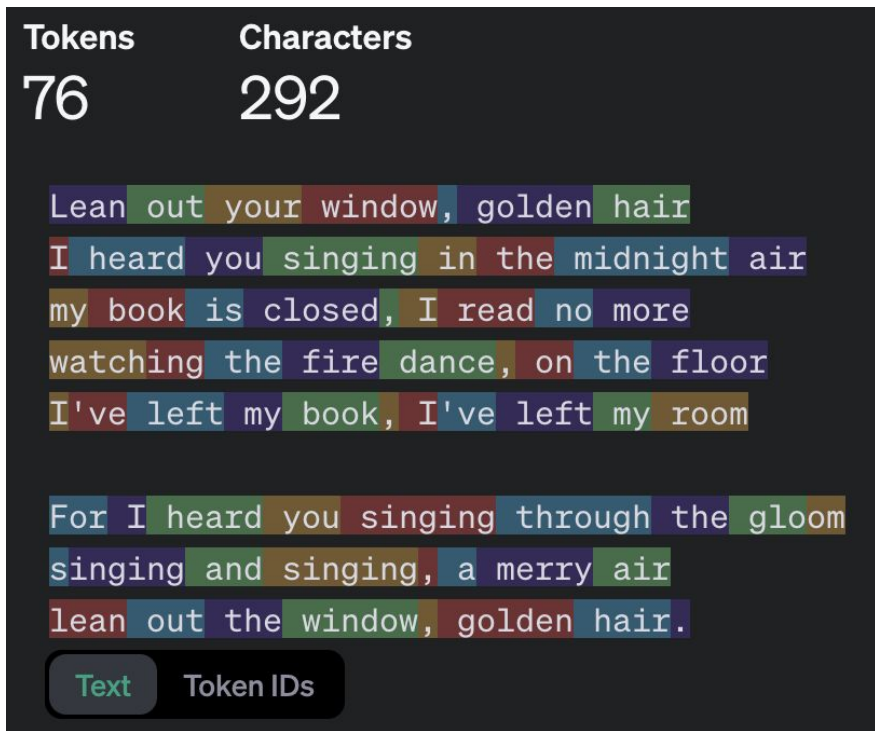
AI Response: "The 2026 Nobel Prize in Physics was awarded to Dr. Jane Doe for her work in quantum computing."

Tokens

A token is a piece of text that has been parsed out for further processing. Tokens can be words, phrases, symbols, or any meaningful piece of text.

These tokens can then be converted into a numeric representation, enabling AI models to process and understand language.

Examples from <https://platform.openai.com/tokenizer>



James Joyce
Lean out your window

Token IDs

Tokens	Characters
76	292

```
[49862, 704, 701, 3321, 11, 21411, 7013, 198, 40, 6755, 499, 26139, 304, 279, 33433, 3805, 198, 2465, 2363, 374, 8036, 11, 358, 1373, 912, 810, 198, 14658, 287, 279, 4027, 15612, 11, 389, 279, 6558, 198, 40, 3077, 2163, 856, 2363, 11, 358, 3077, 2163, 856, 3130, 271, 2520, 358, 6755, 499, 26139, 1555, 279, 14015, 316, 198, 82, 24992, 323, 26139, 11, 264, 93552, 3805, 198, 2742, 704, 279, 3321, 11, 21411, 7013, 13]
```

Text

Token IDs

Retrieval-Augmented Generation (RAG)

Combines the strengths of pre-trained language models with information retrieval systems.

Retriever: Queries a Vector Database

Generator: Produces more informed and contextually relevant responses

Embeddings

An embedding is a vector (list) of floating point numbers. The distance between two vectors measures their relatedness. Small distances suggest high relatedness and large distances suggest low relatedness.

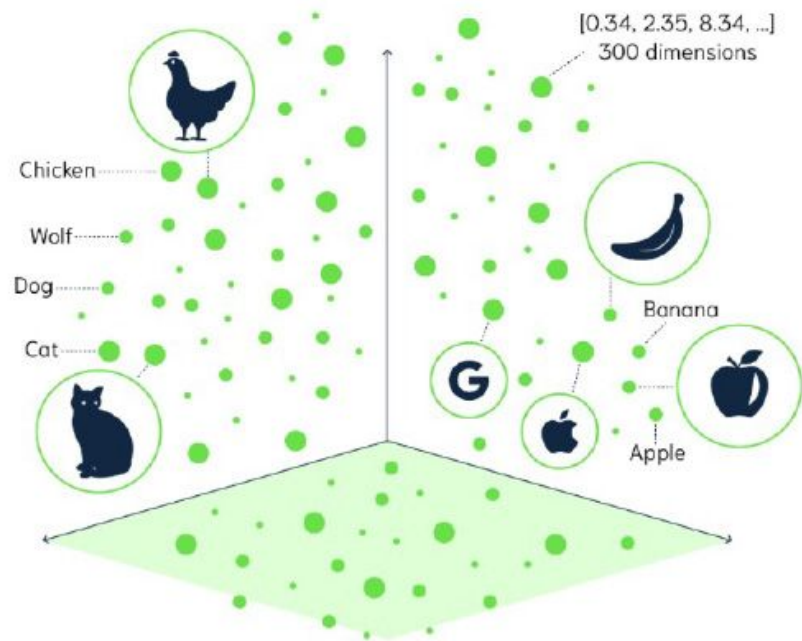
- Search (where results are ranked by relevance to a query string)
- Recommendations (where items with related text strings are recommended)
- Classification (where text strings are classified by their most similar label)

Embedding Example

Imagine we encode words into vector space based on their meaning:

- Point A (2,3) → "Cat"
- Point B (2,4) → "Dog"
- Point C (7,1) → "Car"

Since "cat" and "dog" are semantically related, their vector coordinates are close together, whereas "car" is in a distant location because it belongs to a completely different category.



Pinecone

Pinecone is a managed vector database designed for handling complex AI and machine learning tasks.

It allows for efficient storage, retrieval, and analysis of vector embeddings, which are high-dimensional data points representing items in a way machines can understand.



Pinecone

Original Text

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse tristique metus quis nulla luctus tempus. Suspendisse nec leo a nisi blandit pharetra vel et dolor. In hac habitasse platea dictumst. Cras molestie condimentum nisl et tincidunt. Proin vitae faucibus dui, eget efficitur nulla. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Cras porta ultrices maximus. Cras non vehicula erat. Proin faucibus feugiat leo.

Text Splitter

Chunk of Text

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse tristique metus quis nulla luctus tempus. Suspendisse nec leo a nisi blandit pharetra vel et dolor.

In hac habitasse platea dictumst. Cras molestie condimentum nisl et tincidunt. Proin vitae faucibus dui, eget efficitur nulla. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas.

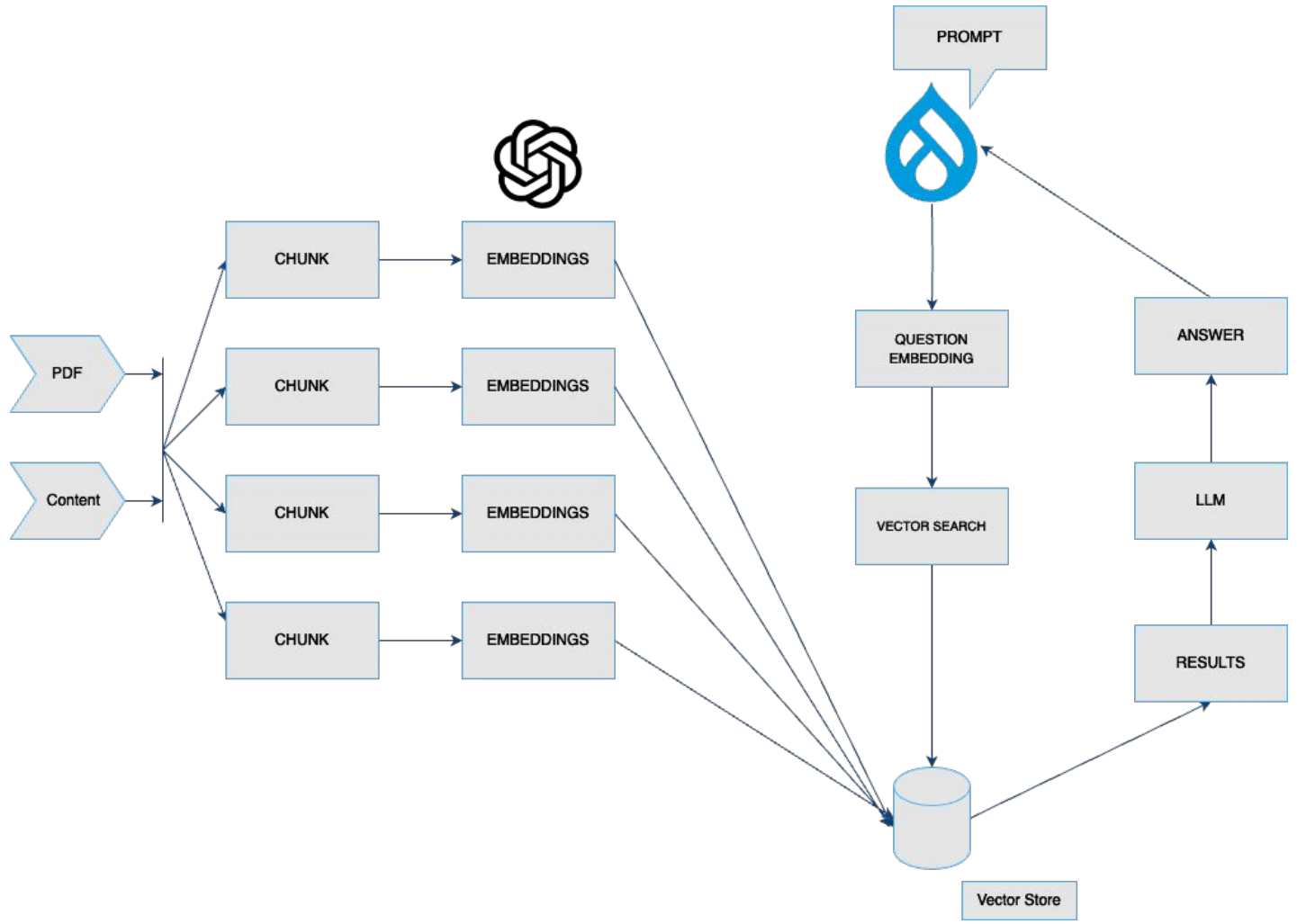
Cras porta ultrices maximus. Cras non vehicula erat. Proin faucibus feugiat leo.

EMBEDDINGS

[0.1, 0.2, ...]

[0.3, 0.1, ...]

[0.2, 0.5, ...]



Introduction to OpenAI

- OpenAI is a research laboratory that develops a safe artificial general intelligence (AGI) for the benefit of humanity.
- Founded in 2015
- Generative AI
 - Generates text
 - Generates Images
 - Generates Video
 - Generates Audio

OpenAI - Models

Dall-E: A model that can generate and edit images

Whispers/TTS: A model that can convert audio into text and viceversa

Embedding: A set of models that can convert text into a numerical form

Moderation: A model that can detect if a text may be sensitive or unsafe

GPT: A set of models that can understand and generate natural language

Sora: A model that can generate videos.

How to Interact with GPT

OpenAI Chat Application: <https://chat.openai.com/>

OpenAI API: <https://platform.openai.com/docs/api-reference>

OpenAI API Key Concepts

Prompts: are essentially how you “program” the model, usually by providing some instructions or a few examples.

Models: models are advanced artificial intelligence systems.

Token: are units of text that the language model reads and processes.

Temperature: What sampling temperature to use, between 0 and 2.

OpenAI API Prerequisites

- OpenAI Account
- API Key
- Payment method

Endpoints examples

Method: POST

Completions: <https://api.openai.com/v1/completions>

Embeddings: <https://api.openai.com/v1/embeddings>

Image generation: <https://api.openai.com/v1/images/generations>

Text to Speech: <https://api.openai.com/v1/audio/speech>

Completion Request body

model: ID of the model to use.

messages: A list of messages comprising the conversation so far. Including roles such as system, developer and user.

Max_completion_token (optional): The maximum number of tokens to generate in the completion.

Temperature (optional): What sampling temperature the LLM should use, between 0 and 2.

Example Call

```
curl https://api.openai.com/v1/chat/completions \  
  -H "Content-Type: application/json" \  
  -H "Authorization: Bearer $OPENAI_API_KEY" \  
  -d '{  
    "model": "gpt-4o",  
    "messages": [  
      {  
        "role": "developer",  
        "content": "You are a helpful assistant."  
      },  
      {  
        "role": "user",  
        "content": "Hello!"  
      }  
    ]  
  }'
```

Example Response

```
{
  "id": "chatcmpl-123",
  "object": "chat.completion",
  "created": 1677652288,
  "model": "gpt-4o",
  "system_fingerprint": "fp_44709d6fcb",
  "choices": [{
    "index": 0,
    "message": {
      "role": "assistant",
      "content": "Hello there, how may I assist you today?",
    },
    "logprobs": null,
    "finish_reason": "stop"
  }],
  "usage": {
    "prompt_tokens": 9,
    "completion_tokens": 12,
    "total_tokens": 21,
    "completion_tokens_details": {
      "reasoning_tokens": 0,
      "accepted_prediction_tokens": 0,
      "rejected_prediction_tokens": 0
    }
  }
}
```

Drupal and AI

— The modules —

AI (Artificial Intelligence)

Supported by FreelyGive. Maintained by Jamie Abraham, Marcus Johansson, Kevin Quillen, Scott Euser, Michal Z Gow and Frederik Wouter.

The Drupal AI module provides a framework for easily integrating Artificial Intelligence on any Drupal site using any kind of AI (from multiple vendors). The AI module aims to provide a suite of modules and an API foundation for generating text content, images, content analysis and more.

AI - Installation

To install the AI module to our Drupal 11 application, run the Composer command to add the AI module to the codebase. It will also download the Drupal module Key

```
composer require 'drupal/ai:^1.0'
```

Then run:

```
drush en ai -y
```

OpenAI Provider - Installation

To install the AI module to our Drupal 11 application, run the Composer command to add the AI module to the codebase.

```
composer require 'drupal/ai_provider_openai:^1.0'
```

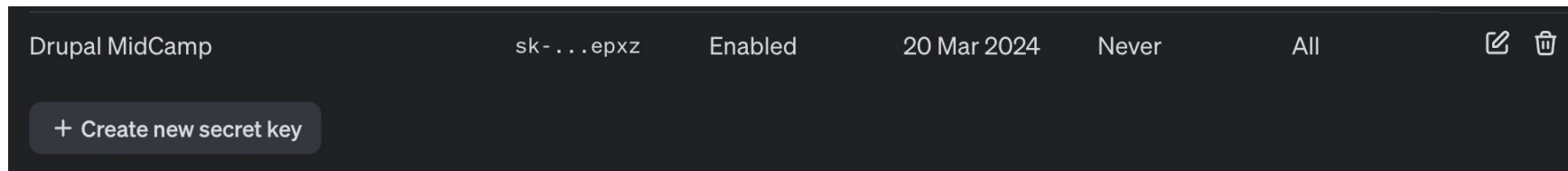
Then run:

```
drush en ai_provider_openai -y
```


Generate an OpenAI key

URL: <https://platform.openai.com/account/api-keys>

Create a New key clicking on the Create new secret key button



Generate an OpenAI key - Name

Create new secret key

Name Optional

Project

Permissions

All Restricted Read Only

Cancel Create secret key

OpenAI Key - Generated

Save your key

Please save this secret key somewhere safe and accessible. For security reasons, **you won't be able to view it again** through your OpenAI account. If you lose this secret key, you'll need to generate a new one.

```
sk-1zQdyQNTnDB2tx7c3j4CT3B1bkFJw9Qb0VC16xWQ
```

 Copy

Permissions

Read and write API resources

Done

Add your key

Configuration -> System -> Keys -> Add (/admin/config/system/keys/add)

Home > Administration > Configuration > System > Keys

Add key ☆

Key name *

OpenAI Machine name: [openai](#) [Edit](#)

Description

A short description of the key.

^ **Type settings**

Key type *

Authentication ▾

A generic key type to use for a password or API key that does not belong to any other defined key type.

^ **Provider settings**

Key provider *

Configuration ▾

The Configuration key provider stores the key in Drupal's configuration system.

^ **Value**

Key value

sk-nj-oiu23469owefjwefalohoinhweolundchq-KHd586wUruIDn-wB5fc


Save

Add the key to the provider settings

Home > Administration > Configuration > AI > AI Providers

Setup OpenAI Authentication

OpenAI API Key

Choose an available key. If the desired key is not listed, [create a new key](#). The API Key. Can be found on <https://platform.openai.com/>.

 **Advanced settings**

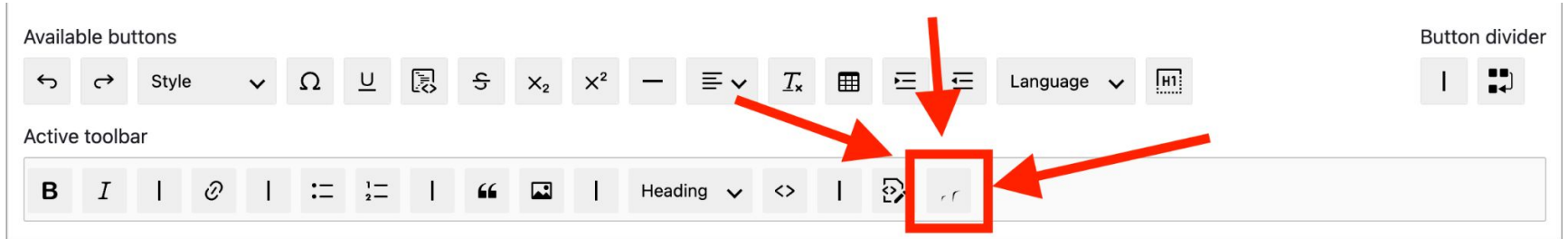
Save configuration

Enable AI modules

- Keys
- AI Core
- AI CKEditor integration
- AI Content Suggestions
- AI Translate
- AI API Explorer
- OpenAI Provider
- AI Search
- VBD Pinecone
- AI Assistant
- AI Chatbot

Add the AI CKEditor button

I will add the AI CKEditor button to the Basic HTML Text Format.
Go to Configuration -> Content authoring -> Text formats and editors
(/admin/config/content/formats/manage/basic_html).



The image shows the CKEditor configuration interface. It is divided into two main sections: "Available buttons" and "Active toolbar".

- Available buttons:** A horizontal row of buttons including undo, redo, style dropdown, Ω, U, a list icon, a link icon, x_2 , x^2 , a minus sign, a list icon with a dropdown arrow, I_x , a table icon, a list icon, a list icon, a language dropdown, and an H1 icon.
- Active toolbar:** A horizontal row of buttons including B, I, a vertical bar, a link icon, a vertical bar, a list icon, a list icon, a vertical bar, a quote icon, an image icon, a vertical bar, a heading dropdown, a code icon, a vertical bar, a list icon, and a button with a red square icon and a vertical bar.

Three red arrows point from the "Available buttons" section to the button with the red square icon in the "Active toolbar". One arrow points to the list icon with a dropdown arrow, another points to the button with the red square icon, and a third points to the button with the red square icon and a vertical bar. The button with the red square icon and a vertical bar is highlighted with a red square border.

On the right side, there is a "Button divider" section with two buttons: a vertical bar and a square with a red square icon.

Add the AI CKEditor button configuration

CKEditor 5 plugin settings

AI tools	Modal dialog options
Headings	Translate
Source editing	Generate with AI
List	Fix spelling
Image Images can only be uploaded.	Summarize
Image resize	Help and Support
	Reformat HTML
	Tone

Add the AI CKEditor button configuration

^ Generate with AI

Get ideas and text completion assistance from AI.

Enabled

Enable this editor feature.

AI provider

OpenAI - gpt-4o



Select which provider to use for this plugin. See the [Provider overview](#) for details about each provider.

Completion pre prompt

This prompt will be prepended before the user prompt. This field may be left empty too.

AI Content Suggestion

/admin/config/ai/suggestions

[Home](#) > [Administration](#) > [Configuration](#) > [AI](#)

AI Content Suggestions settings ☆

Below is a list of the available plugins you can use to analyze your content.

Suggest taxonomy tags

Enable Suggest taxonomy tags.

Suggest taxonomy tags model.

OpenAI - gpt-4

Suggest taxonomy prompt (not limited to vocabulary) *

Suggest no more than five words to classify the following text using the same language as the input text. The words must be nouns or adjectives in a comma delimited list

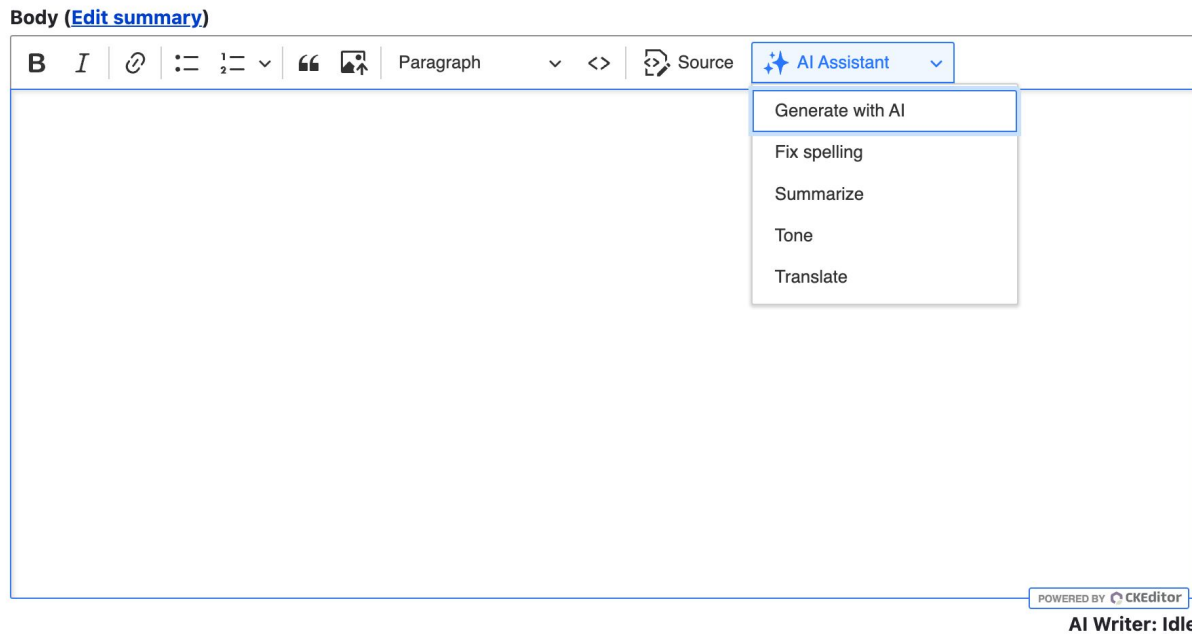
Suggest taxonomy prompt (limited to vocabulary) *

Choose no more than five words to classify the following text using the same language as the input text:

Allow an LLM to suggest SEO-friendly taxonomy tags for the content.

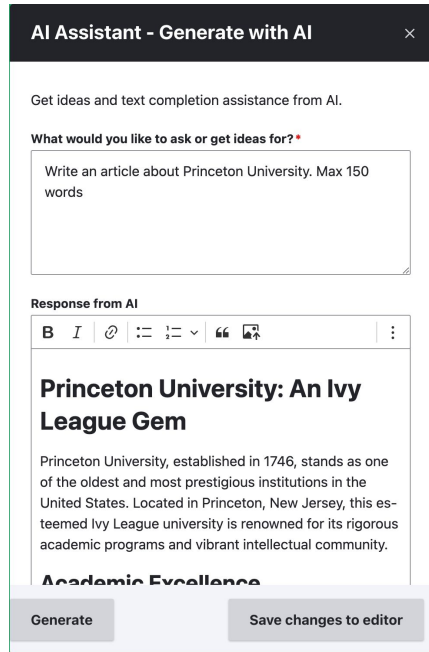
Using the AI CKEditor button

Now that we have the AI button added to our text editor, we can explore the options available:



Text Completion

Choose one of the options, for example Text Completion for now to generate the content.



Change the tone

With this option it is possible to change the tone of the text.

AI Assistant - Tone

Convert tone of selected text.

Choose tone*

- ✓ - Select -
- Friendly
- New York Slang
- Professional

Selected text to convert

Princeton University, a member of the prestigious Ivy League, is renowned for its rich history and tradition of academic excellence. Established in 1446, it is one of the oldest institutions of higher learning in the United States. Located in Princeton, New Jersey, the university is con-

Suggested conversion

B *I* @ :: := v “ ” ↕

Change the tone Save changes to editor

Translate

With this option it is possible to translate the text.

AI Assistant - Translate ×

Translate the selected text into other languages.

Choose language*

- ✓ - Select -
- German
- Spanish

These options will translate the selected text.

Selected text to translate

Princeton University, a member of the prestigious Ivy League, is renowned for its rich history and tradition of academic excellence. Established in 1446, it is one of the oldest institutions of higher learning in the United States. Located in Princeton, New Jersey, the university is consistently ranked among the top universities globally.

Suggested translation

B *I*

Translate Save changes to editor

AI Content Suggestions

This module functionalities will be displayed in the right hand side column of the content form.

✓ Suggest taxonomy tags
✓ Summarise text
✓ Moderate text
✓ Evaluate Readability
✓ Suggest title
✓ Alter tone

Suggest Taxonomy

^ Suggest taxonomy tags

Title
Body
Description

Select the field(s) you wish to send to the LLM

Use source vocabulary
Check this box if you want to use a source vocabulary to suggest terms.

Choose vocabulary

Tags ▾

Optionally, select which vocabulary do you want to find the terms in.

Use source vocabulary's full hierarchy
Check this box if you want to take into account the selected vocabulary's hierarchy, if such exists.

Suggest taxonomy terms

Translation

Home > Princeton University History

Translations of Princeton University History ☆

View

Edit

Delete

Revisions

Translate

<input type="checkbox"/>	Language	Translation	Status	AI Translations	Operations
<input type="checkbox"/>	English (Original language)	Princeton University History	Published	NA	<button>Edit</button>
<input type="checkbox"/>	Italian	n/a	Not translated	Translate using gpt-4	<button>Add</button>

What happens now

— What can we do? —

What can we also do with the content?

- Improve the search experience
- Training GPT with our own data
- Embeddings
- RAG

Search API and VDB Pinecone

Search API integrates Drupal sites with AI-powered search functionalities through AI Search and Pinecone.

Installation:

```
composer require 'drupal/search_api:^1.38'
```

```
composer require 'drupal/ai_vdb_provider_pinecone:^1.0@beta'
```

Generate Pinecone Key

API keys

+ Create API key

Name	Created on	Created by	Value	Permissions	Actions
Demo	04/03/2025	Vincenzo Gambino	pcsk_4fLsHG_*****	All	...
default	29/02/2024	---	*****_*****_*****	All	...
test	29/02/2024	---	*****_*****_*****	All	...

Generate Pinecone API Key

API keys + Create API key

Generate new API key

Specify the name and permissions for your API key.

API key name

DrupalCampNJ

Permissions ⓘ

All Custom

Cancel Create key

Permissions	Actions
All	...
All	...
All	...

Copy the Key

API keys

Name	Permissions
Demo	All
default	All
test	All
DrupalCamp	All

API key generated

Your "**DrupalCampNJ**" API key has been generated.

```
*****_****_****_****_*****
```

⚠ You must copy and save this key immediately
For security reasons, we will not show you this key again after you close this dialog.
Please save it somewhere secure for future use.

Close

Add Key to Drupal

Home > Administration > Configuration > System > Keys

Edit key *Pinecone* ☆

Key name*

Machine name: pinecone

Description

A short description of the key.

^ Type settings

Key type*

A generic key type to use for a password or API key that does not belong to any other defined key type.

^ Provider settings

Key provider*

The Configuration key provider stores the key in Drupal's configuration system.

^ Value

Key value

Save


Delete

Pinecone VDB Provider

Go to [Home](#) > [Administration](#) > [Configuration](#) > [AI](#) > [Vector Database Providers](#)

Configure Pinecone

API Key

Choose an available key. If the desired key is not listed, [create a new key](#). The API key to use for authentication. This can be created and found under "API keys" at <https://app.pinecone.io/>.

Save configuration

Search API Server

Server name *

Machine name: pinecone

Enter the displayed name for the server.

Enabled

Only enabled servers can index items or execute searches.

Description

Enter a description for the server.

Backend *

AI Search

Index items on Vector DB.

Choose a backend to use for this server.

Search API Server

^ Configure *AI Search* backend

Embeddings Engine *

OpenAI | text-embedding-3-small ▾

The service to use for generating the embeddings (the vectorized representations of each chunk of your content). If you change this, everything will be needed to be reindexed. Larger models tend to provide more complete representations of the content and therefore more accurate results, but are however slower (and for paid models, typically with a slightly higher cost). The general idea here is that the engine creates vectorized representations of your chunks of content, then vectorize the user's query in the same manner (i.e., using the same engine) to mathematically compare the vectors and find the nearest matches.

Tokenizer chat counting model

OpenAI - gpt-4 ▾

This is recommended to ensure the right number of tokens is calculated for the embeddings. Depending on the vector database and dimensions, the number of Tokens allowed per chunk of content differs. This service is used to count the number of tokens in your content as accurately as possible to better make use of the available space.

Vector Database *

Pinecone DB ▾

The Vector Database to use. This is where the generated Embeddings (vectorized representations of your content) are stored. The user's queries are then vectorized in the same manner and the mathematical distance between the query and the vectors stored in the database are compared to find the nearest results.

Search API Index

^ General

Label	Machine name	Property path	Type	Indexing option	Operations
<input type="text" value="URI"/>	<input type="text" value="url"/>	search_api_url	String ▾	Contextual content ▾	Edit Remove

^ Content

Label	Machine name	Property path	Type	Indexing option	Operations
<input type="text" value="Body"/>	<input type="text" value="body"/>	body	Fulltext ▾	Main content ▾	Remove
<input type="text" value="Title"/>	<input type="text" value="title"/>	title	String ▾	Contextual content ▾	Remove

How does a record on Pinecone looks like

Edit record

Namespace

drupalmountaincamp

ID

entity:node/1:en:0

Values

-0.016289521,0.0134660043,0.0257374421,0.0598911382,0.04425319

Metadata

content

Princeton University, a member of the prestigious Ivy League, is renowned for its rich history and tradition

Aa



Create the new assistant

Home > Administration > Configuration > AI > AI Assistant configuration

Edit an AI Assistant

Label*

Machine name: search_articles

This is the title of the AI Assistant

Enabled

Instructions

You are a chatbot that can answer questions about XYZ.

Do not have opinions about things that the user asks for, just give back information in a friendly manner. This includes questions about XYZ that you have in your own knowledge.

Always look up from the database what the user is asking, even if it seems absurd and only base your answers on what they write.

This means that when you are in preprompt, you should always respond with a json and the action "search_rag". When the assistant have gotten some results, you should look at them and try to figure out an answer.

You know that results have been given when there is an assistant message with the following start "The following are the results the different actions from the rag action"

Do not use your own knowledge to provider information about XYZ, just the context you are given.

You only ever answer questions based on context from searches you were given, you do not make up your own conclusions. If the context that you receive does not answer the question just answer with "I am sorry, I do not know that, would you like to ask something else?"

⌵ Agent Actions

⬆ RAG Actions

Configure the RAG Actions settings for this AI assistant.

Enable RAG Actions

Add the Assistant as a block

Home > Administration > Structure > Block layout

Configure block ☆

Block description

AI Chatbot

Title *

Machine name: olivero_aichatbot_2

Display title

This code type will be replaced with the AI Deepchat Chatbot, we recommend to use that instead.

AI Assistant

Select the AI Assistant to use for this chat form. You can create new [AI Assistants here](#).

Example


By [admin](#), 4 March 2025

Princeton University History


Princeton University: A Tradition of Excellence

Princeton University, a member of the prestigious Ivy League, is renowned for its rich history and tradition of academic excellence. Established in 1746, it is one of the oldest institutions of higher learning in the United States. Located in Princeton, New Jersey, the university is consistently ranked among the top universities globally.

Search Articles



 **admin** 21:27:59

Talk about Princeton unvieristy

 **Article searcher** 21:27:59

Princeton University is one of the oldest and most prestigious universities in the United States, being a member of the Ivy League. It was established in 1446 and is located in Princeton, New Jersey.

For more detailed information, you can visit the [Princeton University History page](#).

Ask me a question  

Demo Time

Recap

- The status of AI Modules in Drupal.
- How to connect to API and configure the modules.
- How to train the LLM with our data using RAG.

Resources

Drupal AI: <https://www.drupal.org/project/ai>

OpenAI Provider: https://www.drupal.org/project/ai_provider_openai

Pinecone VDB: https://www.drupal.org/project/ai_vdb_provider_pinecone

OpenAI: <https://openai.com>

OpenAI Documentation: <https://platform.openai.com/docs/introduction>

Examples: <https://platform.openai.com/examples>

ChatGPT: <https://chat.openai.com>

Q&A

SAVE THE DATE



Drupal Camp Italy

Rome 2025

20th of June 2025

Sapienza University of Rome