

# Concept | Prompt Studios and Prompt recipe

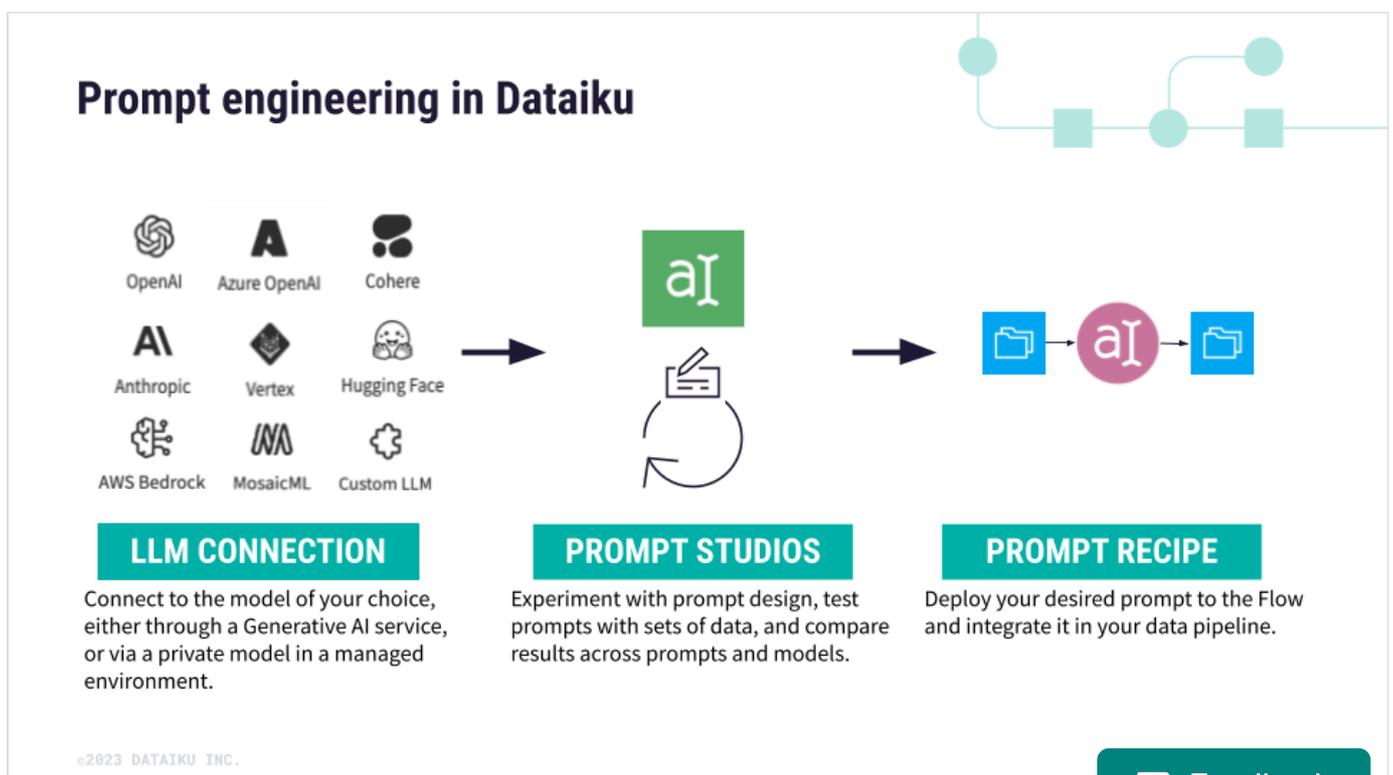
Prompts are the input to a Generative AI model such as a Large Language Model (LLM). You can modify and craft prompts to achieve a desired outcome — a process known as **prompt engineering**.

Prompts guide the model's behavior and output, but they don't modify the underlying model.

## Prompt engineering in Dataiku

In Dataiku, **Prompt Studios** and the **Prompt recipe** work together to test and iterate on LLM prompts, then operationalize them in your data pipeline.

Prompt Studios and the Prompt recipe allow you to securely connect to LLMs either through a Generative AI service or a private model, including open-source models. You can first experiment with different prompt designs in Prompt Studios, before deploying a desired prompt in your data pipeline with the Prompt recipe.

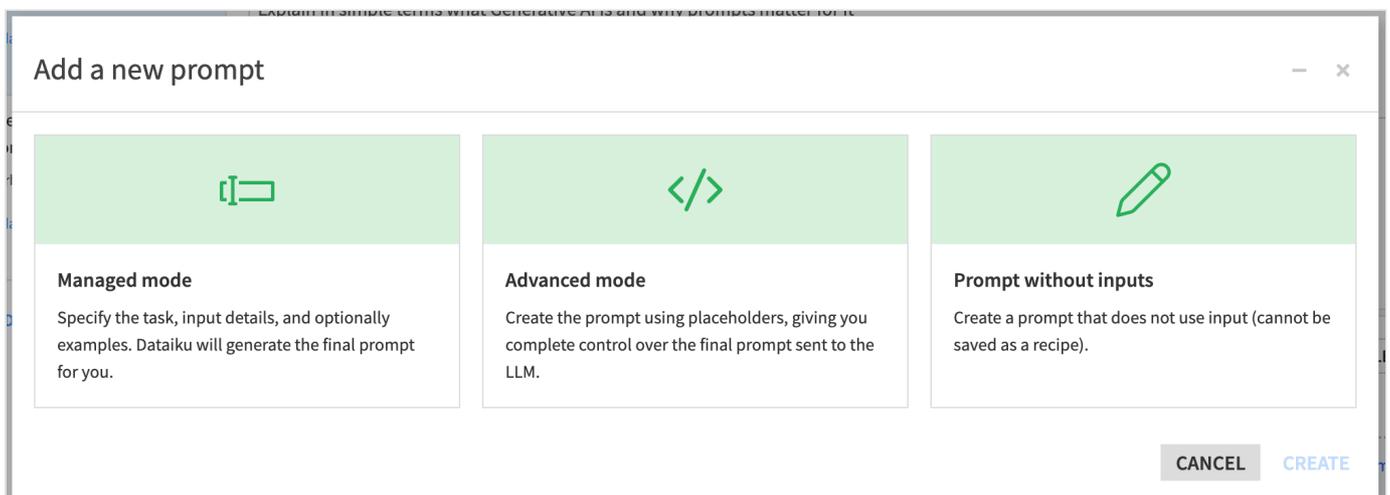


# Prompt Studios

**Prompt Studios** is the playground where you can iterate and test different prompts and models to engineer the optimal prompt for a specific use case. You can find Prompt Studios in the top navigation bar under the Visual Analyses menu.

Within a Prompt Studio, you can create multiple prompts and iterations, and even compare them. For each new prompt, you can choose from three different modes:

Mode	Description
<b>Managed mode</b>	You specify the task using plain language, optionally include other details, and Dataiku generates the final prompt for you. You can use templates for common tasks.
<b>Advanced mode</b>	You create the prompt and examples using placeholders, giving you more control over the final prompt sent to the LLM.
<b>Prompt without inputs</b>	You create a one-off prompt that queries the LLM and is designed for quick experimenting. These prompts can't be converted into a Prompt recipe and their results cannot be compared in Prompt Studio.



## Designing prompts

The **Prompt design** steps will differ depending on the mode you're using. In general, this process includes the elements described below.

Element	Description
Prompt	Includes the main instructions to the LLM.
Inputs	Lists the variables inserted into the prompt. You can map the inputs to columns from a dataset and use values from the dataset as test cases to check the prompt's output. You can also choose to manually write test cases.
Examples	Provides sample input-output pairs to guide the model's expected results.
Test cases	Evaluates prompt performance using either dataset rows or manually written examples.

You can quickly iterate on your prompt design, including updating the task or adding new inputs, examples, and test cases, before deploying your prompt.

Each time you run inference on your prompt and test cases, the model will also return an estimated cost for 1,000 records, helping you manage costs associated with LLM usage.

The screenshot shows the Prompt Studio interface for a project titled "Financial headlines". The interface is divided into several sections:

- Prompt design:** Shows the LLM model as "GPT-4o" and the prompt: "Give the topic of each article based on the headline and short description." It includes an "Inputs from" dropdown set to "reuters\_headlines" and a table mapping "headline" to "Headlines" and "description" to "Description".
- Examples:** Displays a table with columns "Input: headline", "Input: description", and "Output". An example shows the headline "U.S. economy faces significant risks" and the output "economy, pandemic".
- Test cases:** Shows "8 records" with a status of "8 OK (no validation rules)". It includes a table with columns "Headlines", "Description", and "Result from last run". Two rows are visible:
 

Headlines	Description	Result from last run
TikTok considers London and other locations for headquarters	TikTok has been in discussions with the UK government over the past few months to locate its headquarters in London, a source familiar with the matter said, as part of a strategy to distance itself from its Chinese ownership.	business, technology, corporate relocation
Disney cuts ad spending on Facebook amid growing boycott: WSJ	Walt Disney has become the latest company to slash its advertising spending on Facebook Inc as the social media giant faces an ad boycott over its handling of hate speech and controversial content, the	social media, advertising boycott

## Using images in prompts

You can also process images within Prompt Studios for use cases where visual input is useful.

To do this, you will need:

- Dataiku version 13.2 or above.
- A connection to an LLM that supports images as inputs.
- Images stored in a managed folder, which you can connect to within Prompt Studios. (See [How to | Prepare images for use in a model](#) for instructions.)

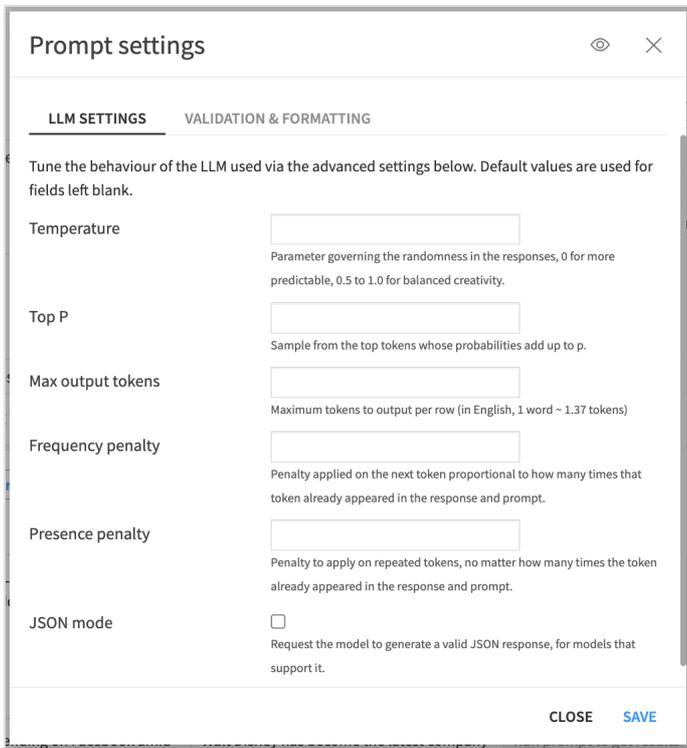
The screenshot displays the Dataiku Prompt Studios interface for a task named "classify dog breeds". The main area shows a "Prompt design" section with a "Prompt" field containing the instruction: "Give the breed of each dog shown in the images." Below the prompt is a large image of two dogs, a Boston Terrier and a Chihuahua, in a room with framed pictures. To the right, a settings panel is open, showing "Inputs from" set to "dog\_breed\_images\_files" and "Image folder" set to "dog\_breed\_images". Below this, a table lists input columns: "image" with a description of "image" and a column of "path". At the bottom, a "Result from last run" section shows a table with four rows of results:

Result from last run
The dogs in the image appear to be a Boston Terrier (black and white) and a Chihuahua (tan).
This dog appears to be a Beagle.
The dog in the image appears to be a Golden Retriever.
This dog appears to be a Border Collie.

## Settings

You can change several hyperparameters or set validation rules for the prompt.

Tab	Description
<b>LLM settings</b>	Changes hyperparameters such as temperature or the maximum output tokens, that control the behavior of the LLM.
<b>Validation &amp; formatting</b>	Sets rules such as validating the expected output format as a JSON object or forbidding certain words from the output. You will receive a warning if the output is not validated.



## Comparing prompts

Prompt Studios track your prompts in the left sidebar. You can navigate between prompts, view each prompt's history, restore previous versions, and duplicate prompts.

When using managed or advanced modes, you can also compare the output of multiple prompts by selecting them in the sidebar and clicking **Compare**. Prompt Studios will build a table comparing the outputs and costs of the selected prompts.

aī
Prompt Studios
←

**COMPARE**

Give the topic of each article based on the headline and short description.  
● GPT 3.5 Chat (Turbo / ChatGPT)  
8 OK (no validation rules) - \$0.247/1K records

Determine whether each topic of the following list of topics is covered in the financial news article provided. List of topics: fed and central banks, company and product news, corporate debt and ear...  
● GPT 3.5 Chat (Turbo / ChatGPT)  
8 OK (no validation rules) - \$0.323/1K records

+ ADD PROMPT

**Comparing**

CLOSE COMPARISON

Give the topic of each article based on the headline and short description.

● openai:maxime\_openAI:gpt-3.5-turbo

8 OK (no validation rules) - \$0.247/1K records

Determine whether each topic of the following list of topics is covered in the financial news article provided. List of topics: fed and central banks, company and product news, corporate debt and ear...

● openai:maxime\_openAI:gpt-3.5-turbo

8 OK (no validation rules) - \$0.324/1K records

Headlines	Description		
TikTok considers London and other locations for headquarters	TikTok has been in discussions with the UK government over the past few months to locate its headquarters in London, a source familiar with the matter said, as part of a strategy to distance itself from its Chinese ownership.	Corporate headquarters and TikTok	<p>Topics covered in the article:</p> <ul style="list-style-type: none"> <li>- Company and product news (TikTok considering London and other locations for headquarters)</li> <li>- Politics (discussions with the UK government)</li> <li>- Legal and regulation (strategy to distance itself from its Chinese ownership)</li> </ul> <p>Topics not covered in the article:</p> <ul style="list-style-type: none"> <li>- Fed and central banks</li> <li>- Corporate debt and earnings</li> <li>- Energy and oil</li> <li>- Currencies</li> <li>- Gold and metals</li> <li>- IPO</li> <li>- M&amp;A and investments</li> <li>- Markets</li> <li>- Stock movement</li> </ul>
Disney cuts ad spending on Facebook amid growing boycott: WSJ	Walt Disney has become the latest company to slash its advertising spending on Facebook Inc as the social media giant faces an ad boycott over its handling of hate speech	Advertising boycott on Facebook	<p>Topics covered in the article: company and product news, corporate debt and earnings, legal and regulation, M&amp;A and investments, politics</p>

i

Tip

Prompt Studios also tracks each run of a prompt as you iterate. To view the version history of a prompt and revert back to any version, click on the clock icon in the top right of the Prompt design screen.

# Prompt recipe

The Prompt recipe puts your prompt into action on a dataset. It generates an output dataset and appears in the Flow.

You can create a new Prompt recipe directly from Prompt Studios by saving a prompt template that is mapped to a dataset as a recipe. This allows you to experiment with your prompts before operationalizing them with the recipe.

You can also create a new Prompt recipe directly from the Flow or from the Actions panel of a dataset. With this method, you can write a Structured prompt or a Text prompt directly in the recipe, with the same settings as a prompt template in a Prompt Studio.

The screenshot shows the 'Prompt and LLM settings' configuration page for a recipe named 'compute\_headline\_summaries'. The interface includes a top navigation bar with 'Settings', 'Input / Output', 'Advanced', and 'History' tabs, and a 'SAVED!' indicator. The main configuration area is divided into several sections:

- LLM:** Set to 'GPT 3.5 Chat (Turbo / ChatGPT) - I'.
- Prompt:** A text area containing 'Give the topic of each article based on the headline and short description.' with a 'SHOW RAW PROMPT' button.
- Inputs:** A table with two columns: 'Description' and 'Column'.

Description	Column
headline	Headlines
description	Description
- Examples:** A table with three columns: 'Input: headline', 'Input: description', and 'Output'.

Input: headline	Input: description	Output
Lured by rising SUV sales, automake	Demand for sport utility vehicles in t	Automotive markets
- LLM completion settings:** Includes fields for 'Temperature', 'Top P', and 'Max output tokens' with descriptive text below each.
- Result validation settings:** (Partially visible at the bottom).

A green 'RUN' button is located at the bottom left of the configuration area.

## What's next?

Continue learning about prompt engineering with LLMs by working through the [Tutorial | Prompt engineering with LLMs](#) article.