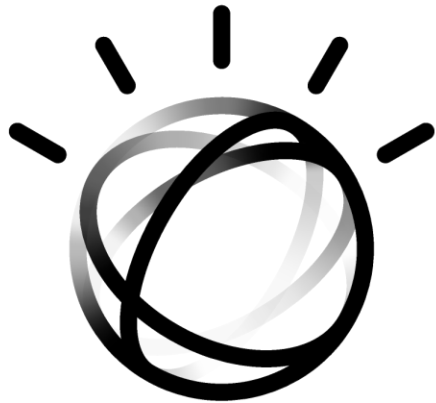


Watsonx AI

cloud.ibm.com

STAT 5350/4350

11/5/2025



IBM Watson®

Recap

- Explain system prompt, user prompt, assistant prompt
- Explain zero, one & few shot prompts
- Define RAG, when is helpful to use it?

Recap

Prompts

...

Compare

Optimize

Evaluate

Save

Your presets

Model

gpt-4.1-mini

response_format: text temp: 1.00 tokens: 2048 top_p: 1.00

store: false

Functions

+ Add

System message

You are a friendly tutor who explains concepts simply and clearly.

Prompt messages

User

Please summarize the following text in 3 bullet points:
{{input_text}}

+ Add message

Your conversation will appear here

Explain quantum computing

Auto-clear

Pre-reqs

- Watsonx.ai Studio
- Watsonx-ai Runtime

Cloud.ibm.com

Catalog /

watsonx.ai Studio

(Formerly known as Watson Studio) Develop powerful AI solutions with an integrated collaborative studio and industry-standard APIs and SDKs.



Create

About

Type
Service

Provider
IBM

Last updated
05/06/2025

Category
AI / Machine
Learning

Compliance
HIPAA Enabled
IAM-enabled

Location
Sydney (au-
syd)
Frankfurt (eu-
de)
London (eu-gb)
Tokyo (jp-tok)
Dallas (us-
south)
Toronto (ca-
tor)

Active Promotion

You've already applied a promo code that gives you \$200 toward this service. [Track your promo](#)

Select a location

Dallas (us-south)

Select a pricing plan

Prices shown are for country or location:

Plan	Features and capabilities	Pricing
Lite	1 authorized user 10 capacity unit-hours monthly limit Environment = # of capacity units required per hour <ul style="list-style-type: none">• 1 vCPU + 4 GB RAM = 0.5• 2 vCPU + 8 GB RAM = 1• 4 vCPU + 16 GB RAM = 2• Decision Optimization + Watson NLP = Environment + 5• Synthetic Data Generator, 2 vCPU + 8 GB RAM = 7 (requires watsonx.ai Runtime)	Free

Cloud.ibm.com

[Catalog](#) /

watsonx.ai Runtime

Formerly known as Watson Machine Learning. Quickly build, run and manage generative AI and machine learning applications with built-in performance and scalability.



Create

About

Type
Service

Provider
IBM

Last updated
11/05/2025

Category
AI / Machine
Learning

Compliance
HIPAA Enabled
IAM-enabled
Service
Endpoint
Supported

Location
London (eu-gb)
Dallas (us-
south)
Sydney (au-
syd)
Toronto (ca-
tor)

Active Promotion
You've already applied a promo code that gives you \$200 toward this service. [Track your promo](#)

Select a location

Dallas (us-south)

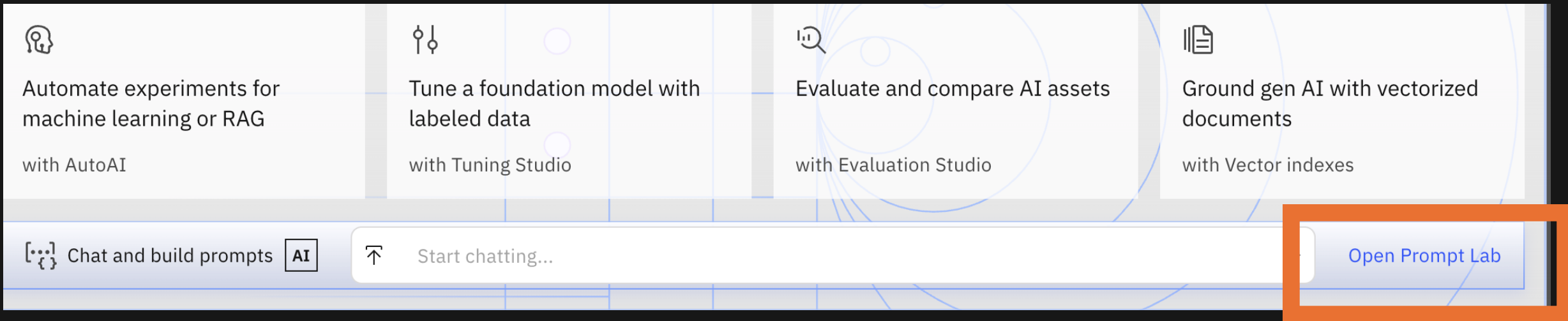
Select a pricing plan

Prices shown are for country or location:

Plan	Features and capabilities	Pricing
Lite	Service instance Instance includes: <ul style="list-style-type: none">• 20 capacity unit-hours (CUH) per month• 50,000 tokens/data points per month• 100 pages per month ----- Foundation models: <ul style="list-style-type: none">• Inferencing for text generation consumes tokens (as Resource Units)	Free



UI Tour




Get started with a task
Select a task based on your goal.

UI Tour

Discover resources


- Find samples & other resources.
- Try a sample project to take a comprehensive tutorial.

Resource hub




Explore foundation models from IBM and other third-parties.


[Explore foundation models →](#)

 Prompts


→

 Data


→

 Projects

→

 Notebooks

→

 Agents

→

[Explore samples →](#)

Developer access

Project or deployment space

Project or deployment space ▾

Project ID

00000000-0000-0000-0000-000000000000

watsonx.ai URL

https://us-south.ml.cloud.ibm.com

Get started and make your first API request to inference a foundation model

Create API key +

[Manage IBM Cloud API keys →](#)

What's new

Migrate assets to inventories in watsonx.governance

Oct 27, 2025 →

Parameter-efficient fine tuning methods are now available in the Toronto and Sydney regions

Oct 23, 2025 →

Work with new deploy on demand models in watsonx.ai

Oct 16, 2025 →

Work with the new text classification API in watsonx.ai on IBM Cloud

Oct 15, 2025 →

UI Tour

Terms & Conditions

- Understand & agree to T&Cs.

Welcome to Prompt Lab

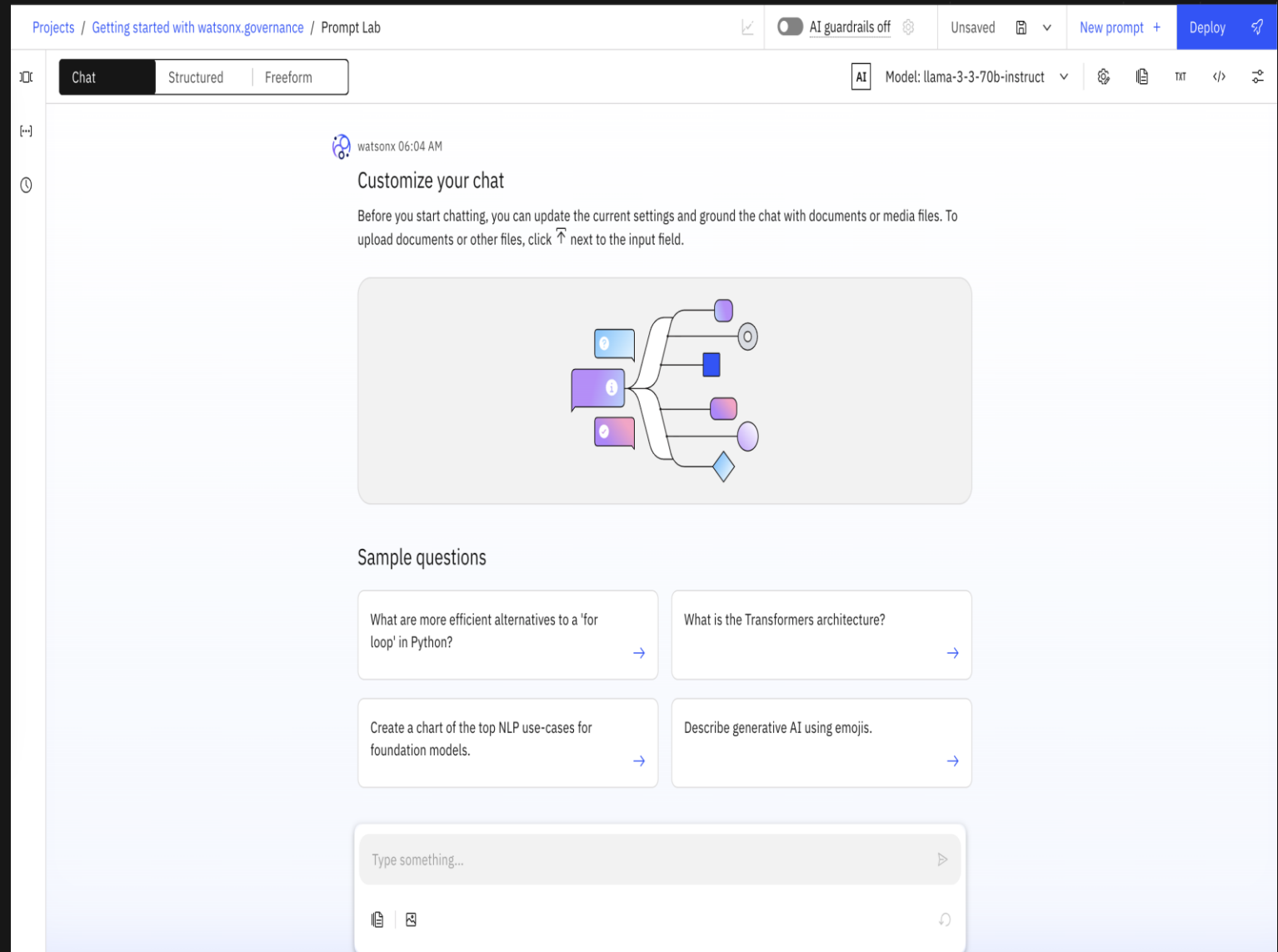
By using any foundation model provided with this Cloud Service, you acknowledge and understand that:

- ☒ Some models included in the Cloud Service are Non-IBM Products. Review the applicable model details on the third party provider and license terms that apply.
- ☐ Models may generate outputs that contain misinformation, obscene or offensive language, or discriminatory content, Client should review the outputs for such information or content prior to re-use. Users should review and validate the outputs generated.
- ☐ The output generated by all models is provided to augment, not replace, human decision-making by the Client.

UI Tour

Main Capabilities

- Chat Type.
- Models Available
- Sample prompts



Lab - 1

Structured prompt

- System prompt.
- Zero, one, few shot examples
- Try different models

[Projects](#) / [Getting started with watsonx.governance](#) / Prompt Lab

Chat

Structured

Freeform

AI

Model: granite-3-8b-instruct

{#}

TXT

</>

Unsaved

New prompt +

Hint: This model works better when you provide at least 1 example.

Set up

Instruction (optional)

Tell the model what to do. For example: Summarize the transcript.

Examples (optional)

Input:	Output:
Enter your example input here.	Enter your desired output.

Add example +

Try

Test your prompt

Input:	Output:
Enter your test input.	Generated output appears here.

New test +

Lab - 2

Freeform prompt

- System prompt.
- Zero, one, few shot examples
- Try different models

Projects / Getting started with watsonx.governance / Prompt Lab

AI guardrails off

Unsaved

New prompt +

ChatStructuredFreeform

AIModel: granite-3-8b-instruct{#}TXT</>

Hint: This model works better when you provide at least 1 example.

Below is an example of how I want you to answer questions.

Example Interaction:
User Question: What is a neural network?
Assistant Answer: A neural network is a type of computer program that learns patterns from data. Think of it like a brain made of layers of "nodes" that help recognize patterns, such as the difference between pictures of cats and dogs.

Now please answer the next question in the **same style**:

User Question: What is quantum superposition?

Stop reason: Max tokens parameter reached
Tokens: 92 input + 200 generated = 292 out of 131072
Time: 2.4 seconds

Clear output

Generate

Lab - 3

Save as:
Jupyter Notebook

[...] Prompt Lab | Part of IBM watsonx.ai®

Prompt notebook

Prompt Notebook with Chat - Prompt Lab Notebook v1.1.0

This notebook contains steps and code to demonstrate inferencing of prompts generated in Prompt Lab in watsonx.ai with a chat format. It introduces Python API commands for authentication using API key and prompt inferencing using WML API.

Note: Notebook code generated using Prompt Lab will execute successfully. If code is modified or reordered, there is no guarantee it will successfully execute. For details, see: [Saving your work in Prompt Lab as a notebook](#).

Some familiarity with Python is helpful. This notebook uses Python 3.10.

Notebook goals

The learning goals of this notebook are:

- Defining a Python function for obtaining credentials from the IBM Cloud personal API key
- Defining parameters of the Model object
- Using the Model object to generate response using the defined model id, parameters and the prompt input

Lab - 4

- Prompt Examples
- One for each category

