Watsonx AI

cloud.ibm.com

STAT 5350/4350

11/5/2025

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 • 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-

Active Promotion
You've already applied a promo code that gives you $200 toward this service. Track your promo

Select a location

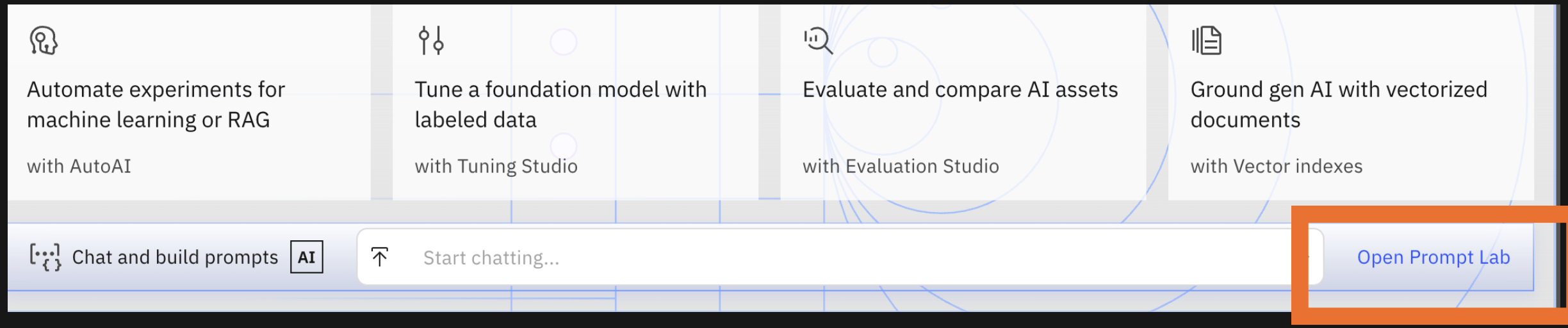
Dallas (us-south)

Select a pricing plan

Prices shown are for country or location:

Plan	Features and capabilities	Pricing
Lite	**Service instance** Instance includes: • 20 capacity unit-hours (CUH) per month • 50,000 tokens/data points per month • 100 pages per month --- Foundation models: • Inferencing for text generation consumes tokens (as Resource Units)	Free

UI Tour



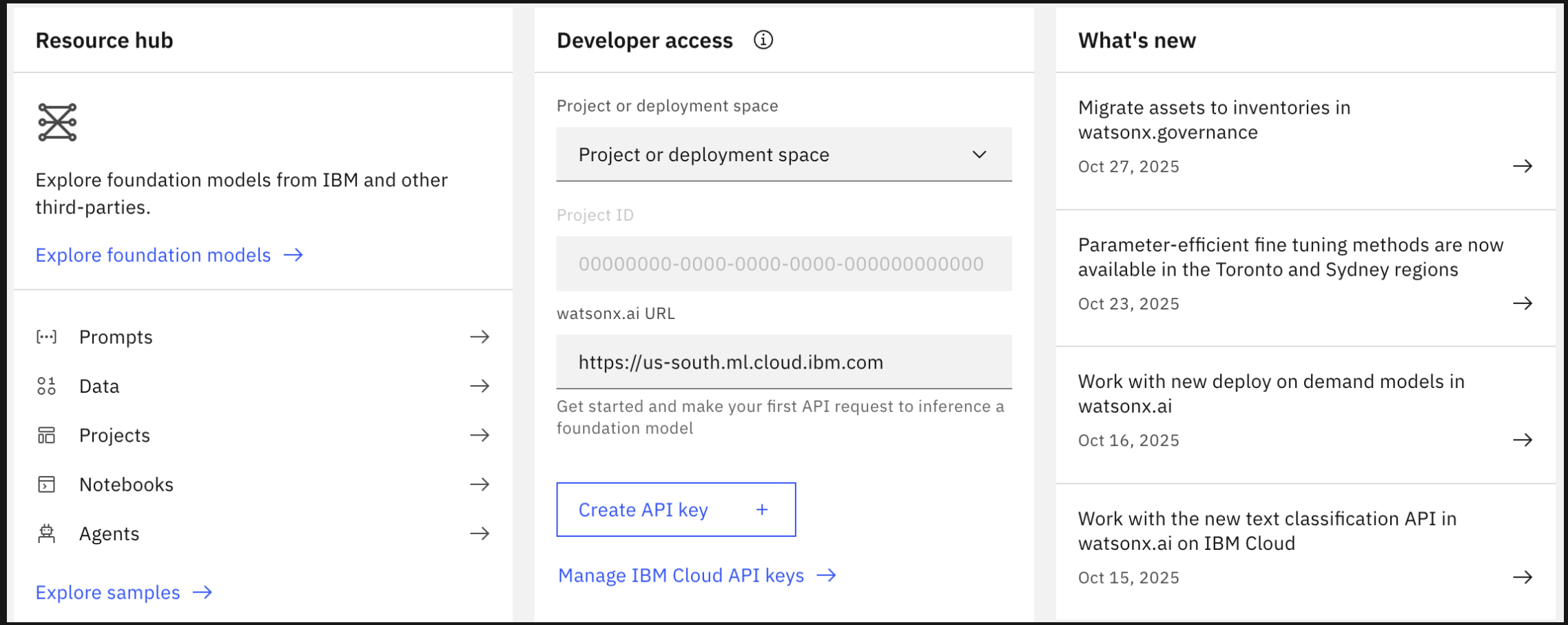
Get started with a task

Select a task based on your goal.

UI Tour

Discover resources

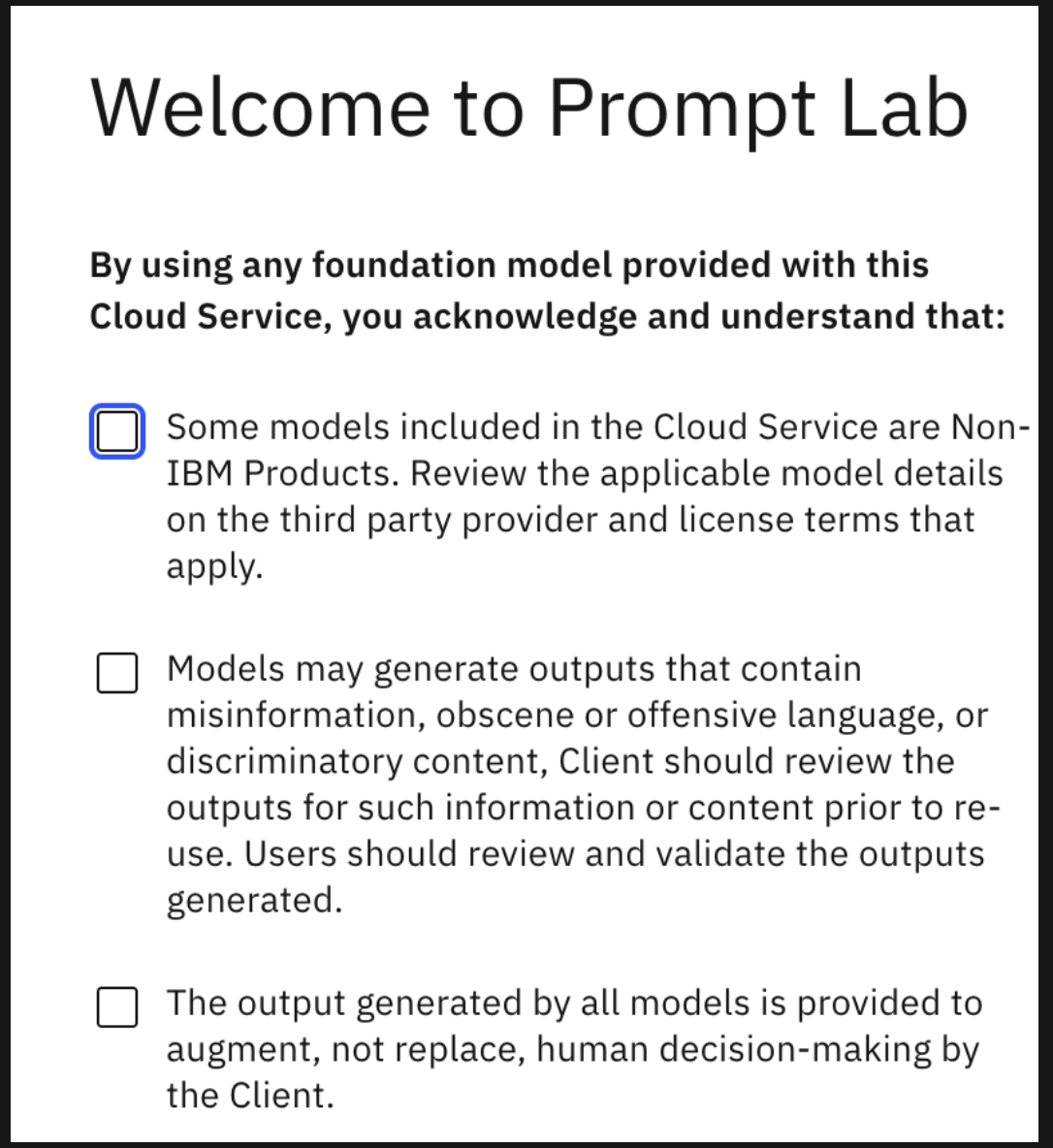
- Find samples & other resources.
- Try a sample project to take a comprehensive tutorial.



UI Tour

Terms & Conditions

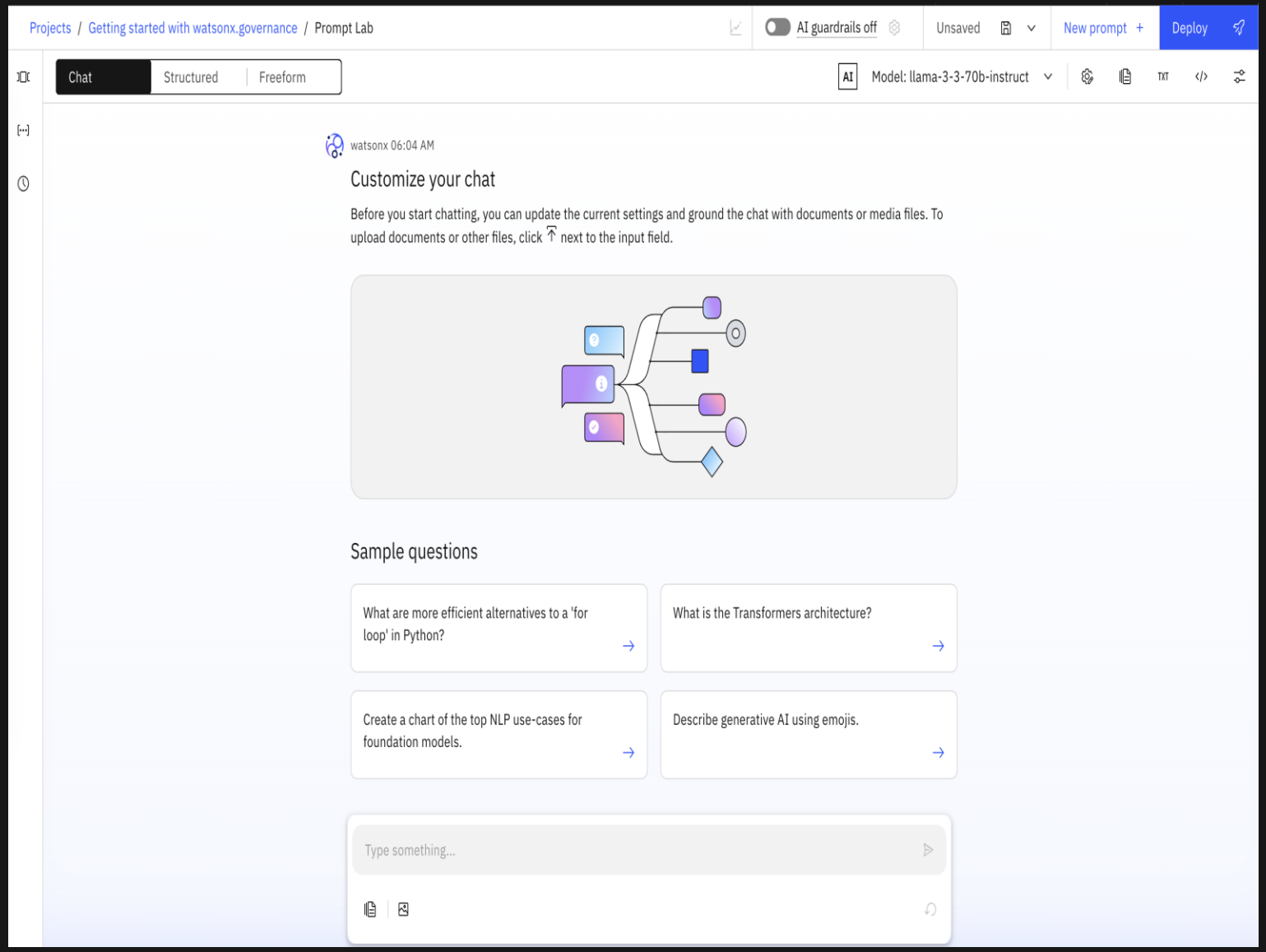
- Understand & agree to T&Cs.



UI Tour

Main Capabilities

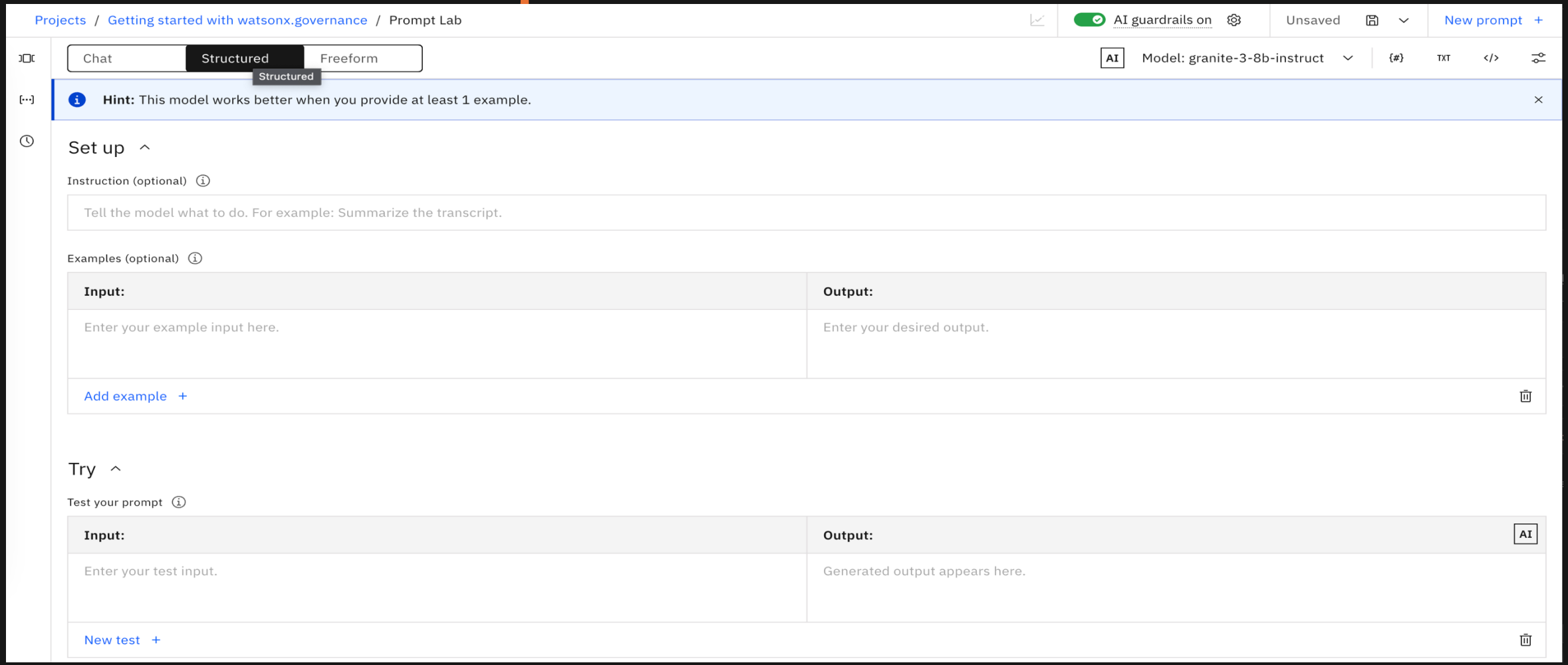
- Chat Type.
- Models Available
- Sample prompts



Lab - 1

Structured prompt

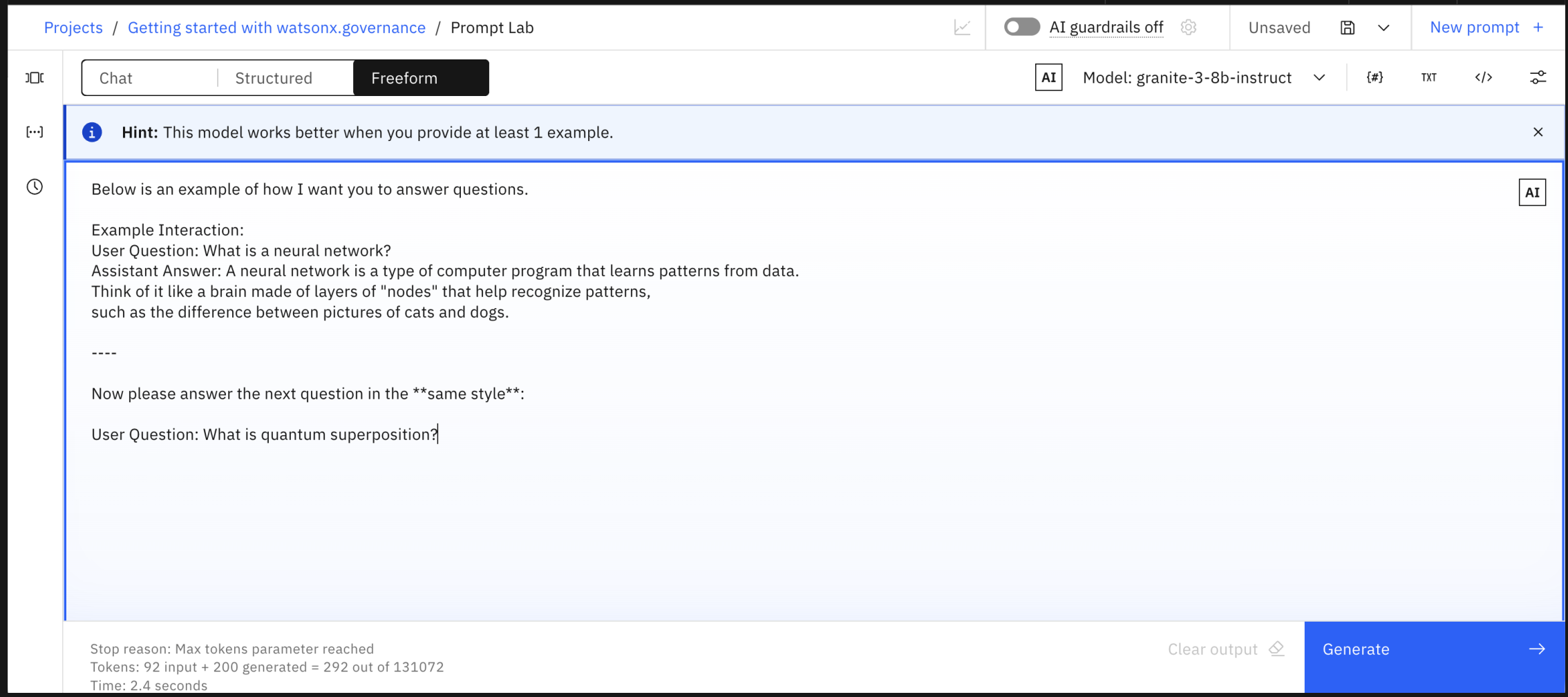
- System prompt.
- Zero, one, few shot examples
- Try different models



Lab - 2

Freeform prompt

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



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**