



Vertex AI Studio

In this module, you learn to ...

- 01 Use Vertex AI Studio to implement generative AI use cases
- 02 Explore the available models and sample use cases in Vertex AI Studio
- 03 Write and test prompts using the Google Cloud Console
- 04 Protect your data while using Vertex AI Studio



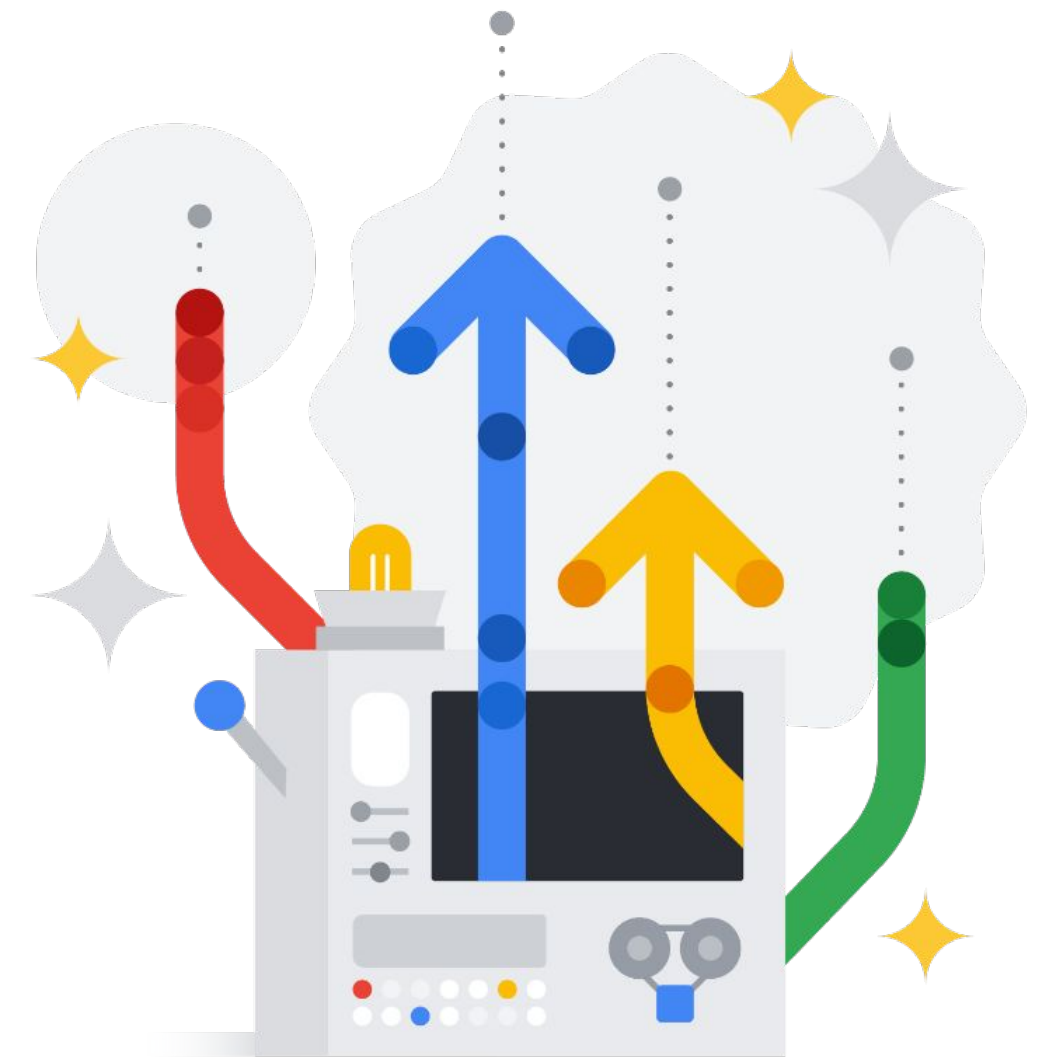
Topics

01	Vertex AI Studio
02	Designing and testing prompts
03	Data governance in Vertex Studio



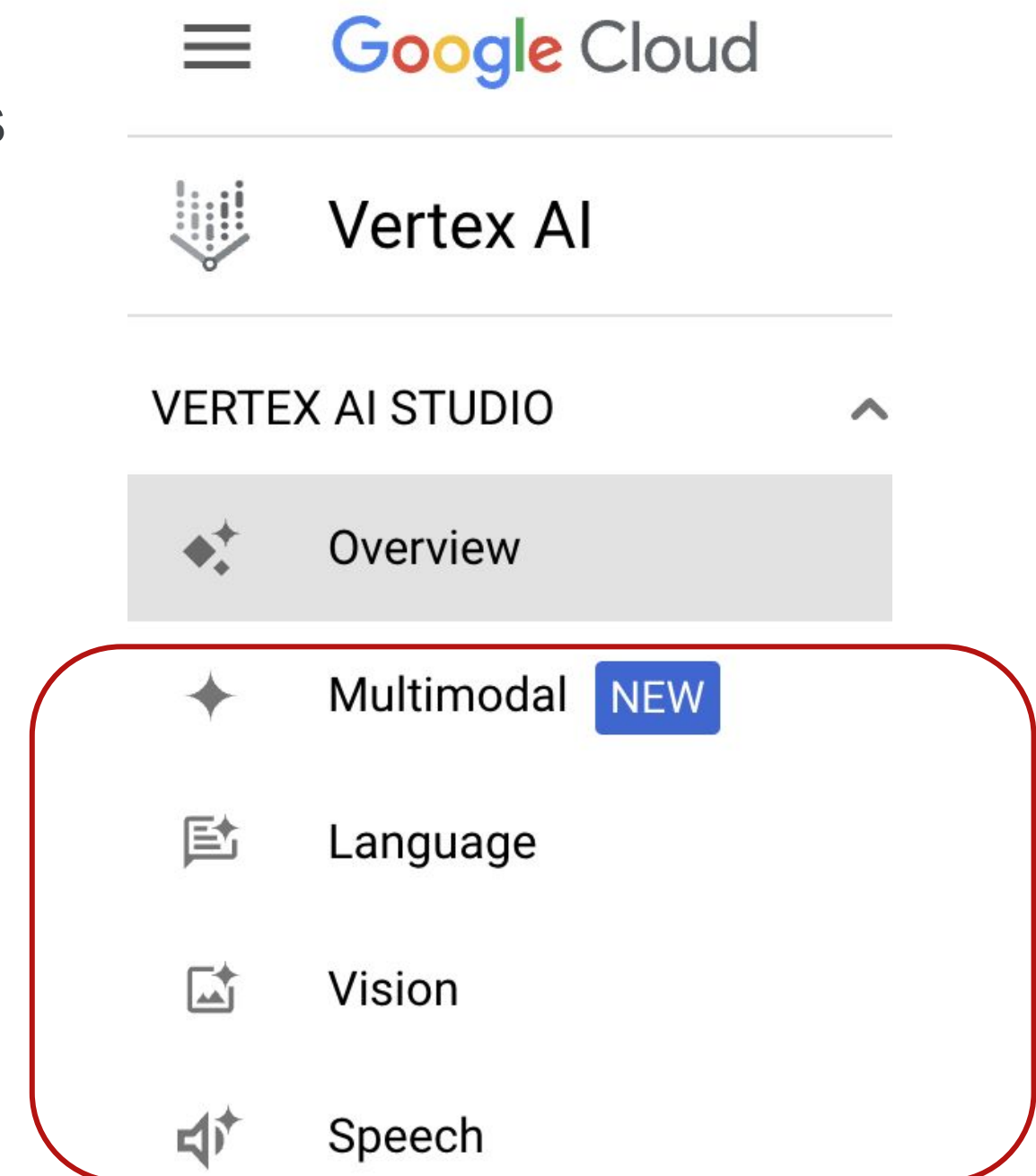
Vertex AI Studio provides an intuitive interface to prompt, tune, and deploy Google foundation models

- Supports Google foundational models
 - PaLM 2 for text and chat
 - Codey for code generation and chat
 - Chirp for speech to text and text to speech
 - Imagen for text to image generation
 - Gemini for multimodal
- Allows users to easily experiment with prompts
 - Simple, intuitive design
 - Easily experiment with parameters
 - Add context and examples

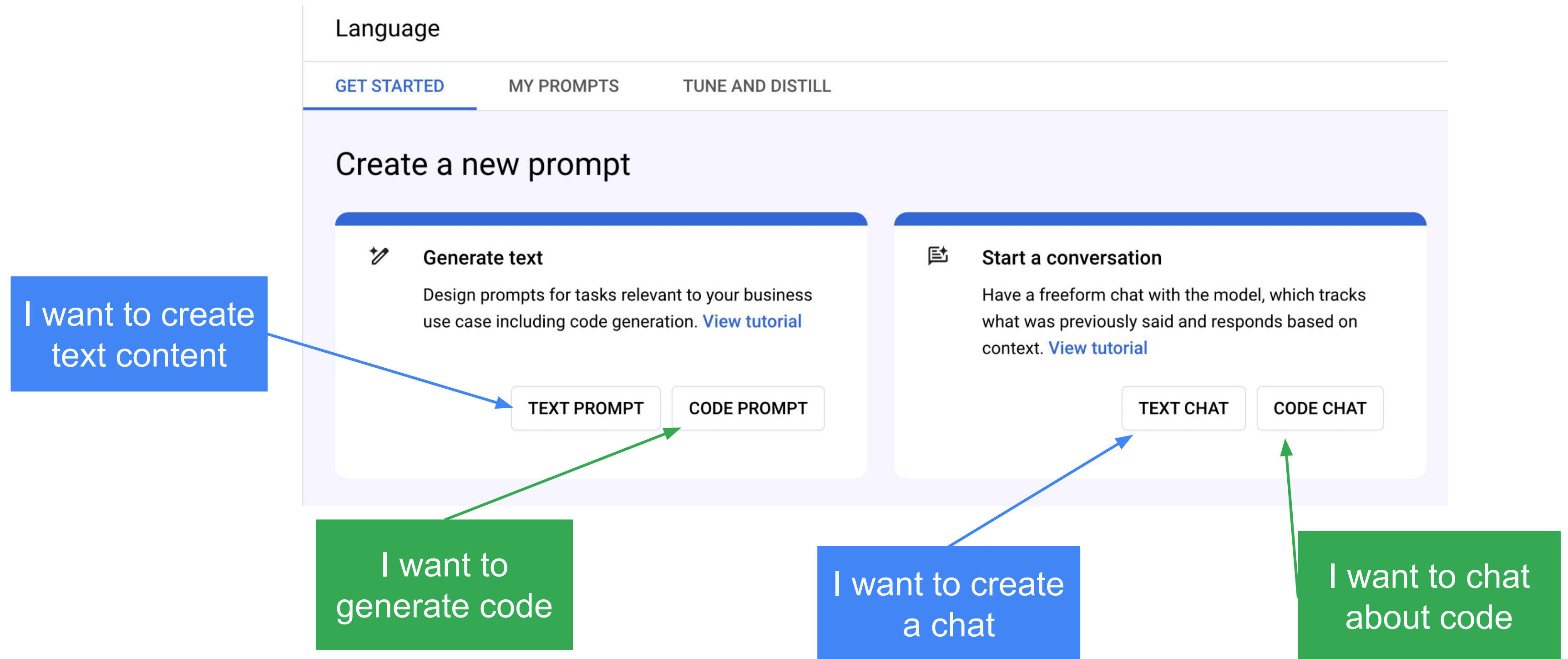


Vertex AI Studio is available as a feature of Google Cloud Vertex AI

- Choose from Language, Vision, Speech or Multimodal models



To get started, choose a task based on your goal



Vertex AI Studio UI

Save the model

Enter a Prompt

Select the foundational model

Adjust the parameters

The response

Submit the prompt

← grace-hopper-prompt ✎

FREEFORMSTRUCTURED

SAVEVIEW CODEDELETE

Prompt

Who was Grace Hopper?

Response

Grace Hopper, known as "Amazing Grace", was an American computer scientist and United States Navy rear admiral. She was a pioneer in the development of the COBOL programming language and was also one of the first programmers of the Harvard Mark I computer. She is widely considered to be one of the most important figures in the history of computing.

Model may display inaccurate or offensive information that doesn't represent Google's view. Not all languages are supported. [Learn more.](#)

Region

us-central1 (Iowa)

Model

text-bison (latest)

Temperature

0.2

Advanced

Token limit

1024

Top-K

40

Top-P

0.8

Max responses

1

Add stop sequence

Press Enter after each sequence

Streaming responses

Print responses as they're generated

Safety filter threshold

Block few

SUBMIT

RESET PARAMETERS

Vertex AI Studio UI for Chat

Save the model

Select the foundational model

Adjust the parameters

Submit the prompt

Enter a Prompt

The response

Add Examples

Add Context

hawaii-chat

Context (Optional)

Write your context here...

Examples (Optional)

USER

Write an input to add a new row

AI

Write an output to add a new row

Responses

USER

tell me about hawaii

AI

Hawaii, the 50th and most recent state of the United States, is an archipelago of volcanic islands in the central Pacific Ocean. Its eight main islands, from northwest to southeast, are Niihau, Kauai, Oahu, Molokai, Lanai, Kahoolawe, Maui, and the Island of Hawaii (also called the Big Island).

Sources

1.

<https://dtmag.com/thelibrary/a-sense-of-place-in-the-pacific-explori...>

Continue the conversation

Model may display inaccurate or offensive information that doesn't represent Google's view. Not all languages are

SAVE

VIEW CODE

DELETE

We want your feed

Region

us-central1 (Iowa)

Model

chat-bison (latest)

Temperature

0.2

Advanced

Token limit

1024

Top-K

40

Top-P

0.8

Max responses

1

Add stop sequence

Press Enter after each sequence

Safety filter

Block few

Google Cloud

You can use Vertex AI Studio to fine-tune foundational models

- You supply your own training data in JSON format
- The foundational model is customized for your particular use case
 - A new model is created based on PaLM 2
 - Uses a process called “Transfer Learning”
- The training job runs on Google infrastructure
- Your model and data are stored separately in your own project



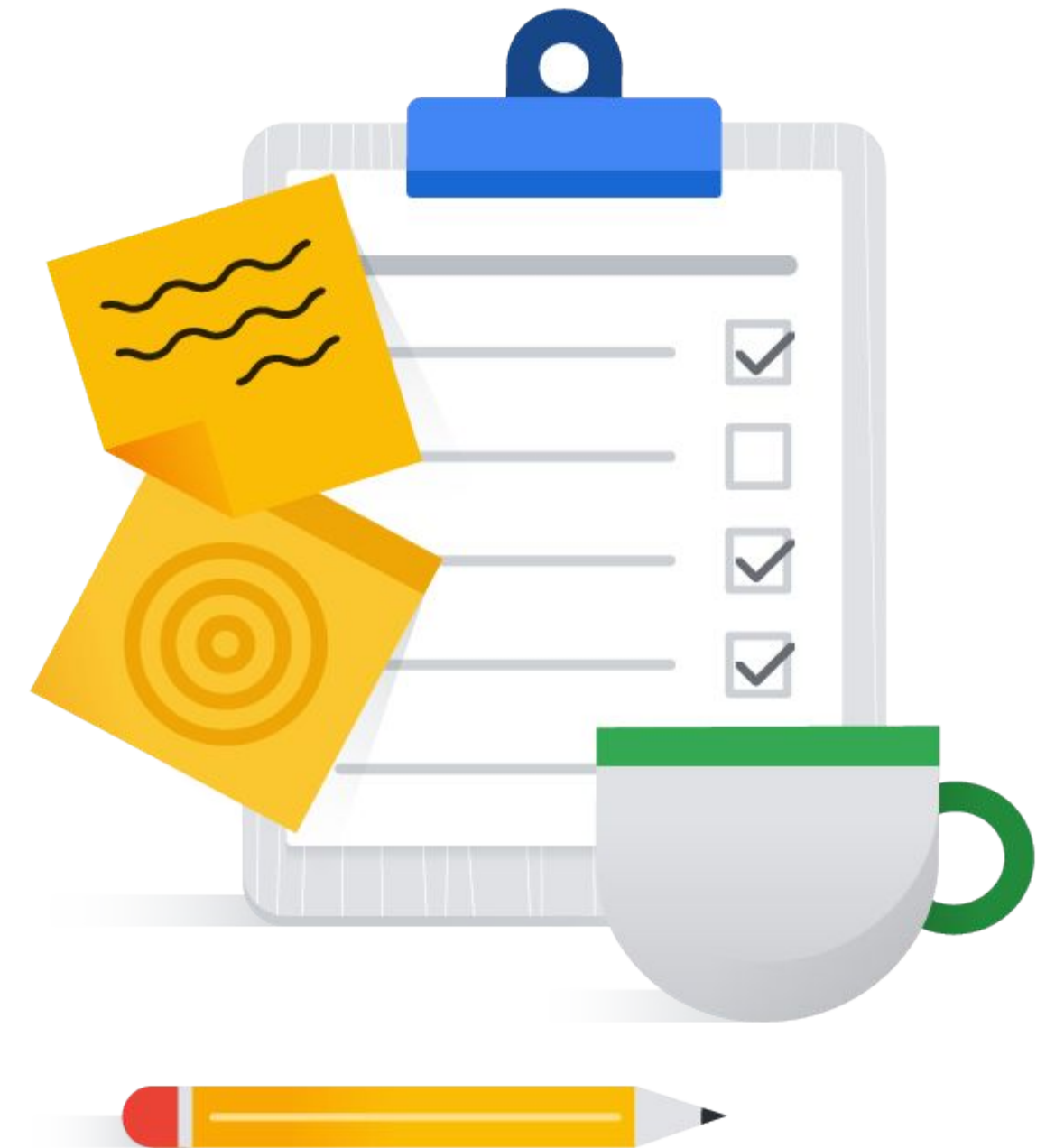
Tune a model

Tune a model so it's better equipped for your use case, then deploy to an endpoint to get predictions or test it in prompt design. Take a tutorial on [creating a tuned model](#).

[+ CREATE TUNED MODEL](#)

Topics

01	Vertex AI Studio
02	Designing and testing prompts
03	Data governance in Vertex Studio



Start with a simple text prompt and click the Submit button to see the results

Tell me about Hawaii





Responses

USER

Tell me about Hawaii

AI

Hawaii is a state of the United States located in the Pacific Ocean. It is the only state in the United States that is an archipelago, meaning that it is made up of a group of islands. Hawaii is the 50th state and the only one located outside of North America. The state is made up of eight main islands: Hawaii, Maui, Oahu, Kauai, Molokai, Lanai, Niihau, and Kahoolawe. Hawaii is also

How about a code prompt?

What is the Terraform code to create a Google Cloud Virtual Machine



e

AI

Here is the Terraform code to create a Google Cloud Virtual Machine:

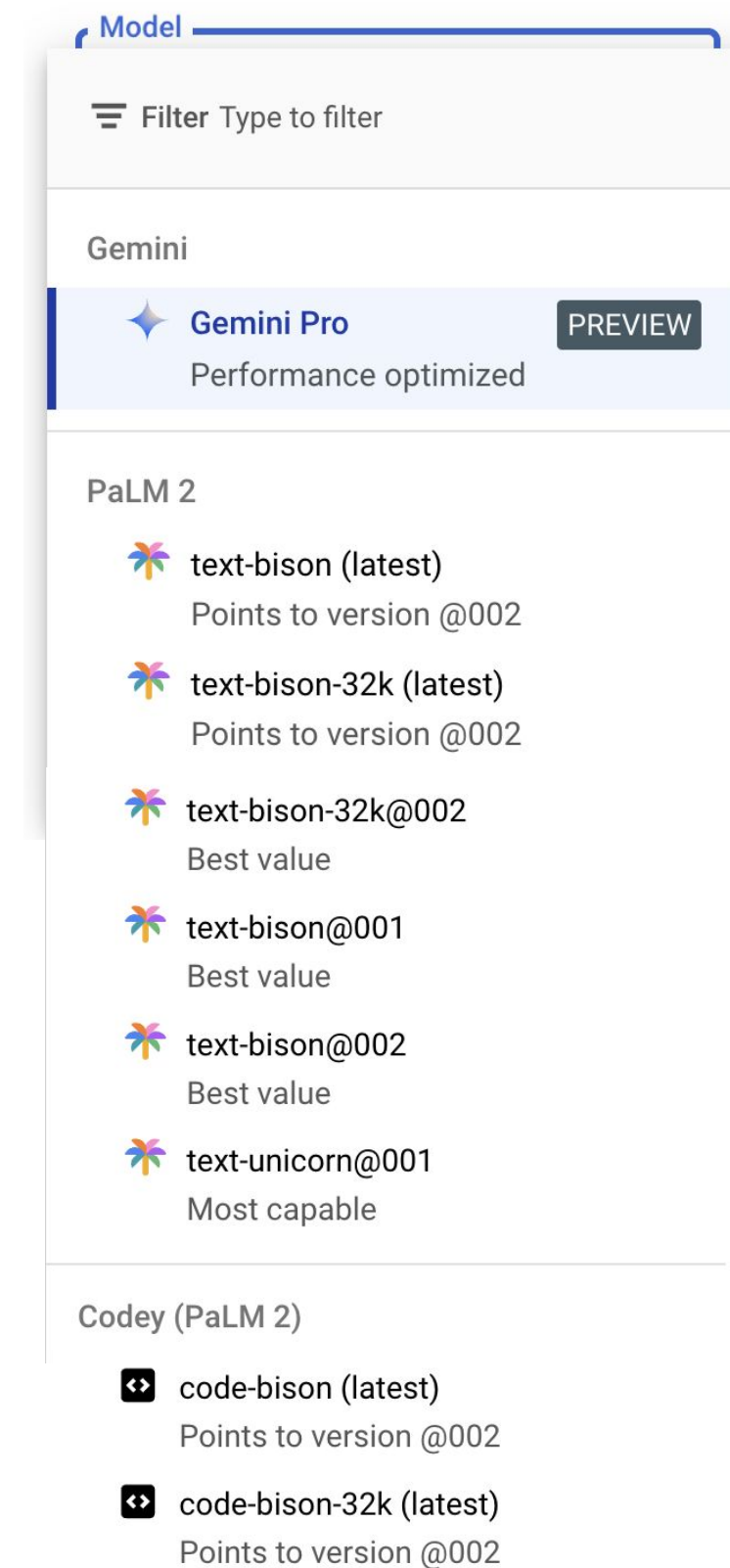
```
resource "google_compute_instance" "default" {
  name = "my-vm"
  machine_type = "n1-standard-1"
  zone = "us-central1-a"

  boot_disk {
    initialize_params {
      image = "debian-cloud/debian-9"
    }
  }

  network_interface {
    network = "default"
    subnetwork = "default"
  }
}
```

In Vertex AI Studio, choose the right model for text or code

- Use different model versions depending on the use case
 - The latest versions are more capable, but also more expensive
- PaLM 2 is for text
 - text-bison (latest) is the best, but most expensive
 - text-bison@### is not the latest, but cheaper
- Codey is for code
 - code-bison@### is the best
 - code-gecko@### is less expensive
- Gemini is for multimodal
- There are smaller models that will run entirely on mobile devices when building apps



Select a chat to have a conversation, use Codey to talk about code

- Chats (multi-turn conversations) remember the history of what was said in the session

USER Write me the code to create a Google Cloud Storage bucket

AI To create a Google Cloud Storage bucket, you can use the following code:

```
import google.cloud.storage as gcs

# Create a client
client = gcs.Client()

# Create a bucket
bucket = client.create_bucket('my_bucket')
```

This code will create a bucket named my_bucket in your Google Cloud Storage account. You can then use this bucket to store data.

USER How about using a shell script

AI Sure, here is the code for creating a Google Cloud Storage bucket using a shell script:

```
#!/bin/bash

# Create the bucket
gsutil mb gs://my_bucket
```

This code will create a bucket named my_bucket in your Google Cloud Storage account. You can then use this bucket to store data.

PaLM 2 is good at creating summaries

List some fun things to do in Hawaii



USER

List some fun things to do in Hawaii


AI

There are many fun things to do in Hawaii, including:

- **Visiting the beaches.** Hawaii is home to some of the most beautiful beaches in the world, including Waikiki Beach, Hanauma Bay, and Waimea Bay.
- **Snorkeling and scuba diving.** Hawaii's clear waters and abundant marine life make it a great place for snorkeling and scuba diving.
- **Hiking.** Hawaii has many hiking trails, ranging from

How about sentiment analysis...

Responses



USER

Is the following comment about Hawaii positive or negative: We had a great time in Hawaii. The food was great. Though the flight was long.

AI

The comment is positive. The person said they had a great time in Hawaii, and that the food was great. The only negative thing they mentioned was that the flight was long. However, they still had a great time overall.

Ideation...

USER

I want to write a blog post about vacationing in Hawaii. What are some good titles?

AI

Here are some good titles for a blog post about vacationing in Hawaii:

- The Best Things to Do in Hawaii
- The Best Beaches in Hawaii
- The Best Hikes in Hawaii
- The Best Restaurants in Hawaii
- The Best Things to See in Hawaii
- The Best Places to Stay in Hawaii
- How to Plan a Trip to Hawaii

Classification...

USER

Is the following message
Spam or Not Spam: 2 Days
and Nights FREE in Hawaii.
Just listen to our 2-hour
exclusive member sharing
presentation.

AI

This message is spam. It is
trying to get you to listen to
a presentation in order to
get a free trip to Hawaii.
This is a common tactic
used by scammers to get
your personal information.

Content creation...

USER

Can you write me a 1 page blog post on the history of Hawaii

AI

Sure, here is a 1 page blog post on the history of Hawaii:

The history of Hawaii begins with the arrival of the first Polynesians around 1,000 years ago. These early settlers were from the Marquesas Islands, and they brought with them their language, culture, and religion. Over time, the Hawaiian people developed their own unique

Extraction...

USER

Can you extract and list the key dates from that history

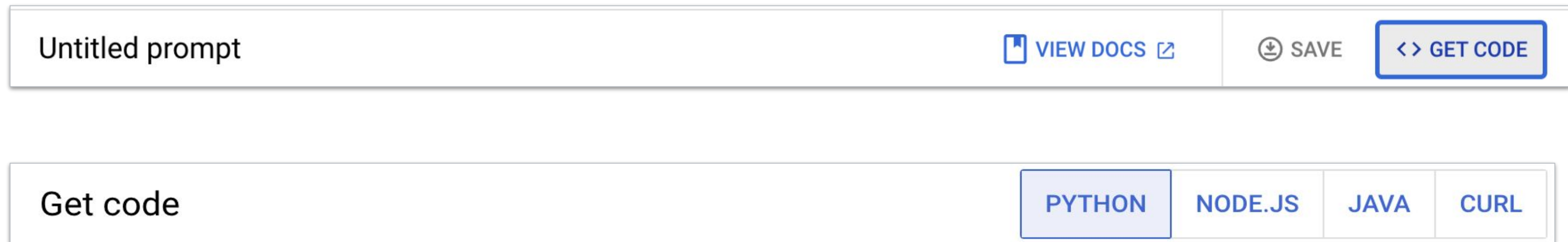
AI

Sure, here are some of the key dates in the history of Hawaii:

- 1,000 years ago: The first Polynesians arrive in Hawaii.
- 1778: British explorer James Cook becomes the first European to visit Hawaii.
- 1893: A group of American businessmen and sugar planters overthrow the Hawaiian monarchy.
- 1898: The United States annexes Hawaii.
- 1959: Hawaii becomes a state of the United States.

Ok, but how do I use this in my app?

- In Vertex AI Studio, click the **View Code** button
 - Returns the code in Python, Node.js, Java or CURL



Python code example

View code

PYTHON

Use this script to request a model response in your application.

1. Set up the [Vertex AI SDK for Python](#)
2. Use the following code in your application to request a model response

```
import vertexai
from vertexai.language_models import TextGenerationModel

vertexai.init(project="vertex-ai-dar", location="us-central1")
parameters = {
    "temperature": 0.2,
    "max_output_tokens": 256,
    "top_p": 0.8,
    "top_k": 40
}
model = TextGenerationModel.from_pretrained("text-bison@001")
response = model.predict(
    """Is the following Spam or Not Spam:

Hi Doug, I read this article from Wikipedia on Grace Hopper. Thought you might
    **parameters
)
print(f"Response from Model: {response.text}")
```

cURL code example

View code

PYTHON PYTHON

Use the command-line interface (CLI) to request a model response

1. Install Google Cloud SDK [Google Cloud SDK](#) if you haven't already or open Cloud Shell and skip to step 3.
2. Run the following command to authenticate using your Google account.

```
$ gcloud auth application-default login
```

3. Enter the following to request a model response

```
API_ENDPOINT="us-central1-aiplatform.googleapis.com"
PROJECT_ID="vertex-ai-dar"
MODEL_ID="text-bison@001"

curl \
-X POST \
-H "Authorization: Bearer $(gcloud auth print-access-token)" \
-H "Content-Type: application/json" \
"https://${API_ENDPOINT}/v1/projects/${PROJECT_ID}/locations/us-central1/publishers/google/models/${MODEL_ID}:predict" \
$' {
  "instances": [
    {
      "content": "Is the following Spam or Not Spam:

Hi Doug, I read this article from Wikipedia on Grace Hopper. Thought you might be interested."
    }
  ],
  "parameters": {
    "temperature": 0.2,
    "maxOutputTokens": 256,
    "topP": 0.8,
    "topK": 40
  }
}
```

But I want a different language... Ask Codey!

▼ Prompt

Can you convert the following Python code into Go?

```
import vertexai
from vertexai.language_models import TextGenerationModel

vertexai.init(project="vertext-ai-dar", location="us-central1")
parameters = {
    "temperature": 0.2,
    "max_output_tokens": 256,
    "top_p": 0.8
```


▼ Response Markdown

```
package main

import (
    "fmt"

    "google.golang.org/api/vertexai/v1alpha1"
)

func main() {
    // Initialize the client.
    client, err := vertexai.NewClient(vertexai.ClientConfig{
        ProjectID: "vertext-ai-dar",
        Location:  "us-central1",
    })
    if err != nil {
        panic(err)
    }
}
```



You can explore sample use cases and prompts from Vertex AI Vertex AI Studio

Prompt examples

Try a sample prompt to test a foundation model with a practical use case

Just getting started? See [Introduction to prompt design](#) to learn more.

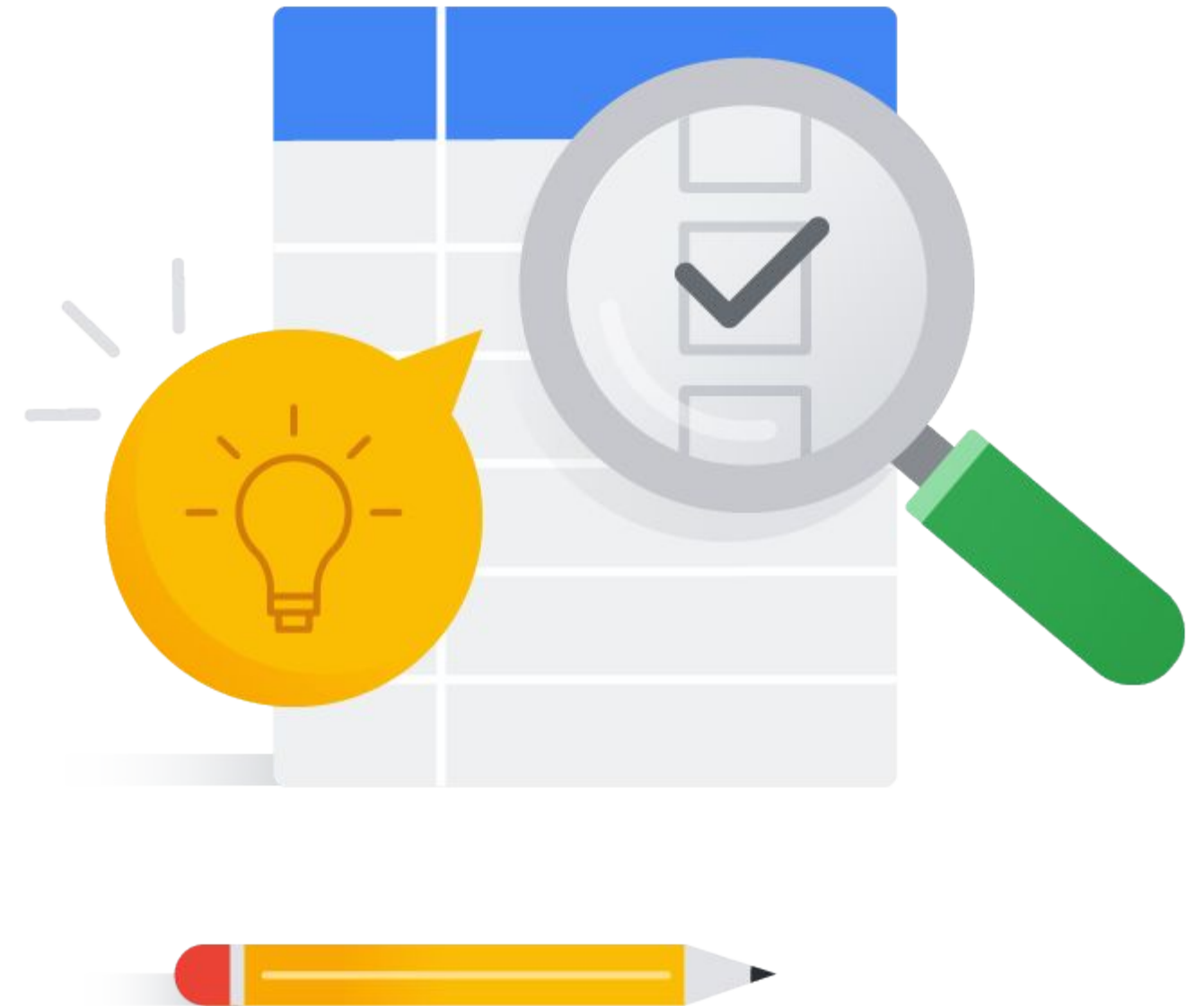
Summarization

<div>Freeform</div> <div>Financial table summariz...</div> <div>Summarize key insights from a financial report table with heavy numerical info.</div> <div>OPEN</div>	<div>Freeform</div> <div>Support chat summary</div> <div>Summarize a customer support call.</div> <div>OPEN</div>	<div>Freeform</div> <div>Support chat summary - ...</div> <div>Summarize a customer support chat in French.</div> <div>OPEN</div>	<div>Freeform</div> <div>Article summary</div> <div>Summarize a news article.</div> <div>OPEN</div>	<div>Freeform</div> <div>Support call summary</div> <div>Summarize a customer support call from the agent perspective.</div> <div>OPEN</div>
<div>Freeform</div> <div>Support call next steps</div> <div>Summarize actions taken during a custom support call and agent next steps.</div> <div>OPEN</div>	<div>Structured</div> <div>Transcript summary</div> <div>Summarize a transcript using examples.</div> <div>OPEN</div>	<div>Structured</div> <div>Dialog summary</div> <div>Summarize a conversation using examples.</div> <div>OPEN</div>	<div>Structured</div> <div>Hashtag tokenization</div> <div>Create and tokenize hashtags.</div> <div>OPEN</div>	<div>Structured</div> <div>Title generation</div> <div>Create a title for an article using examples.</div> <div>OPEN</div>

Lab

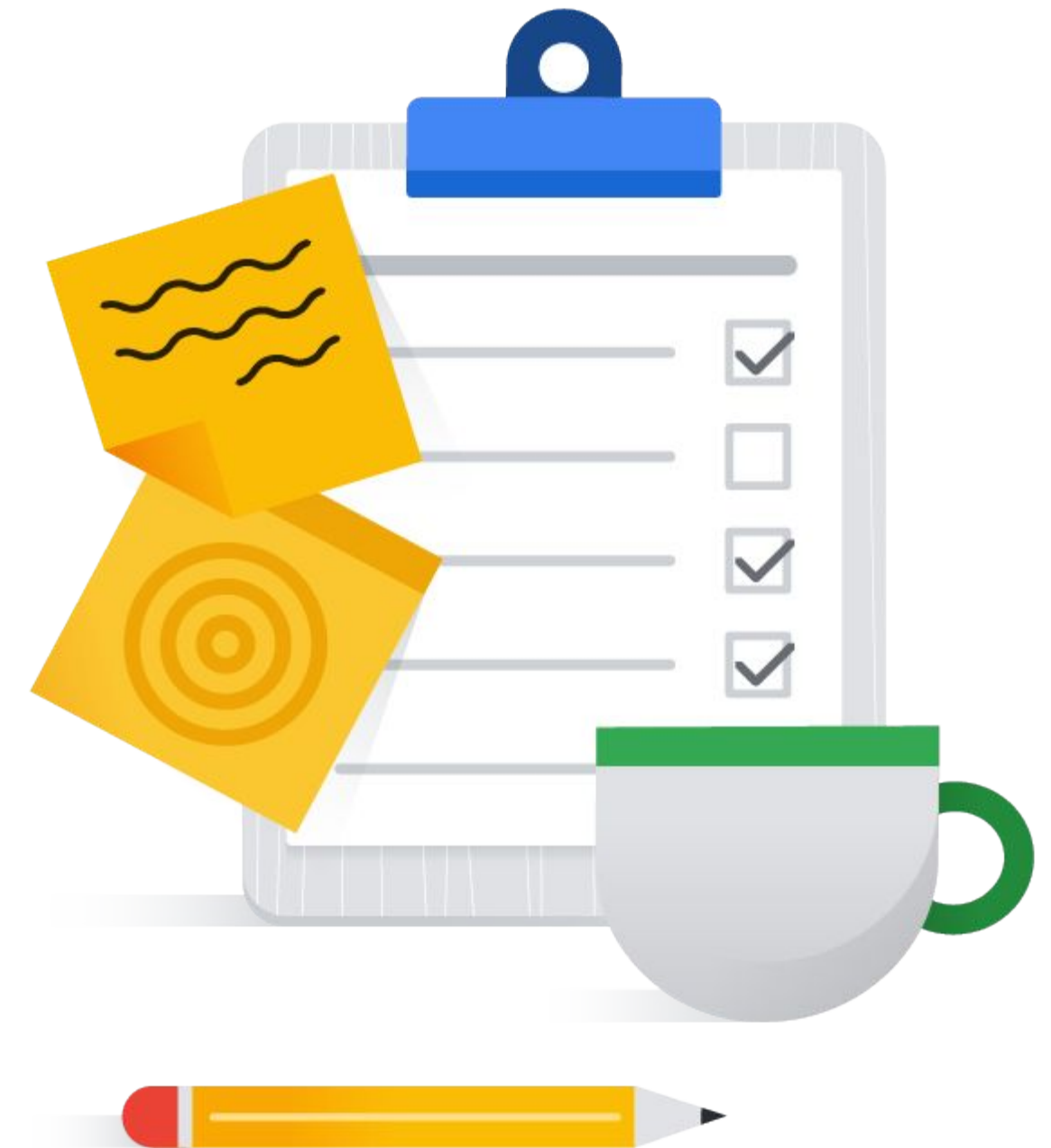
🕒 30 min ⚙️

Getting started with Vertex AI Vertex AI Studio's User Interface



Topics

01	Vertex AI Studio
02	Designing and testing prompts
03	Data governance in Vertex AI Studio



AI Data Governance

Google Cloud's approach to governance of customer data for Cloud AI Large Models and Generative AI



We have no access to customer data, in accordance with [GCP Terms](#) and [Cloud Data Processing Addendum](#).



We're committed to transparency, compliance with regulations like the EU General Data Protection Regulation (GDPR), and privacy best practices.



We will not use customer data to train our models



Security Generative AI Risk Pillars

01

Privacy & Safety

Transparency on how Google LLMs operate is core to the Google's mission for Responsible AI

02

Prompt Hacking

Large Language Models can be manipulated to output responses that are not aligned with its objectives

03

Sec Operations

Generative AI can be leveraged to monitor for malicious activities

04

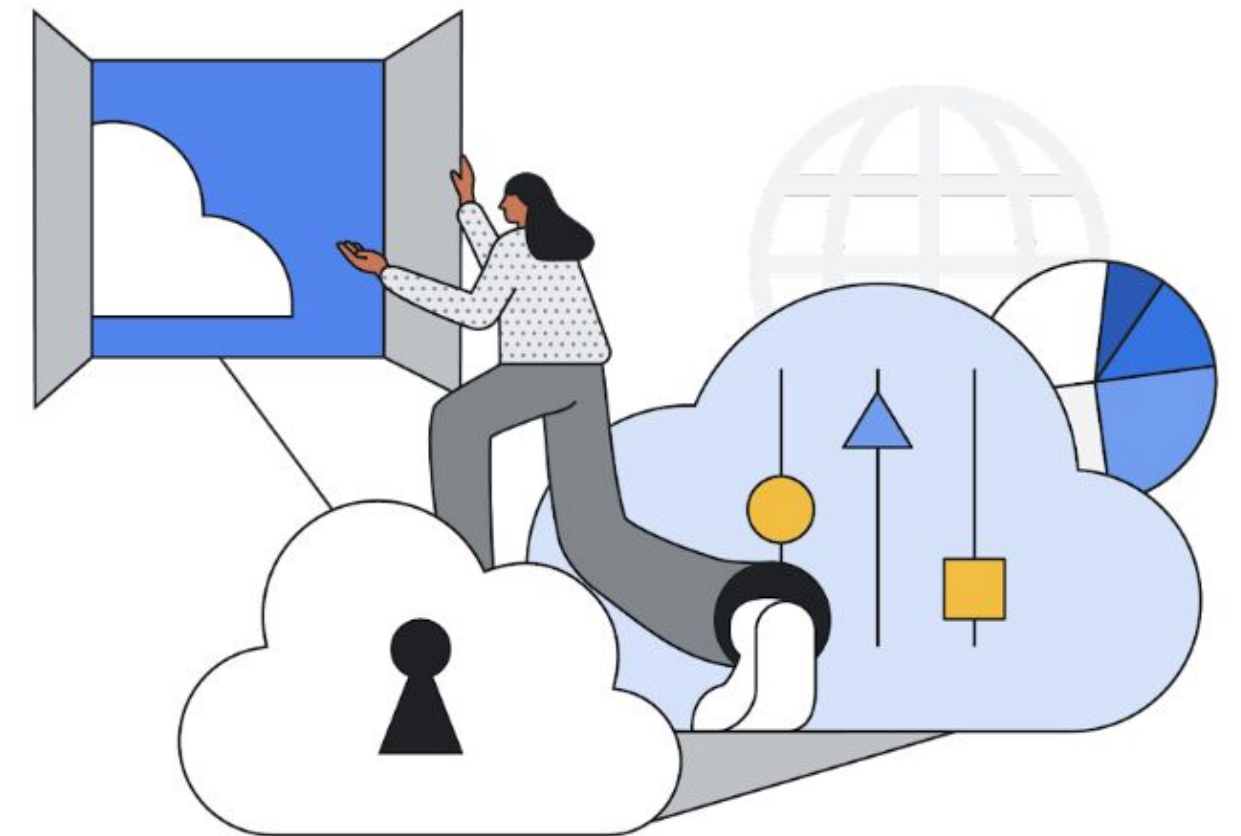
Cloud Resources

Gen AI workloads require the same security controls as traditional workloads



Customer data is kept private during prompt design

- Customer data is encrypted in-transit when a request is submitted to a foundational model
 - Uses industry-standard cryptographic protocols to protect the confidentiality of customer data
- The foundation model processes customer data to generate a response
 - Does not retain the data for any other purpose
- Google does not use customer data to train its models without the express consent of its customers



Model tuning data is not retained by Google

Fine tuning does not rebuild foundational models

- Creates additional parameters, called “adapter weights”
- Passed to foundational model during inference

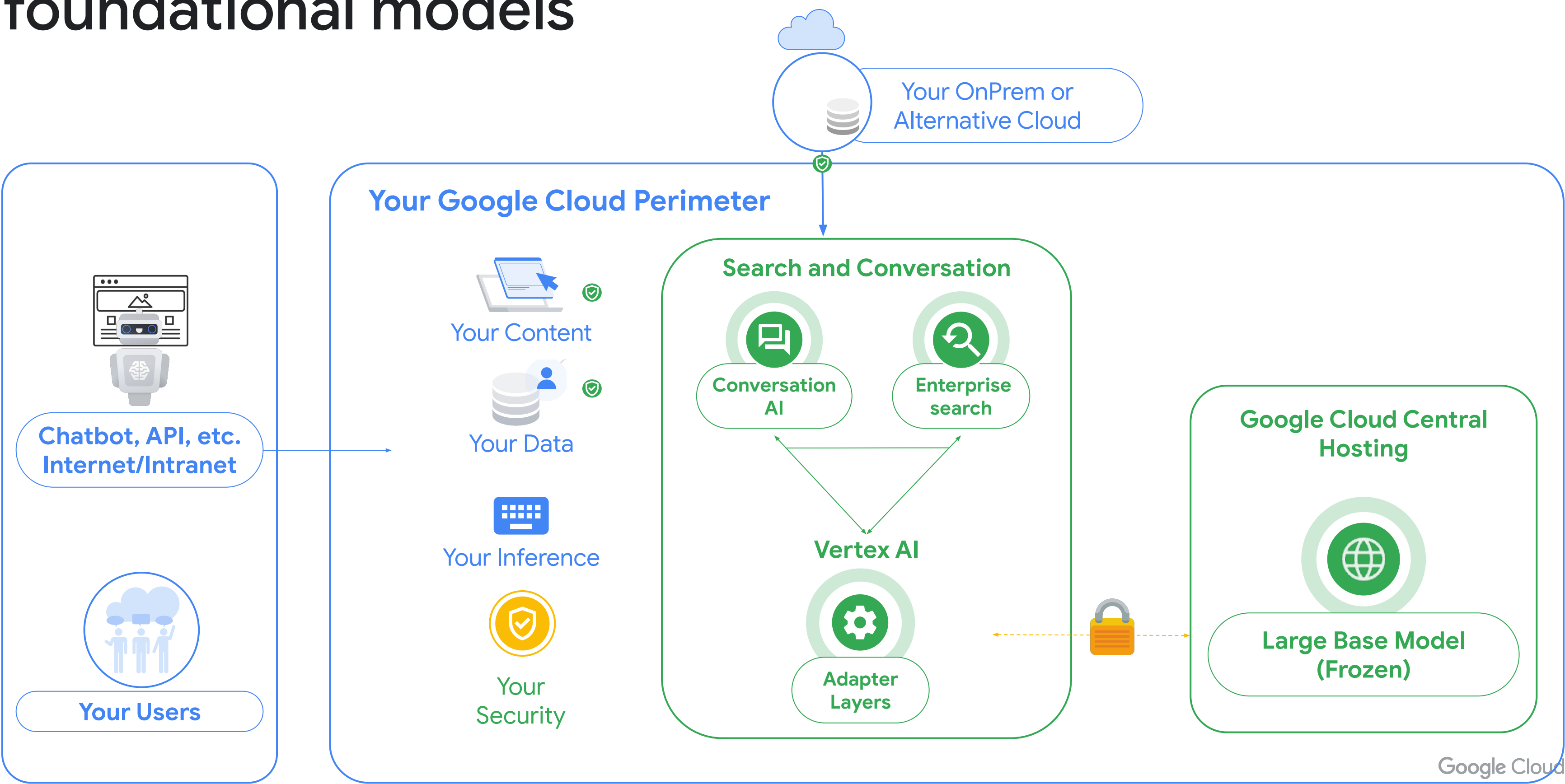
Customers have sole access to tuned models

- Can control encryption keys
- Can delete the model

Customer data is secure

- All data is encrypted at rest and in-transit
- Prompts and training data aren’t logged

Your fine-tuned models are separated from the foundational models

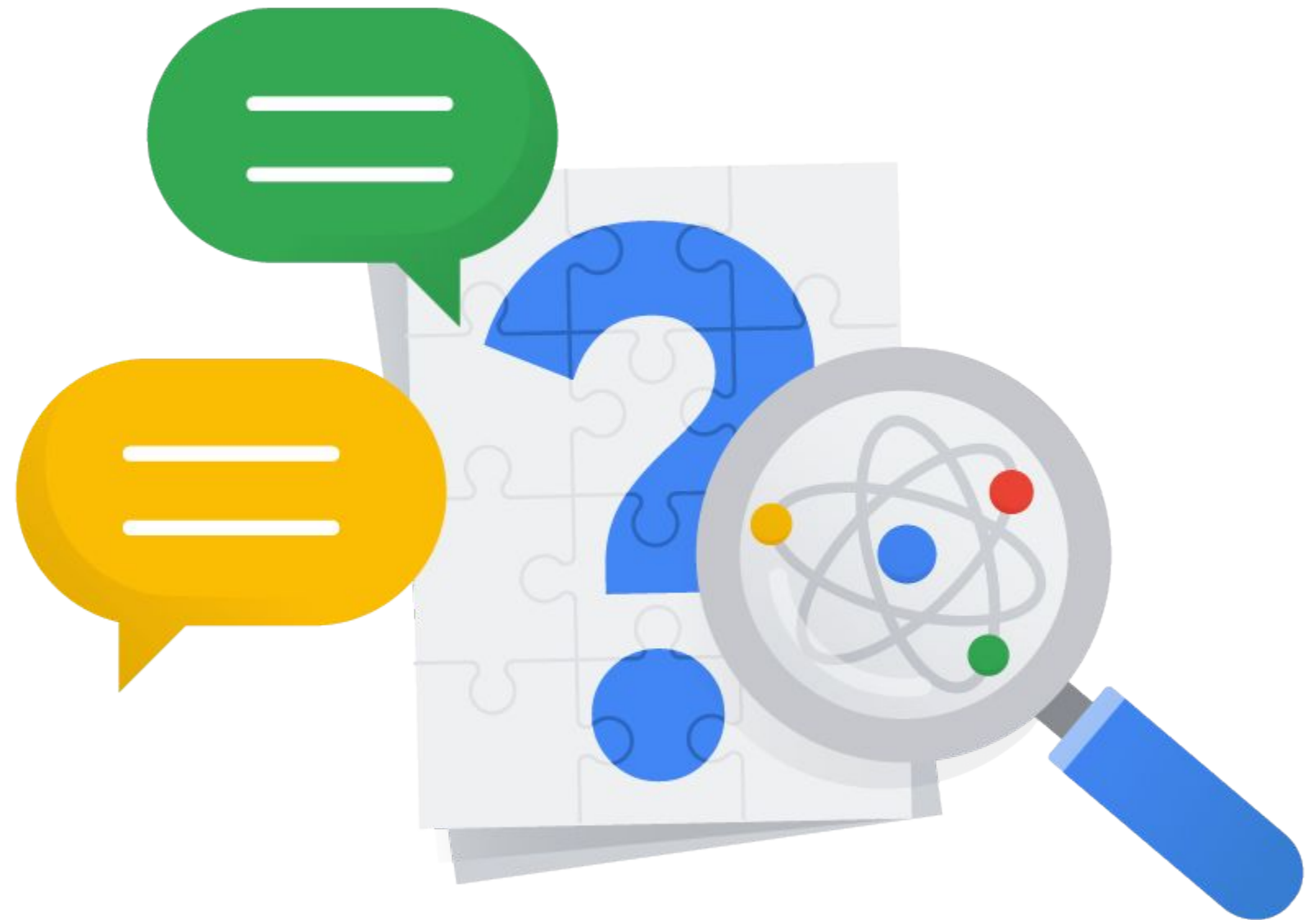


In this module, you learned to ...

- 01 Use Vertex AI Studio to implement generative AI use cases
- 02 Explore the available models and sample use cases in Vertex AI Studio
- 03 Write and test prompts using the Google Cloud Console
- 04 Protect your data while using Vertex AI Studio



Questions and answers



Quiz question

Match the model to the task

PaLM 2

Codey

Chirp

Imagen

Gemini

Generate text from speech

Generate code

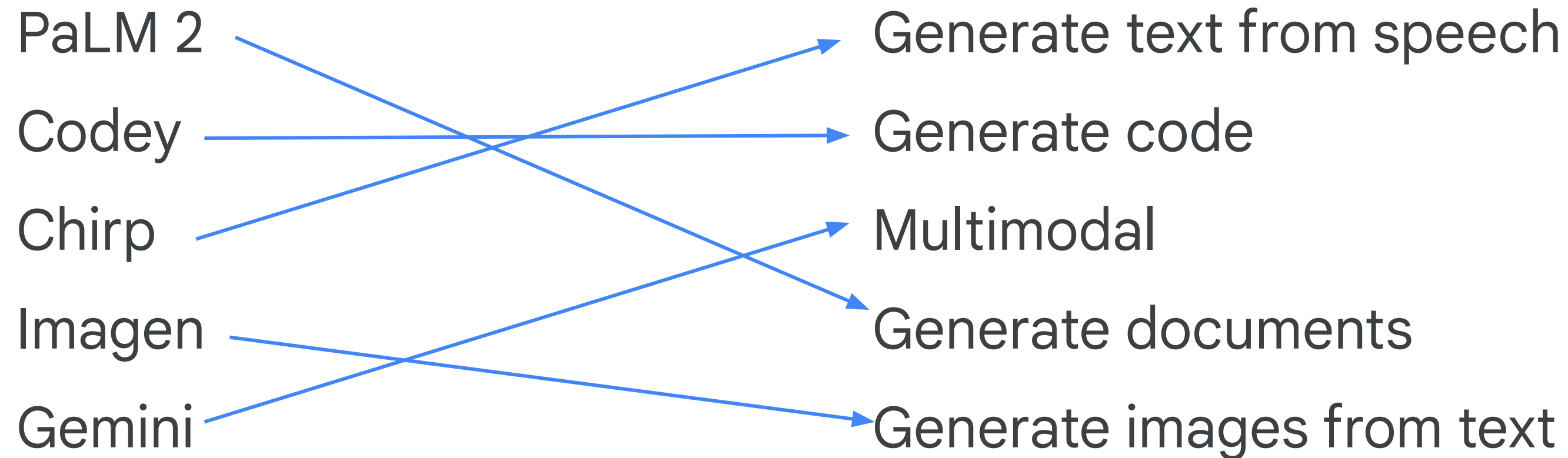
Multimodal

Generate documents

Generate images from text

Quiz question

Match the model to the task



Quiz question

Vertex AI Studio will generate code in which languages? (choose all that apply)

A: Python

B: IPython (Jupyter)

C: Go

D: Node.js

E: cURL

F: Java

Quiz question

Vertex AI Studio will generate code in which languages? (choose all that apply)

A: Python

B: IPython (Jupyter)

C: Go

D: Node.js

E: cURL

F: Java

Quiz question

List some use cases for PaLM 2 language models:

Quiz question

List some use cases for PaLM 2 language models:

Writing

Sentiment analysis

Classification

Extraction

Summarization

Ideation

Chat

