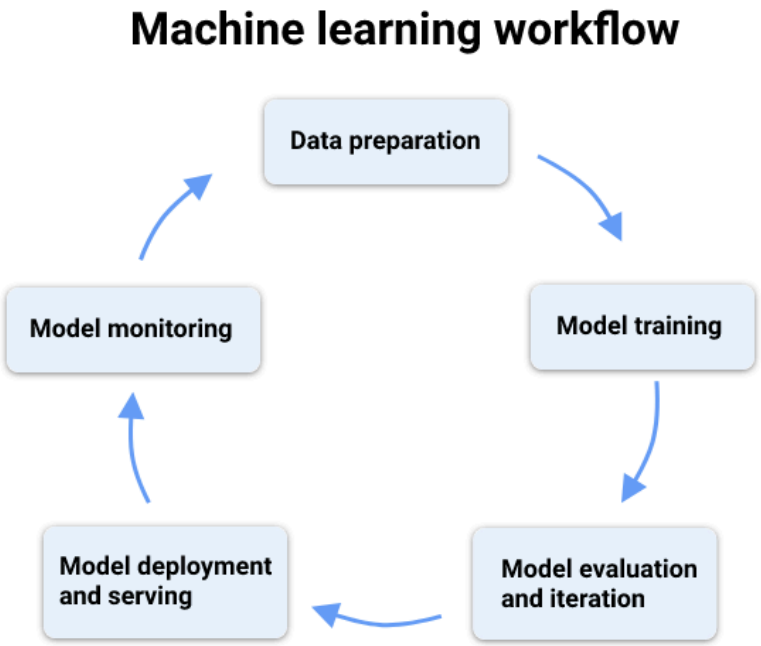
Getting Started with Agent Builder - PDF Q&A

For section-specific guidelines and examples, see below

Vertex AI Overview

Vertex AI is a machine learning (ML) platform that lets you train and deploy ML models and AI applications. Vertex AI combines data engineering, data science, and ML engineering workflows, enabling your teams to collaborate using a common toolset



What are LLMs

Large language models (LLMs) are deep learning models trained on massive datasets of text. LLMs can translate language, summarize text, generate creative writing, generate code, power chatbots and virtual assistants, and complement search engines and recommendation systems. Creating an LLM requires massive amounts of data, significant compute resources, and specialized skills. Because LLMs require a big investment to create, they target broad rather than specific use cases. On Vertex AI, you can customize a foundation model for more specific tasks or knowledge domains by using prompt design and model tuning.

Build powerful AI agents, no code required. For complex goals, you can easily stitch together multiple agents, with one agent functioning as the main agent and others as subagents. Train with your data, automate tasks, and iterate with ease. Launch and analyze - all within a user-friendly platform.

Objectives:

- ✅ Introduce Agent Builder
- ✅ Introduce DataStore
- ✅ Build a use-case with a no-code/low-code platform

Pre-Requirements

Before you click the Start Lab button

Read these instructions. Labs are timed and you cannot pause them. The timer, which starts when you click **Start Lab**, shows how long Google Cloud resources will be made available to you.

This hands-on lab lets you do the lab activities yourself in a real cloud environment, not in a simulation or demo environment. It does so by giving you new, temporary credentials that you use to sign in and access Google Cloud for the duration of the lab.

What you need

To complete this lab, you need:

- Access to a standard internet browser (Chrome browser recommended).
- Time to complete the lab.

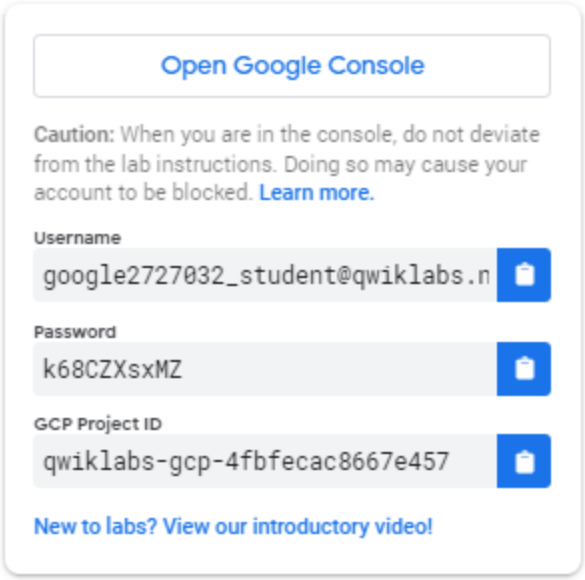
Note: If you already have your own personal Google Cloud account or project, do not use it for this lab.

Note: If you are using a Chrome OS device, open an Incognito window or use a guest profile to run this lab.

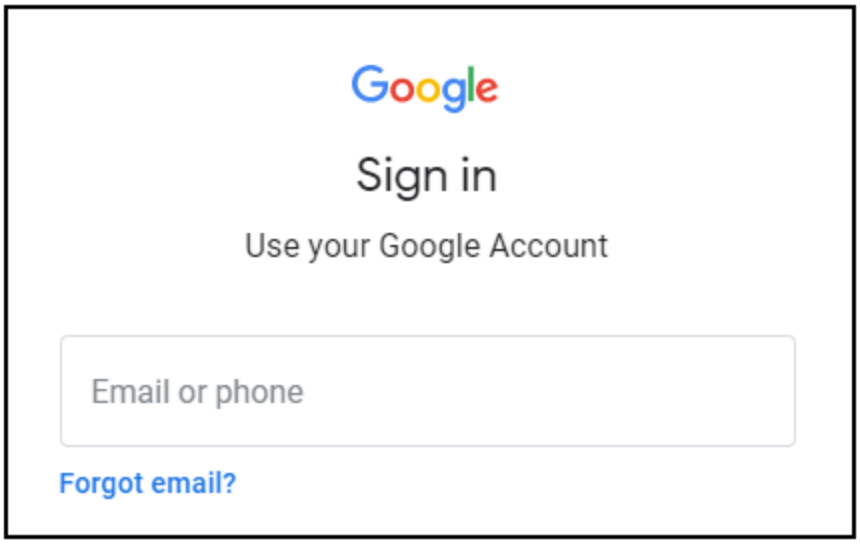
How to start your lab and sign in to the Google Cloud Console

1. Click the **Start Lab** button. If you need to pay for the lab, a pop-up opens for you to select your payment method. On the left is a panel populated with the temporary credentials that you must use for this lab.

2. Copy the username, and then click **Open Google Console**. The lab spins up

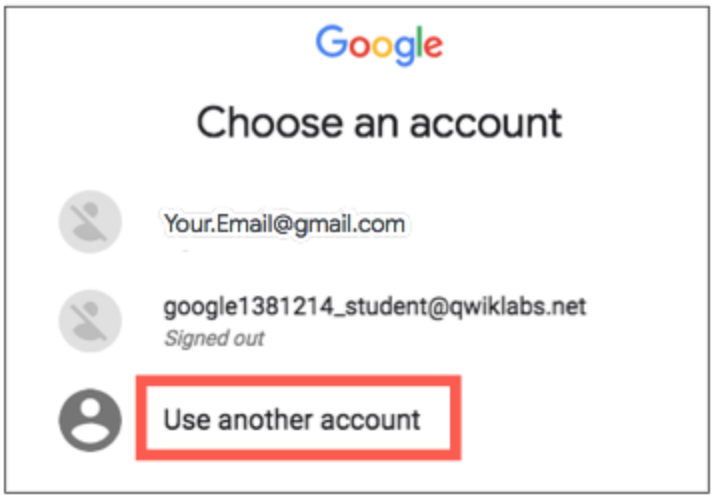


resources, and then opens another tab that shows the **Sign in** page.



Tip: Open the tabs in separate windows, side-by-side.

If you see the **Choose an account** page, click **Use Another Account**.

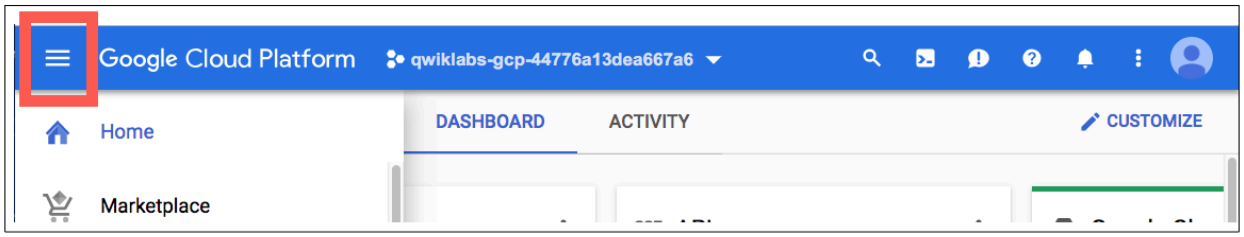


3. In the **Sign in** page, paste the username that you copied from the left panel. Then copy and paste the password.

   Important: You must use the credentials from the left panel. Do not use your Google Cloud Training credentials. If you have your own Google Cloud account, do not use it for this lab (avoids incurring charges).
4. Click through the subsequent pages:
   - Accept the terms and conditions.
   - Do not add recovery options or two-factor authentication (because this is a temporary account).
   - Do not sign up for free trials.

After a few moments, the Cloud Console opens in this tab.

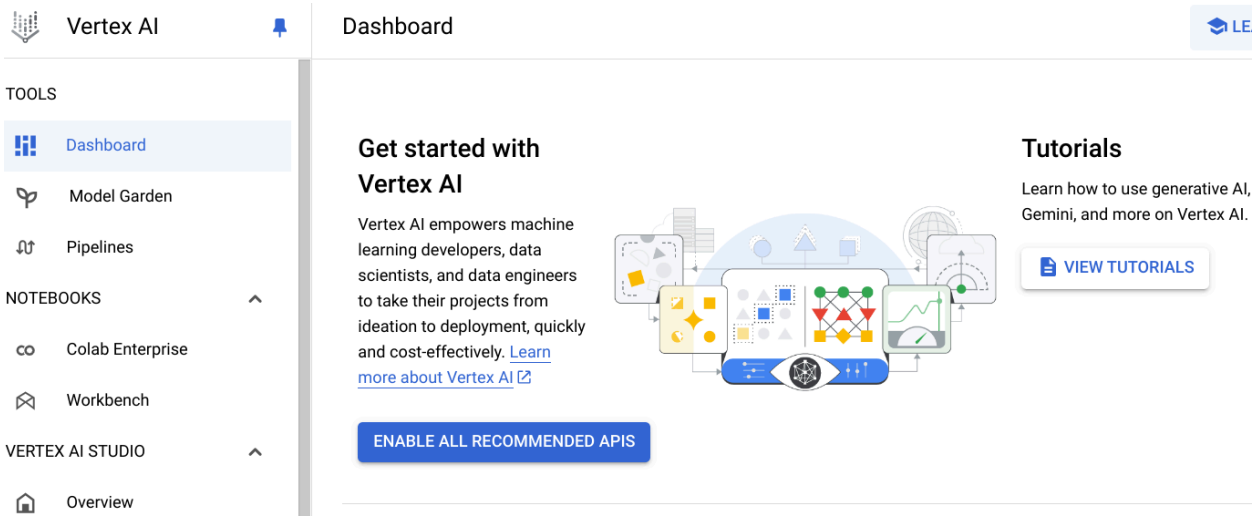
Note: You can view the menu with a list of Google Cloud Products and Services by clicking the **Navigation menu** at the top-left.



Set up your environment

Enable the Vertex AI API

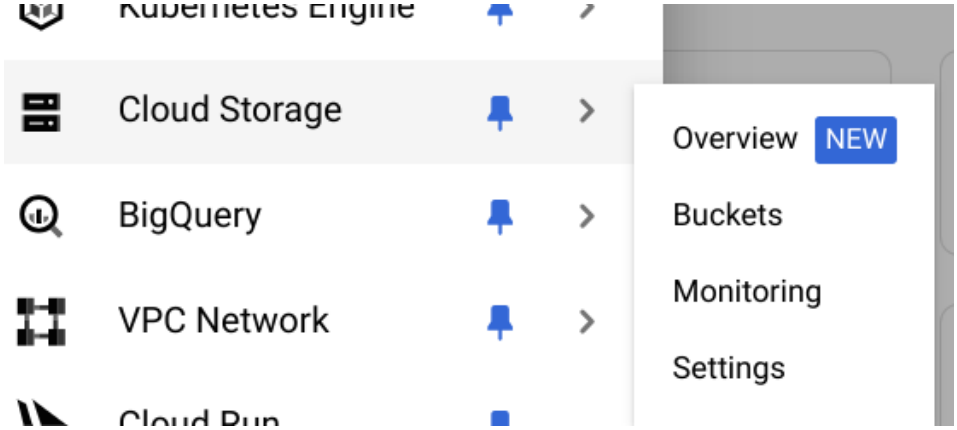
1. In the Google Cloud Console, on the **Navigation menu**, click **Vertex AI** or Utilize the search bar to find **Vertex AI.**
2. Within the "Get Started with Vertex AI" section, locate and click on the option that says "ENABLE ALL RECOMMENDED API."



Q&A over PDF documents

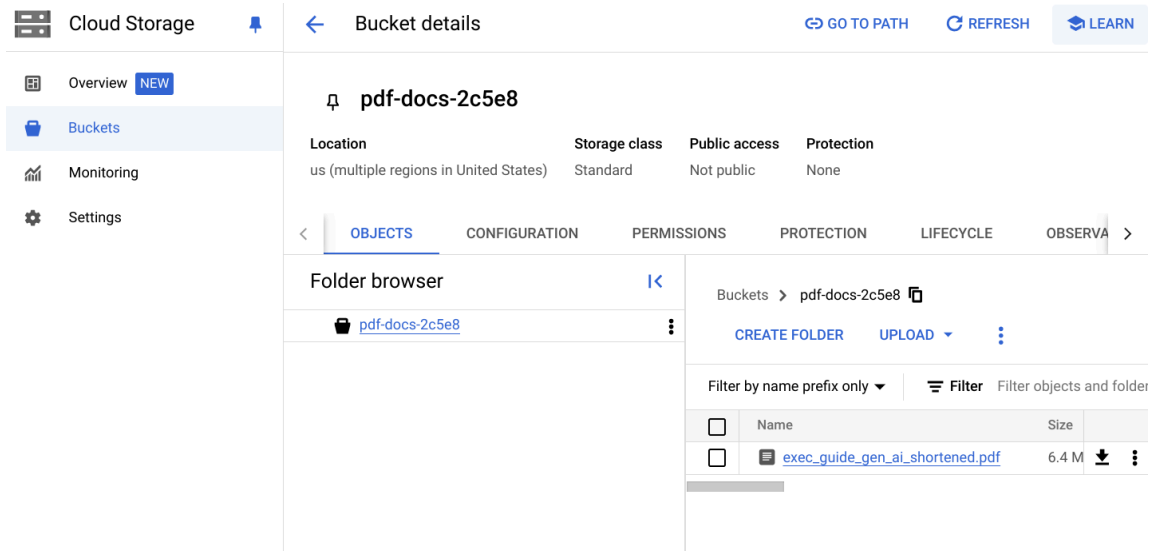
Task 1: Create Cloud Storage Bucket

1. Navigate to Cloud Storage



2. Create a new Bucket
    - For bucket name type: "`pdf-docs`" + last 5 digits of your GCP project.
    - Location type: `multi-region, us`.
    - Storage class: `Standard`
    - Access control: `Uniform`
3. Data protection: `uncheck soft delete policy`
4. Click "`Create`".
5. Confirm "Public access will be prevented".
6. Download PDF report and upload it to the bucket. https://storage.googleapis.com/ancient-sandbox-322523-external-bucket/exec_guide_gen_ai_shortened.pdf

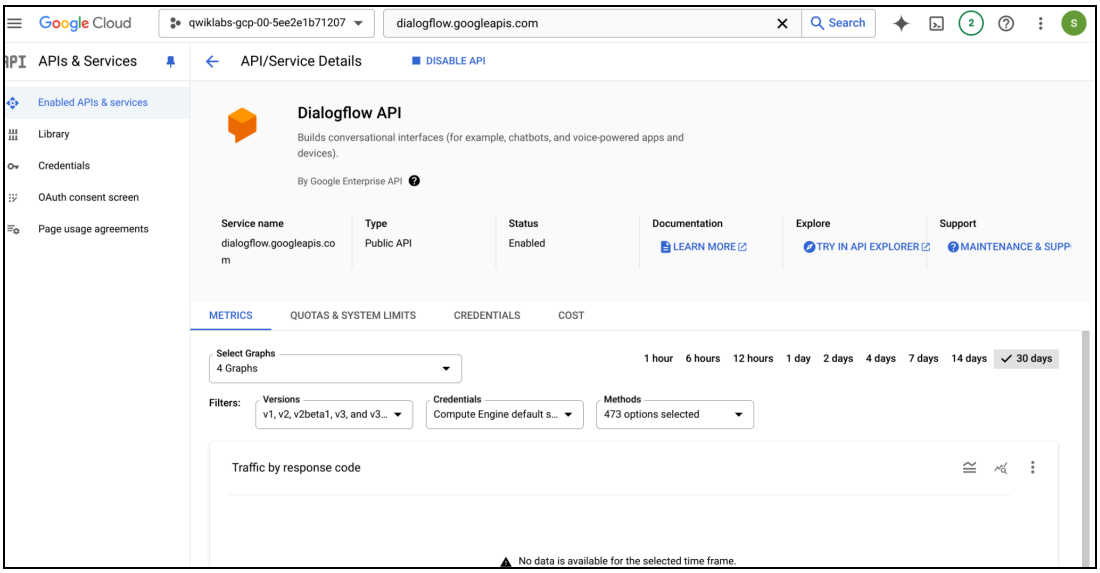
Bucket with uploaded file view:



Now, let us enable the DialogFlow APIs.

On the top search bar, search for *dialogflow.googleapis.com* api.

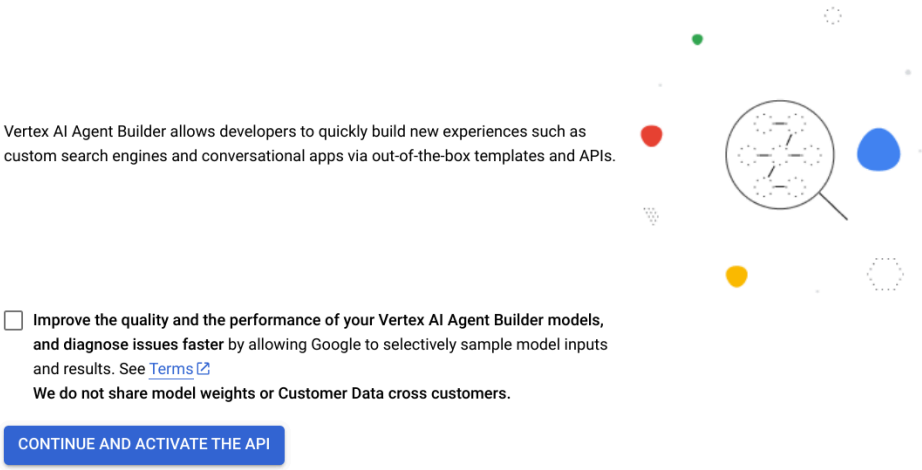
Enable the API, if not already enabled.



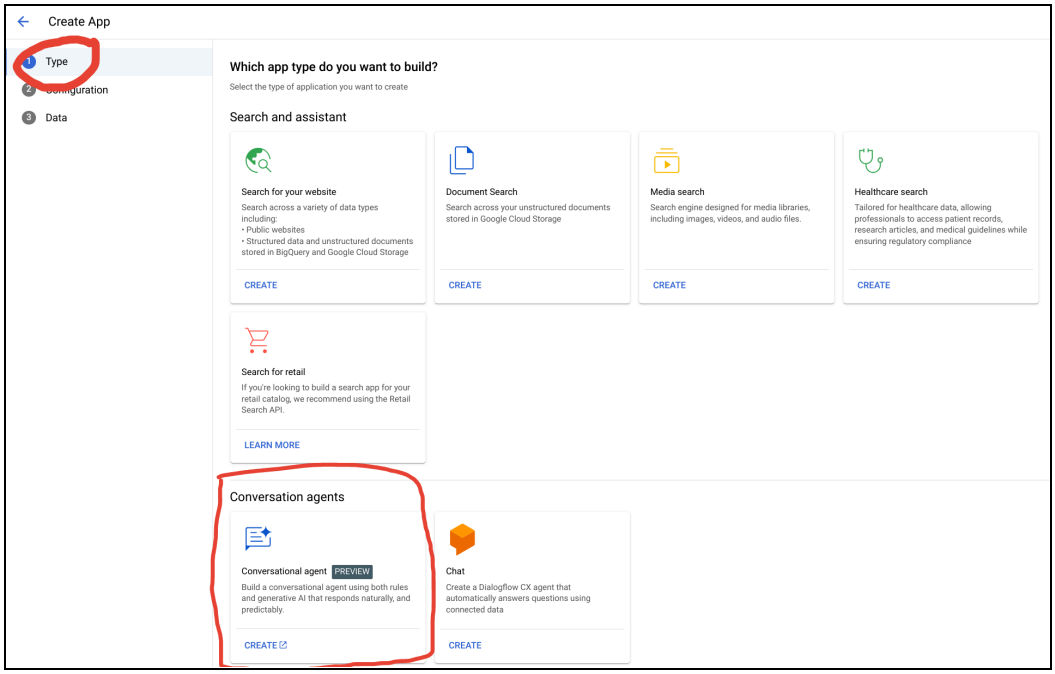
Task 2: Data store configuration

Search or Navigate to "`Agent Builder`"

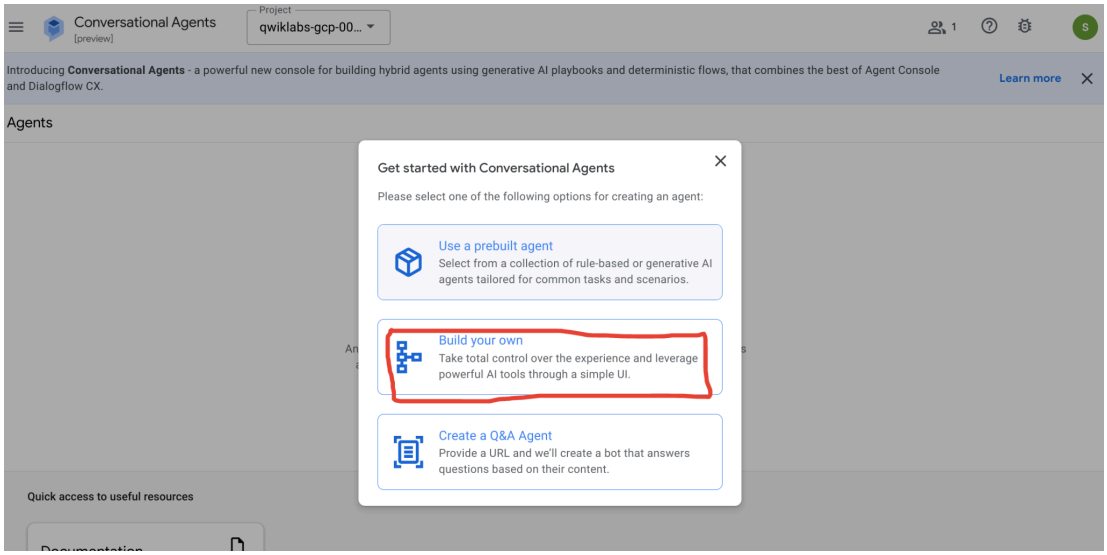
1. Click Continue and Activate the API



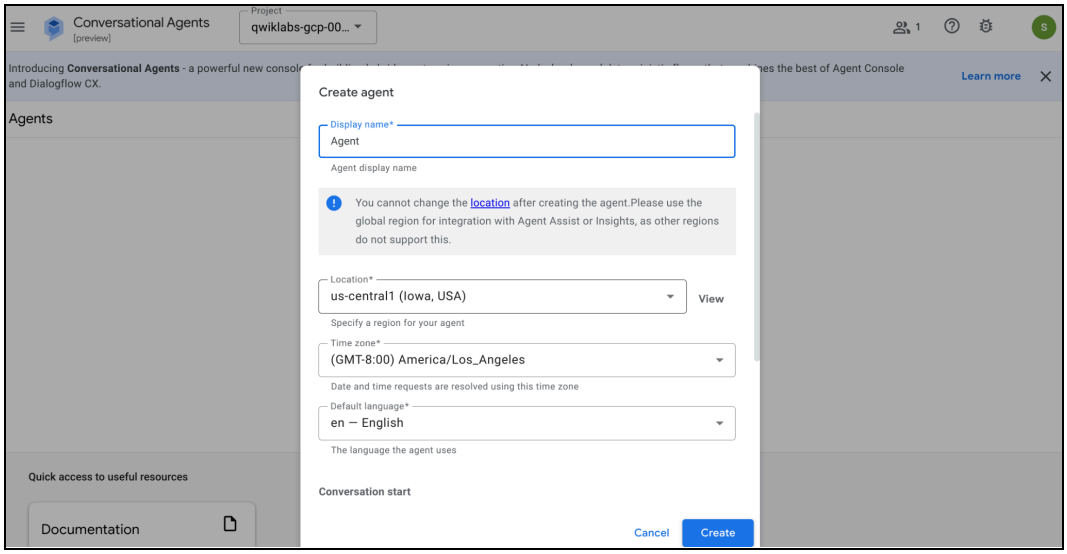
2. In Create App, select Conversation agents



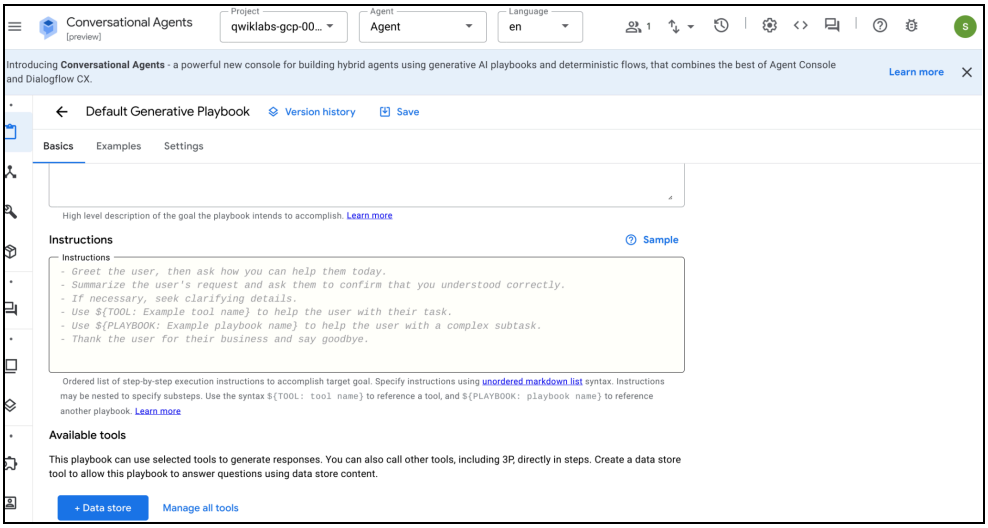
Select “Build your own”



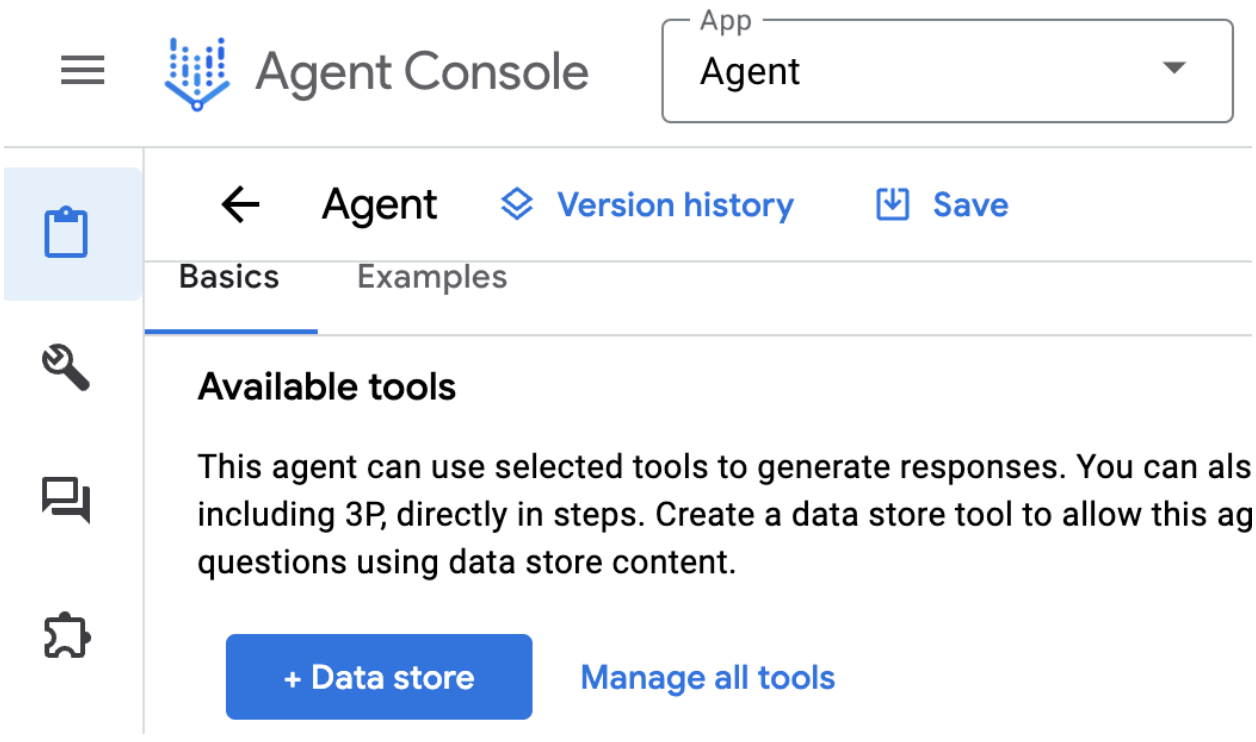
Click “Create”. Please note that it is a Playbook - Generative AI agent.



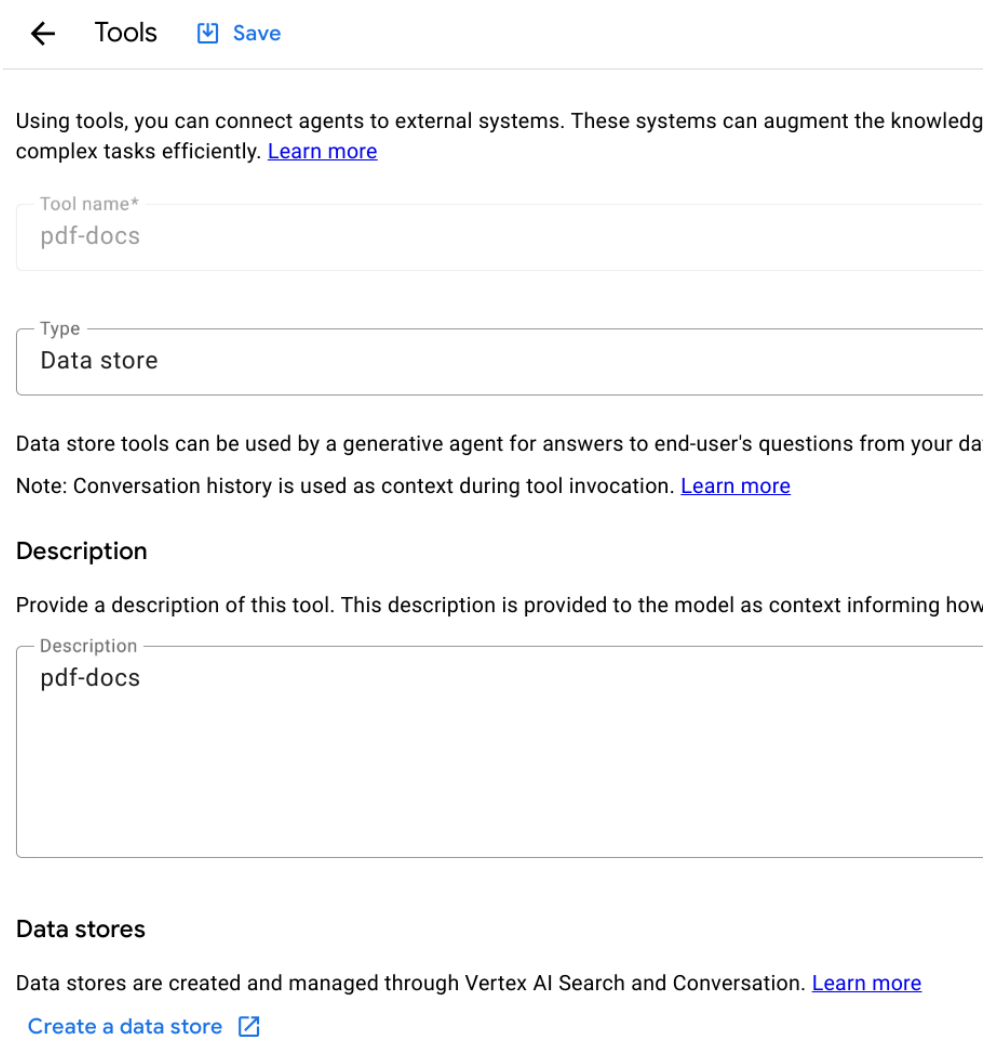
Scroll down to the bottom of this page



3. Click the +"`Data store`" icon



4. Use following values:
    - Tool name: `pdf-docs`
    - Type: `Data store`
    - Description: `pdf-docs`
5. Click "`Save`"



6. Click the "`Create a data store`" at the bottom on the page.
7. Click "`AGREE`" when asked about "Do you agree to have your search & conversation data stores in the us region?"
8. Type "`Google`" in the "Provide Company" field and press `Continue`
9. On the next screen, click "`CREATE DATA STORE`".
10. Select "`Cloud Storage`" as data source.

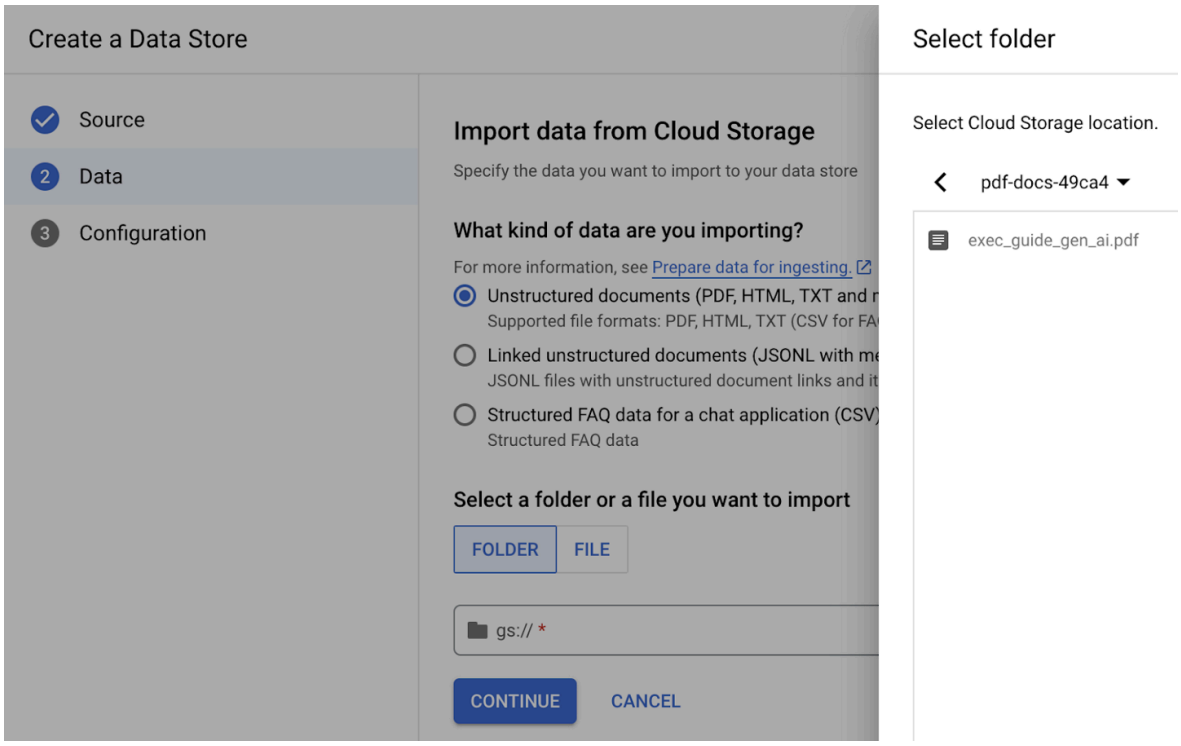
NOTE: *Prepare data for ingesting*

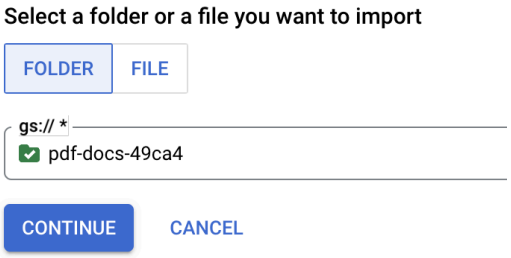
- *https://cloud.google.com/generative-ai-app-builder/docs/prepare-data*
- *HTML and TXT files must be 2.5 MB or smaller.*

  PDF, PPTX, and DOCX files must be 100 MB or smaller.

You can import up to 100,000 files at a time.

11. Select: `unstructured documents`
12. Select your GCS bucket/folder.





13. Click continue:
14. For data store name type: "`pdf-docs`"
15. Expand Document Processing Options
16. Check `Include ancestor headings in chunks`.
17. Click "`Create`".

Configure your data store

Configure additional settings for your data store

Location of your data store

Multi-region

us (multiple regions in United States)

Your data store name

Data store name *

pdf-docs

ID: **pdf-docs_1729710903895**. It **cannot be changed later.** EDIT

☐ This data store contains access control information

▲ DOCUMENT PROCESSING OPTIONS

☑ Enable document processing config

Document parsing

Select a parser to apply to your documents. Learn more about parsing and chunking documents.

Default document parser

Layout Parser

Document chunking

You can't change document chunking after creating the data store

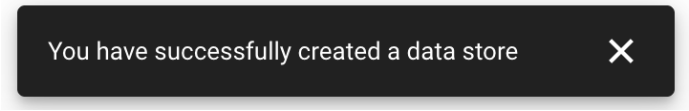
☑ Enable advanced chunking configuration

Chunk size limit

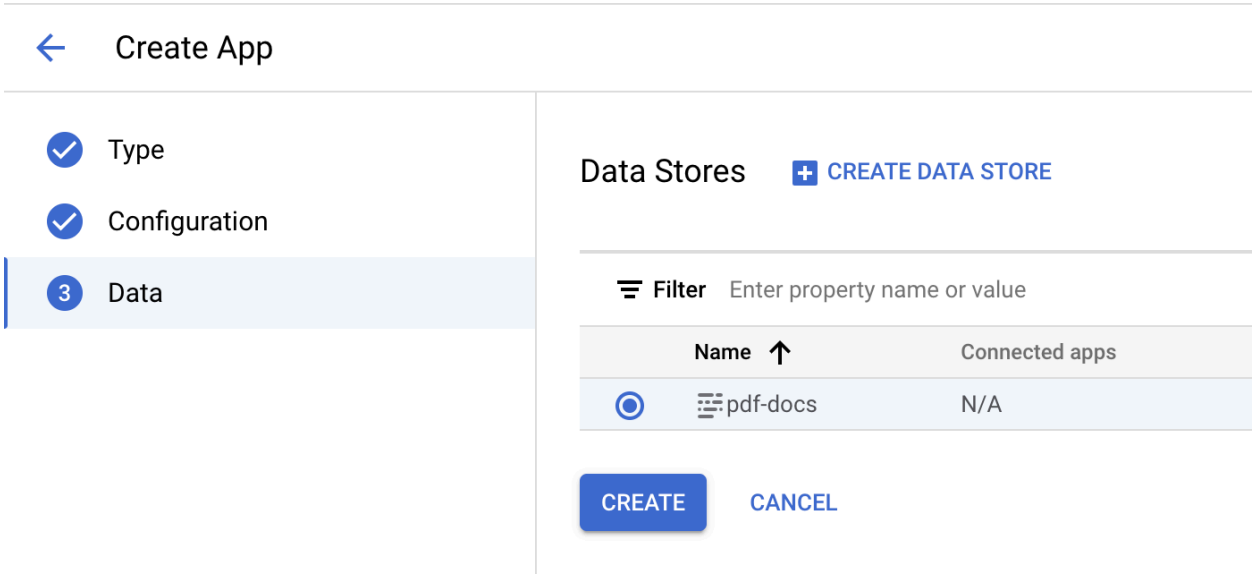
500 Tokens

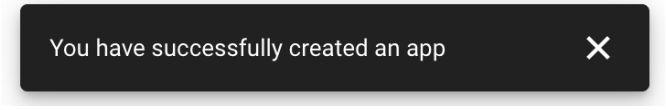
☑ Include ancestor headings in chunks

CREATE CANCEL



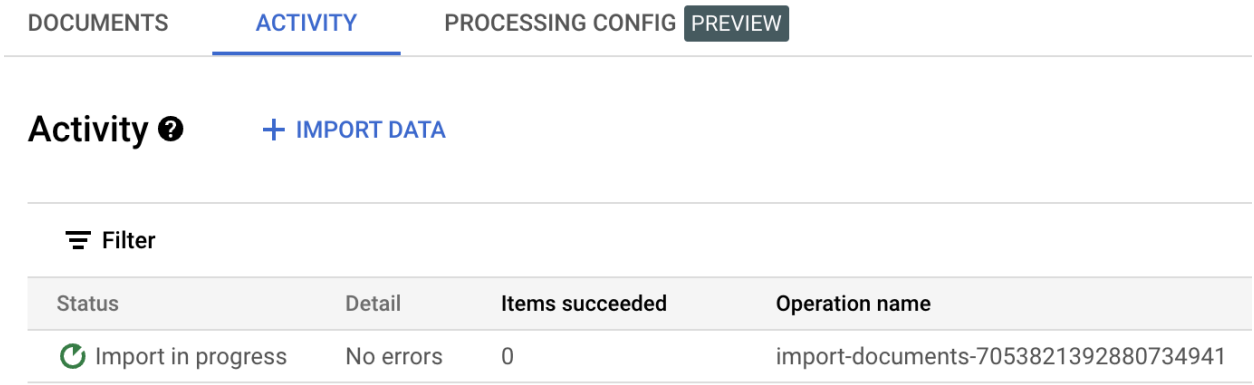
18. Select data store and click "`Create`"



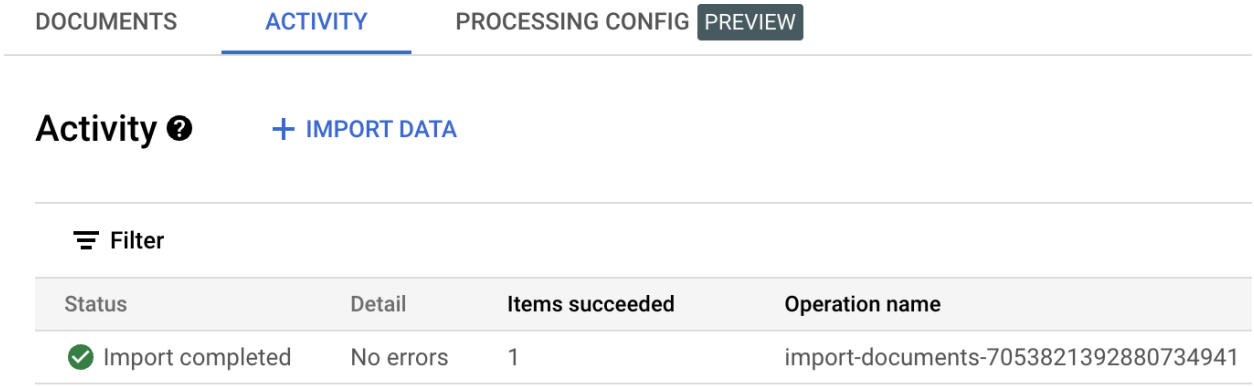


19. Click on the data store that is created, scroll down

20. Click on the data store and review Documents, Activity and Processing Config.



It will take ~5-10 minutes to complete the import.



Parsing and Chunking options

You can control content parsing in the following ways:

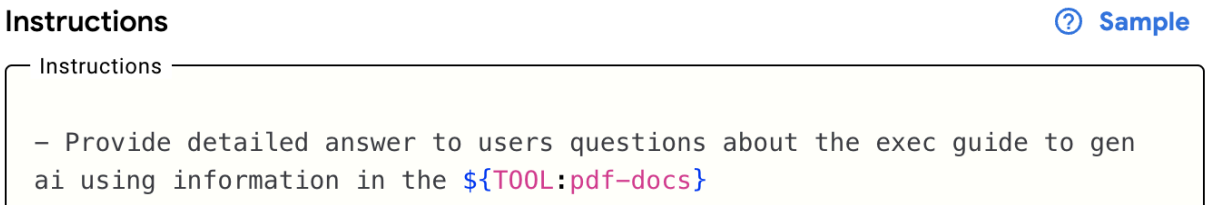
- *Digital parser. The digital parser is on by default for all file types unless a different parser type is specified. The digital parser processes ingested documents if no other default parser is specified for the data store or if the specified parser doesn't support the file type of an ingested document.*
- *OCR parsing for PDFs. Public preview. If you plan to upload scanned PDFs or PDFs with text inside images, you can turn on the OCR parser to improve PDF indexing. See About OCR parsing for PDFs.*
- *Layout parser. Public preview. Turn on the layout parser for HTML, PDF, or DOCX files if you plan to use Vertex AI Search for RAG. See Chunk documents for RAG for information about this parser and how to turn it on.*

Learn more about parsing and chunking documents.

Task 3: Agent's instructions configuration

1. Go back to the Agent Builder console, then Click on the *Playbook* icon in the navigation bar
2. Click on the Default Generative Agent
3. Add new instruction:

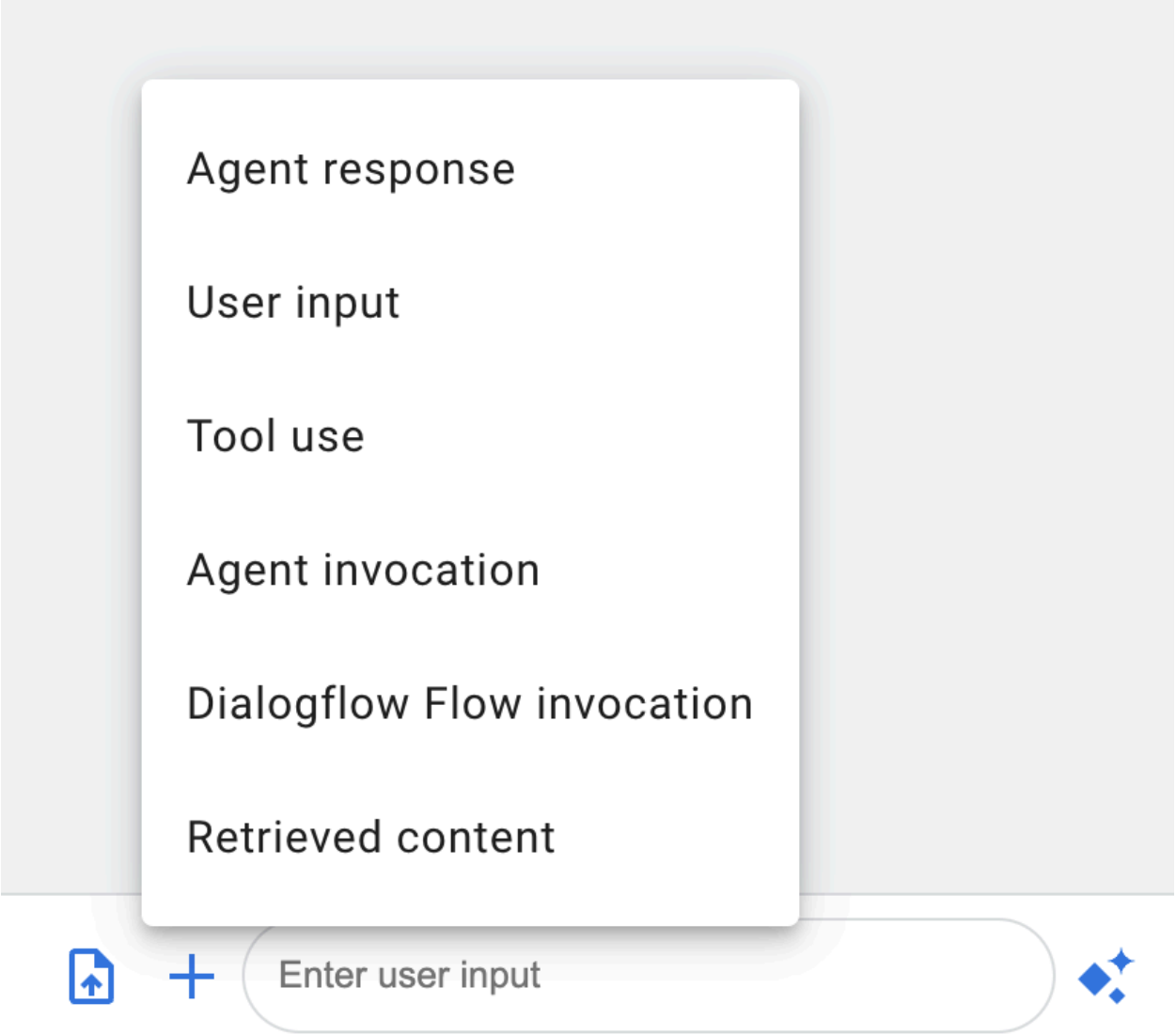
```
Unset
- Provide detailed answer to users questions about the exec guide to gen ai using information in the ${TOOL:pdf-docs}
```



4. Select the tool “pdf-docs” from the tools at the bottom of the page.
5. Save configuration.

Task 4: Create an example for PDF-Docs tool

1. Switch to the Examples tab. Create a new example.
2. Using actions "+":



3. Give it the Display Name “user-question-flow-1”
4. Add "User input":

```
What are the main capabilities?
```

5. Click the Add action and Add "Tool use".
    - Tool & Action should both be: "`pdf-docs`"
6. Check the Include Input (requestBody)
7. Add the following code to the Tool input:

Unset

```
{

  "query": "Main capabilities",

  "filter": "",

  "userMetadata": {},

  "fallback": ""

}
```

8. Add the following code to Tool Output:

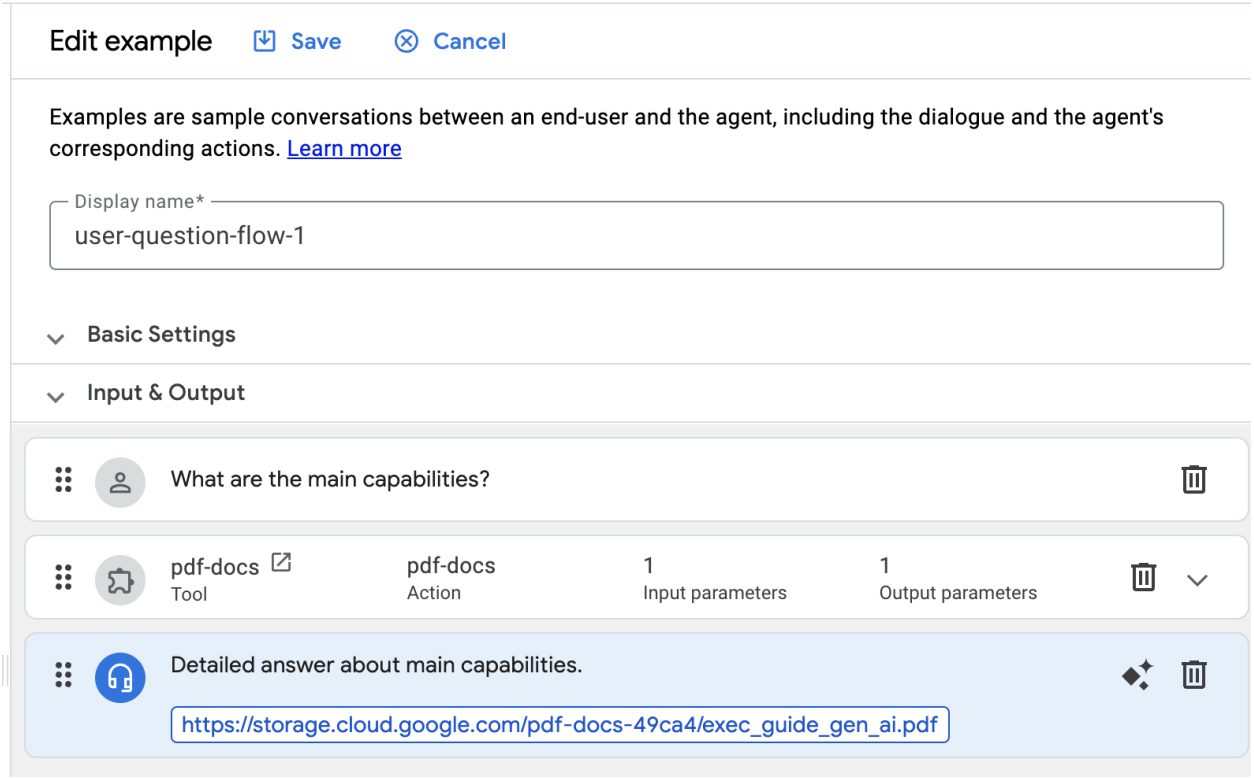
Unset

```
{

  "answer": "Detailed answer about main capabilities",

  "snippets": [

    {

      "uri":
"https://storage.googleapis.com/ancient-sandbox-322523-external-buck
et/exec_guide_gen_ai_shortened.pdf",

      "text": "Detailed answer about main capabilities",

      "title": "exec_guide_gen_ai"

    }

  ]

}
```

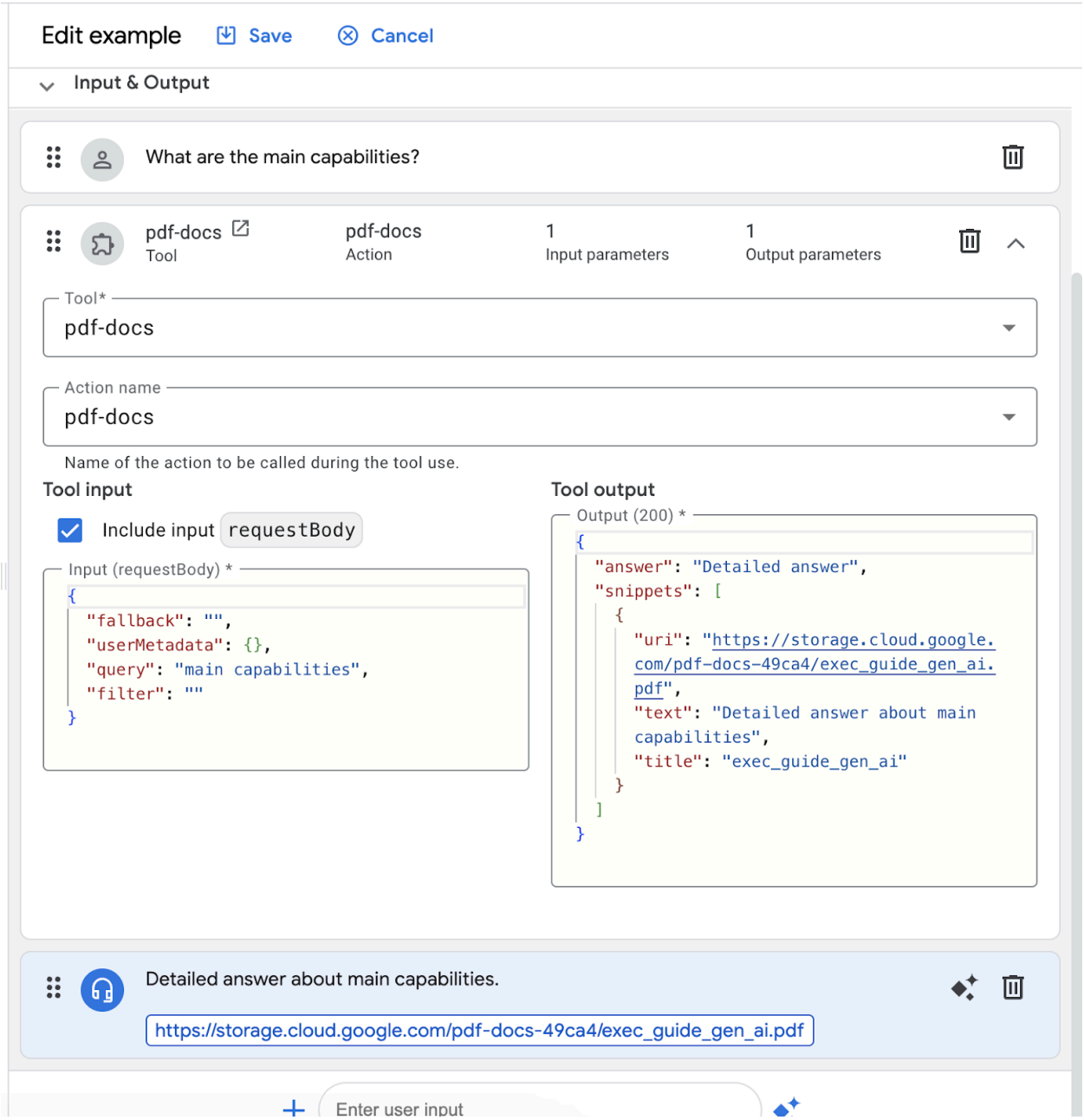
9. Add Action: "Agent response"
10. Add the following to Agent Response:

```
Unset
Detailed answer about main capabilities.


https://storage.googleapis.com/ancient-sandbox-322523-external-bucket/exec_guide_gen_ai_shortened.pdf
```

- Configured example:



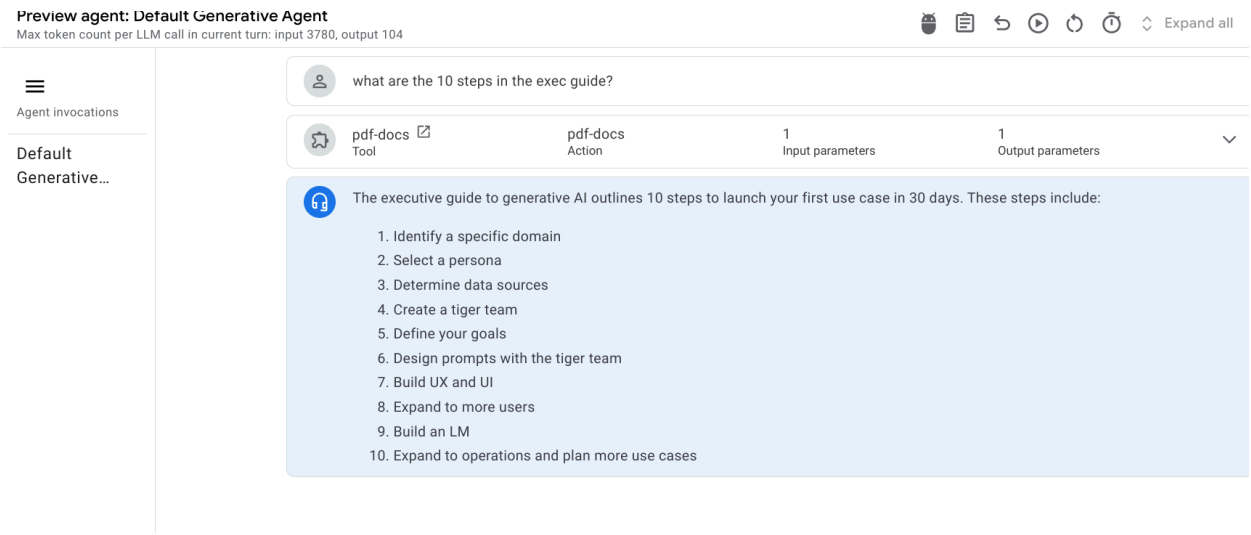
- Tool invocation configuration example:



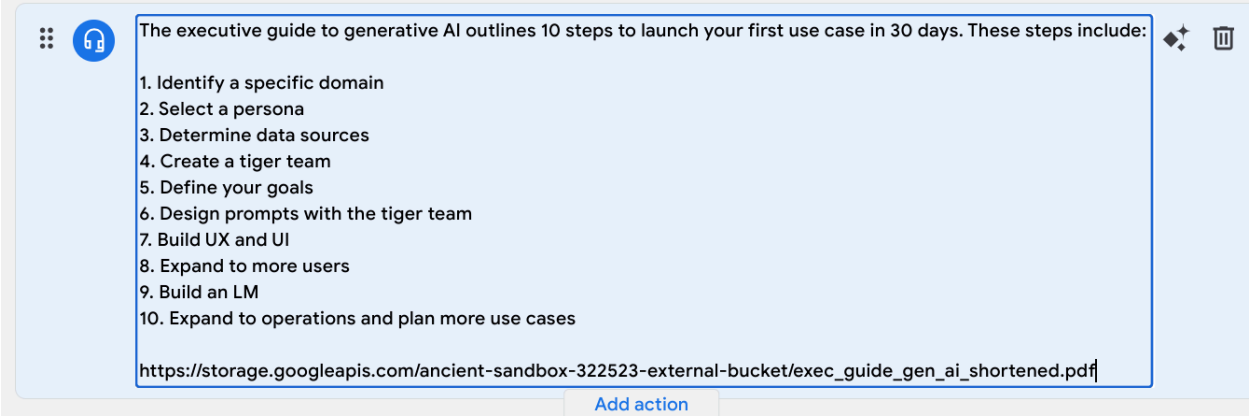
11. Once saved, click cancel to exit the editor.
12. Test the configuration by sending a question to the Agent in the Preview Agent emulator.
13. Ask the question:

Unset

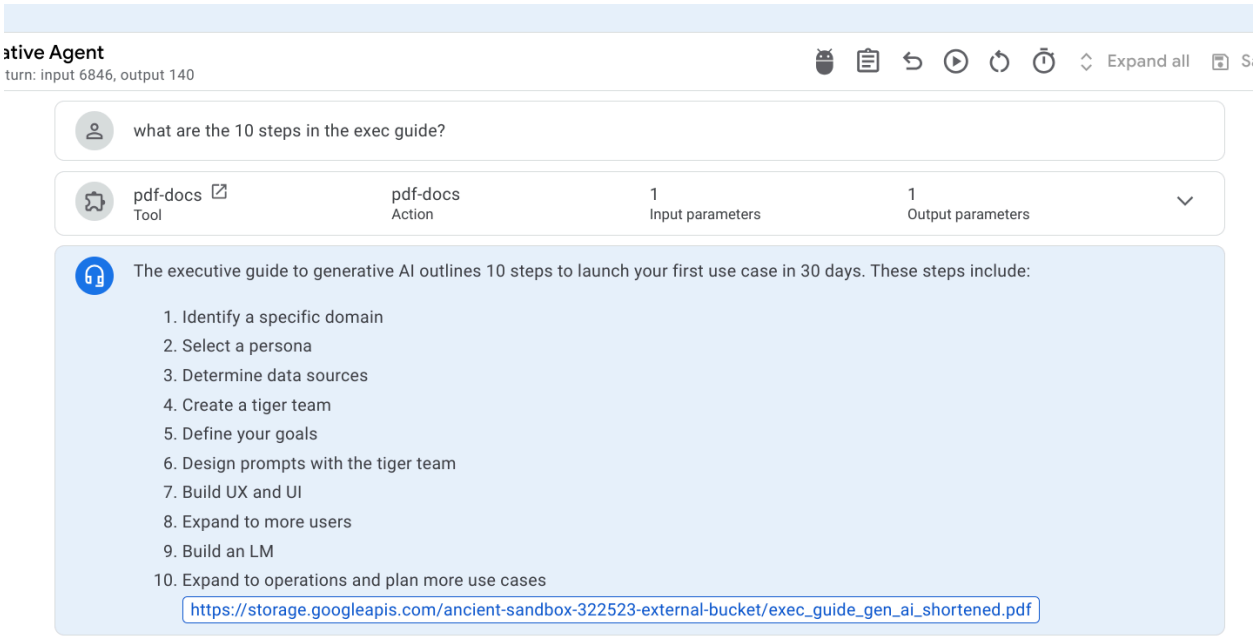
```
What are the 10 steps in the exec guide?
```



14. Select "`Default Generative Agent`" and click "`Save example`".
15. Provide a name "`user-question-flow`" and save.
16. Format agent response and include link to the pdf doc from the tool output section.

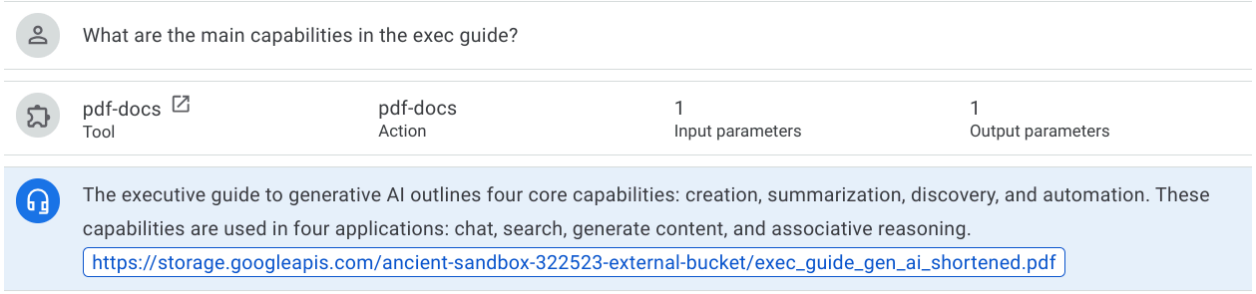


17. Click Create to save the example.
18. Click on the “Conversation History” from the tool bar on the left.
19. Click "`Replay conversation`" button.
20. Validate the updated response format.

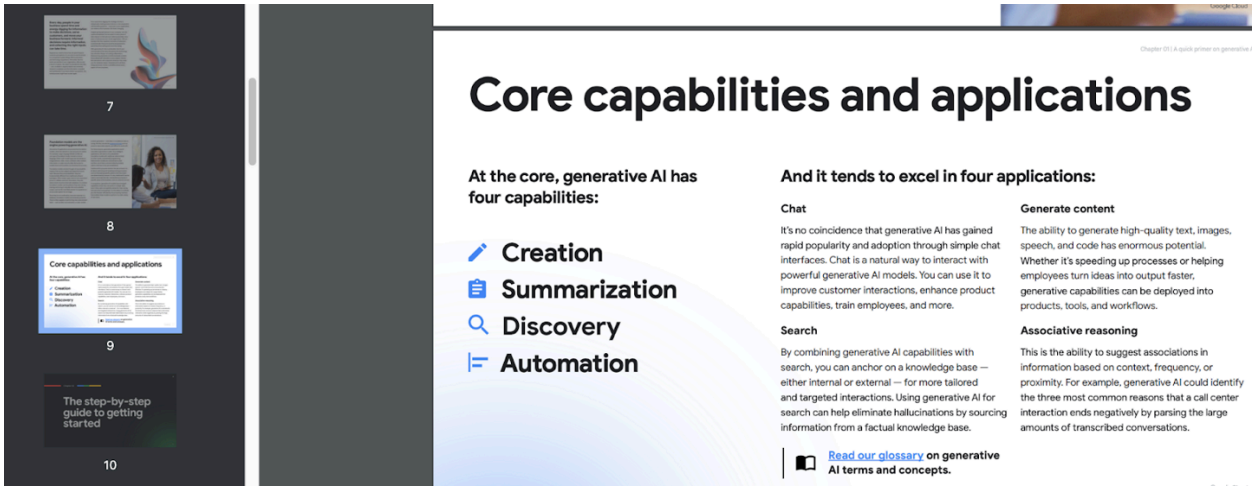


21. Ask another question:

```
Unset
What are the main capabilities in the exec guide?
```



- Source PDF document.



22. Question:

Unset

```
What should I consider when evaluating projects?
```

What should I consider when evaluating projects?

pdf-docs
Tool

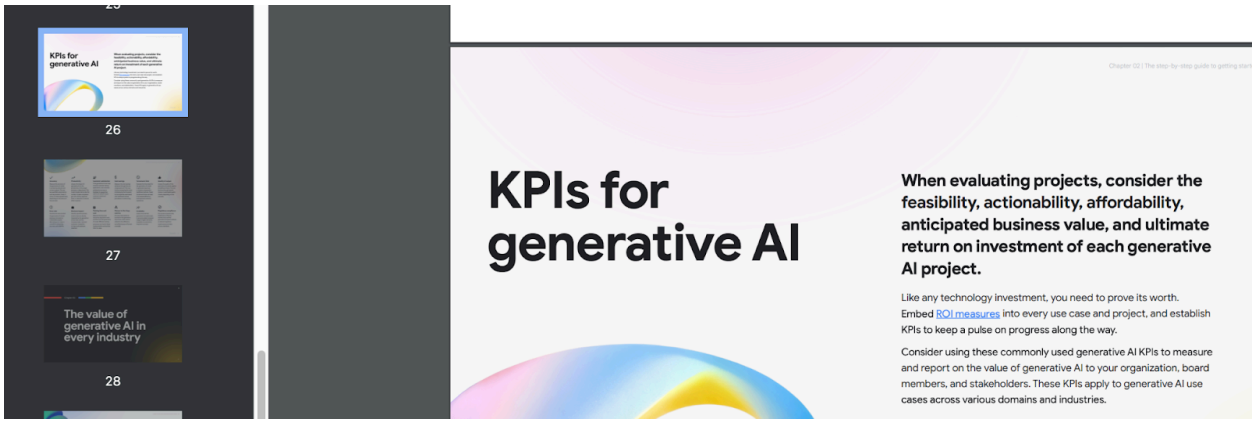
pdf-docs
Action

1
Input parameters

1
Output parameters

When evaluating generative AI projects, consider the feasibility, actionability, affordability, anticipated business value, and ultimate return on investment. You should also embed ROI measures into every use case and project, and establish KPIs to keep a pulse on progress.
https://storage.googleapis.com/ancient-sandbox-322523-external-bucket/exec_guide_gen_ai_shortened.pdf

- Source PDF document.



23. Question:

Unset

```
What are the priority use cases in Retail and CPG in the exec guide?
```

What are the priority use cases in Retail and CPG in the exec guide?

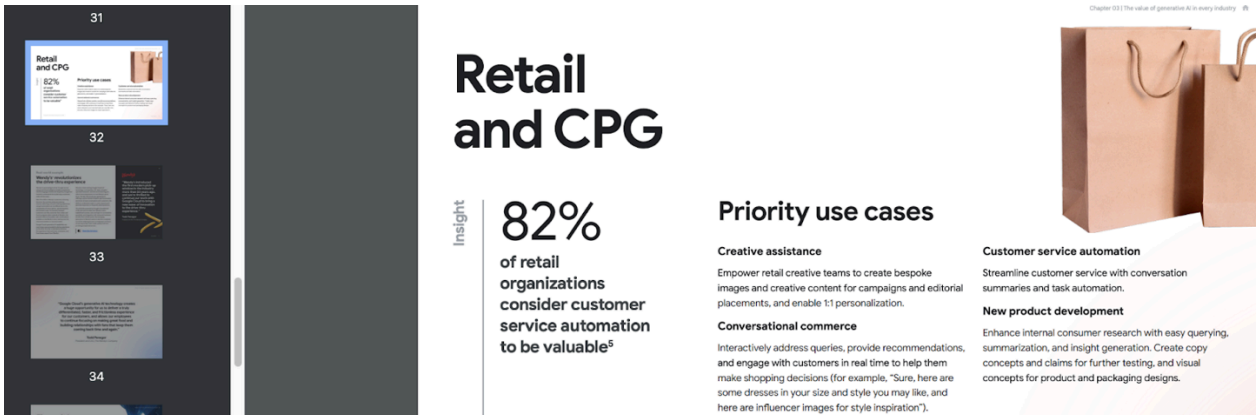
pdf-docs Tool	pdf-docs Action	1 Input parameters	1 Output parameters

The executive guide to generative AI outlines three priority use cases for retail and CPG:

- Creative assistance
- Conversational commerce
- Customer service automation

https://storage.googleapis.com/ancient-sandbox-322523-external-bucket/exec_guide_gen_ai_shortened.pdf

- Source PDF document.



Congratulations!

In this lab, you have successfully learnt about how to build an agent using the Agent Builder on unstructured datasets.

Lab Authors: Naresh Jasotani(nareshjasotani@google.com) / William Dufrin (Wdufrin@google.com)


You're Awesome!


Kudos