



Understanding Generative AI

In this module, you learn to ...

01

Differentiate Generative AI (GenAI) from other Artificial Intelligence and Machine Learning solutions

02

Recognize how GenAI uniquely adds value to key business processes

03

Assess the impact GenAI can have on your organization



Topics

- 01 Generative AI in Action
- 02 GenAI vs. AI/ML
- 03 GenAI Applications
- 04 Business Impact of GenAI



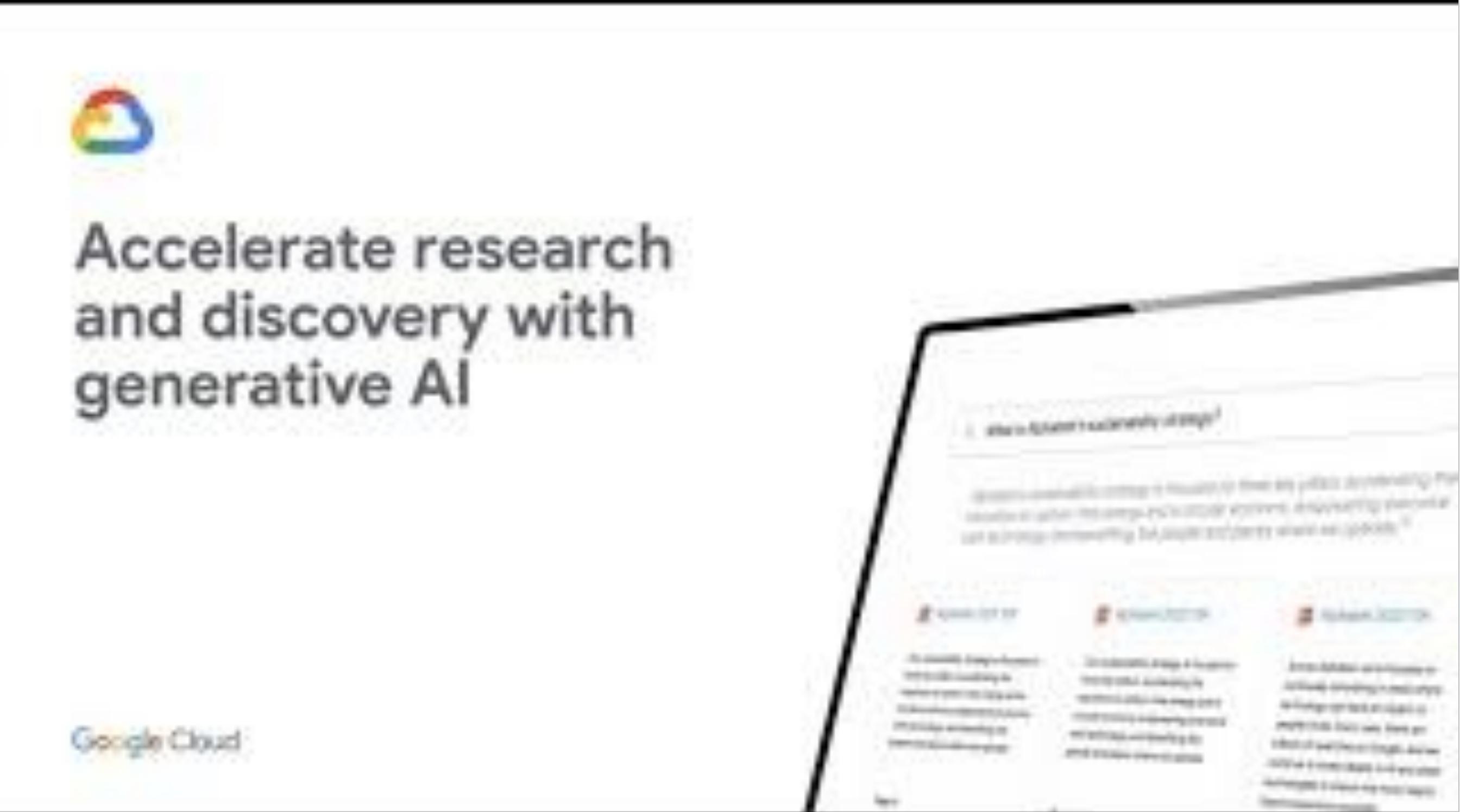
GenAI in Retail

The slide features a white background with a thin black border. In the top left corner is the Google Cloud logo, which consists of a multi-colored cloud icon above the word "Cloud". The main title, "Accelerate product catalog management with generative AI", is centered in large, bold, dark gray font. Below the title is a smaller, faint "Google Cloud" watermark. To the right of the text is a screenshot of a retail website's product catalog page. The page displays a grid of product images, including shoes, clothing items, and a slide. A black curved arrow points from the text "Accelerate product catalog management with generative AI" towards the screenshot.

Accelerate product catalog management with generative AI

Google Cloud

GenAI in Financial Services



The image shows a screenshot of a Google Cloud landing page titled "GenAI in Financial Services". The page features a large, colorful cloud icon at the top left. Below it, the main headline reads "Accelerate research and discovery with generative AI". To the right, there's a section with three cards, each containing a title and a short description. The first card is titled "Generative AI for Research", the second "Generative AI for Compliance", and the third "Generative AI for Risk". At the bottom left, the "Google Cloud" logo is visible.

Accelerate research and discovery with generative AI

Generative AI for Research

Generative AI for Compliance

Generative AI for Risk

Google Cloud

GenAI in Customer Service

The image shows a white smartphone with a black border. On the screen, there is a message from Google Cloud. The message reads:

It takes 10 times less time to resolve customer issues using AI than with your current team.

You're experiencing service disruptions, which is severely impacting control. It's currently impossible. Who can you work with to fix this together? Do you want to continue?

At the bottom of the screen, there are two blue circular buttons labeled "Cancel" and "Fix".

Topics

01

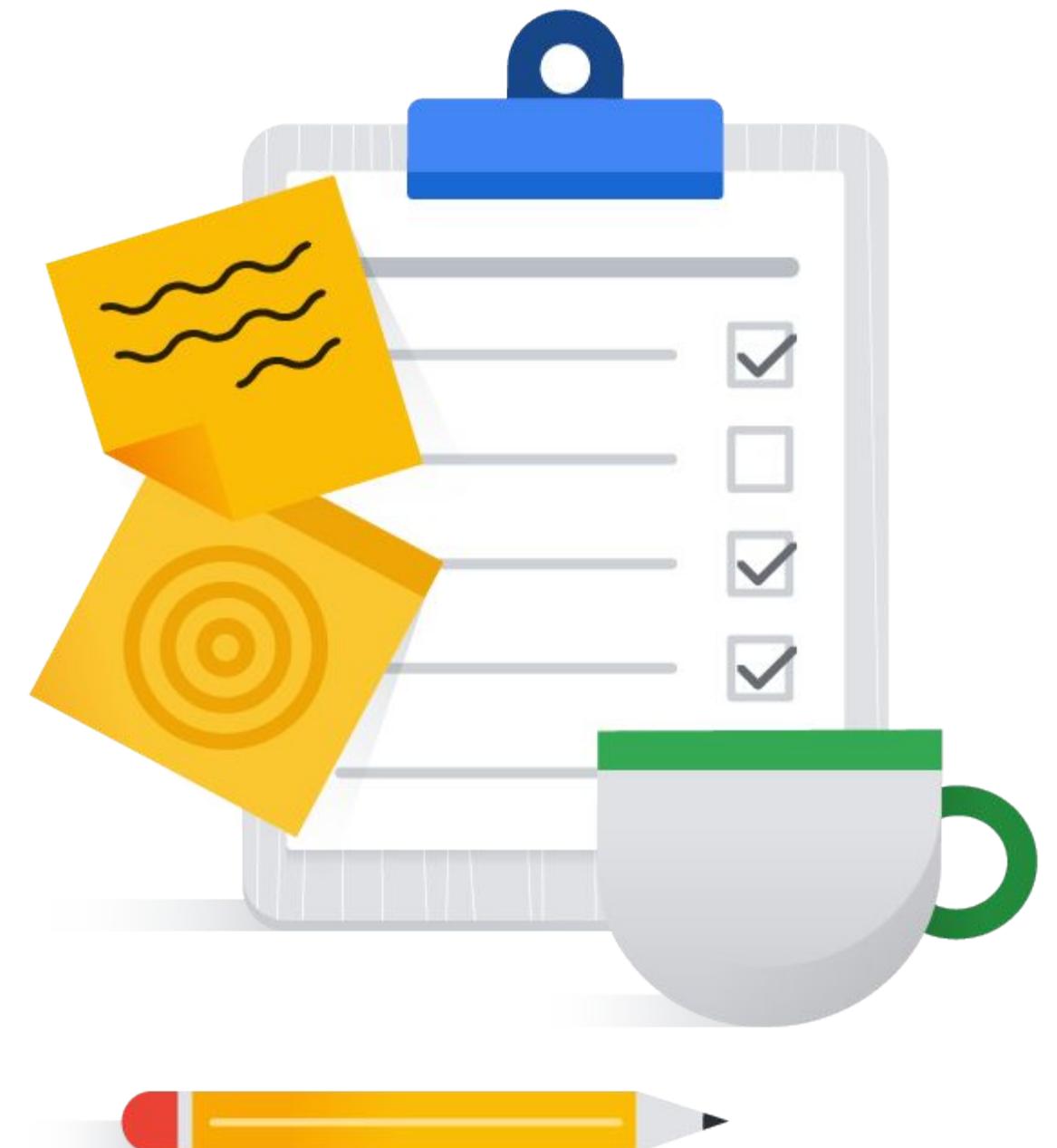
GenAI vs. AI/ML

02

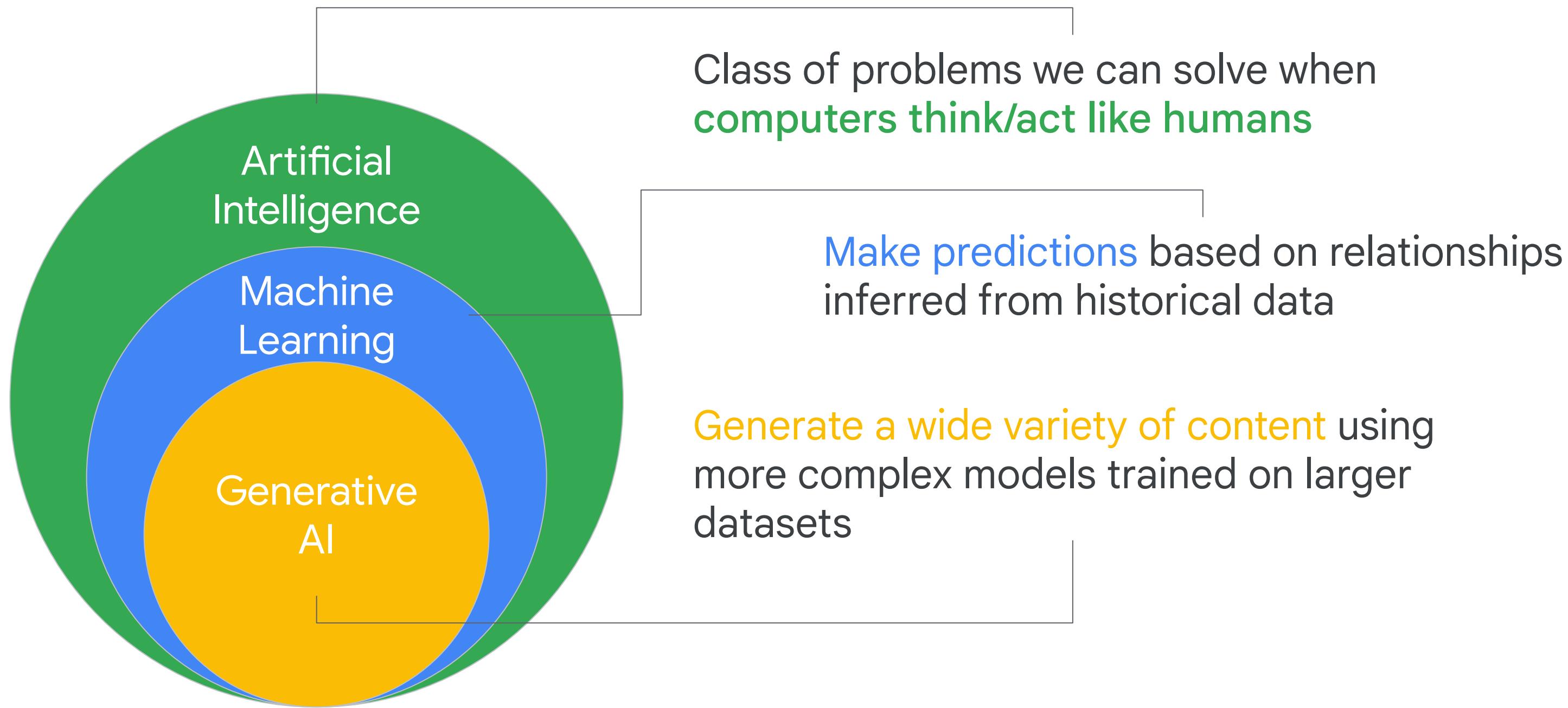
GenAI Applications

03

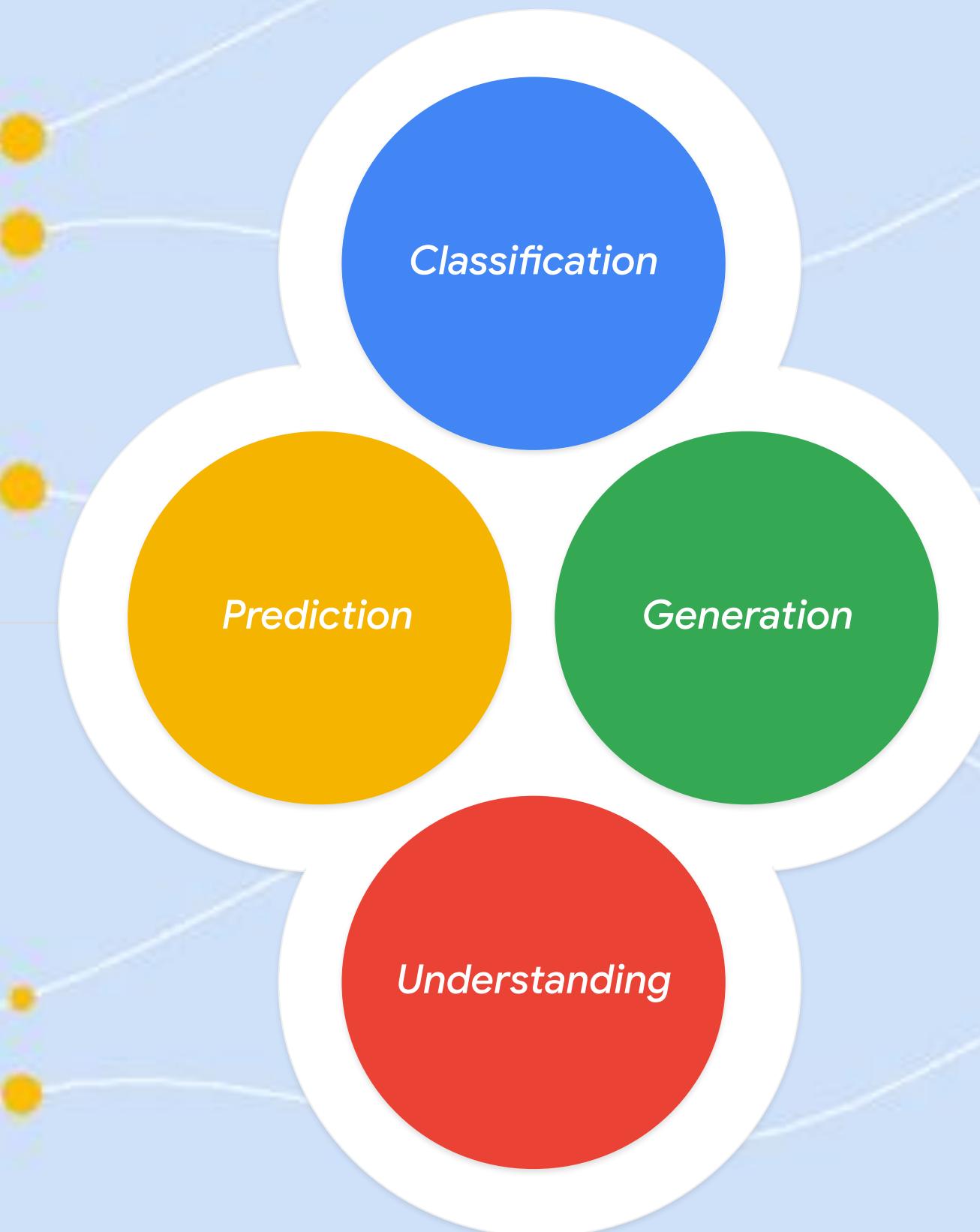
Business Impact of GenAI



Machine Learning is a type of AI, and Generative AI (GenAI) is a type of machine learning



AI is evolving from analysis to creation



Large language models are - large!



ML algorithms that can **recognize, predict, and generate** human languages



Pre-trained on petabyte scale text-based datasets resulting in large models with **10s to 100s of billions of parameters**



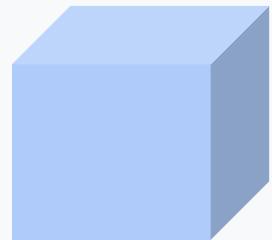
LLMs are normally **pre-trained on a large collection of text** followed by fine-tuning on a specific task



LLMs can also be called **Large Models** (includes all types of data modality) and **Generative AI** (a model that produces content)

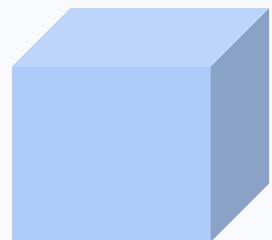


Go read this huuuuuuge pile of books.



So, you've learned about cats and millions of other concepts ... what's a cat?

A cat is a small, domesticated carnivorous mammal.



Generative language models

LaMDA, PaLM, GPT-3, etc.

Why are large language models different?



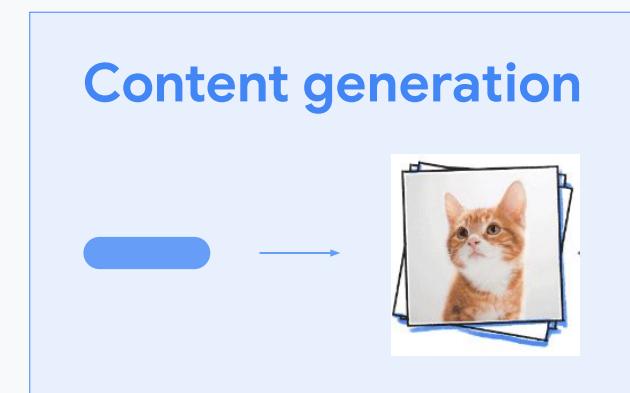
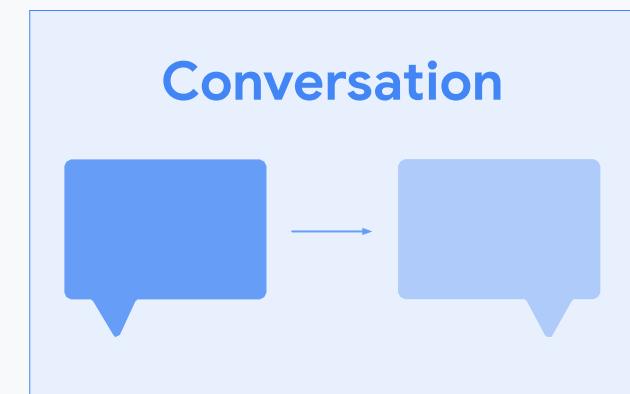
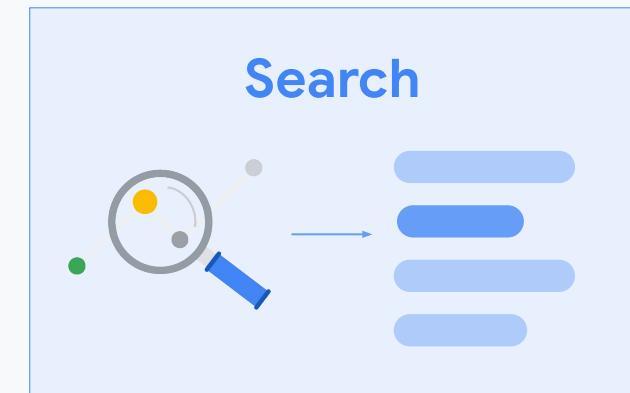
LLMs are characterized by **emergent abilities**, or the ability to perform tasks that were not included in their training examples



LLMs contextual understanding of human language **changes how we interact** with data and intelligent systems



LLMs can find patterns and connections in **massive, disparate data corpora**



Topics

01

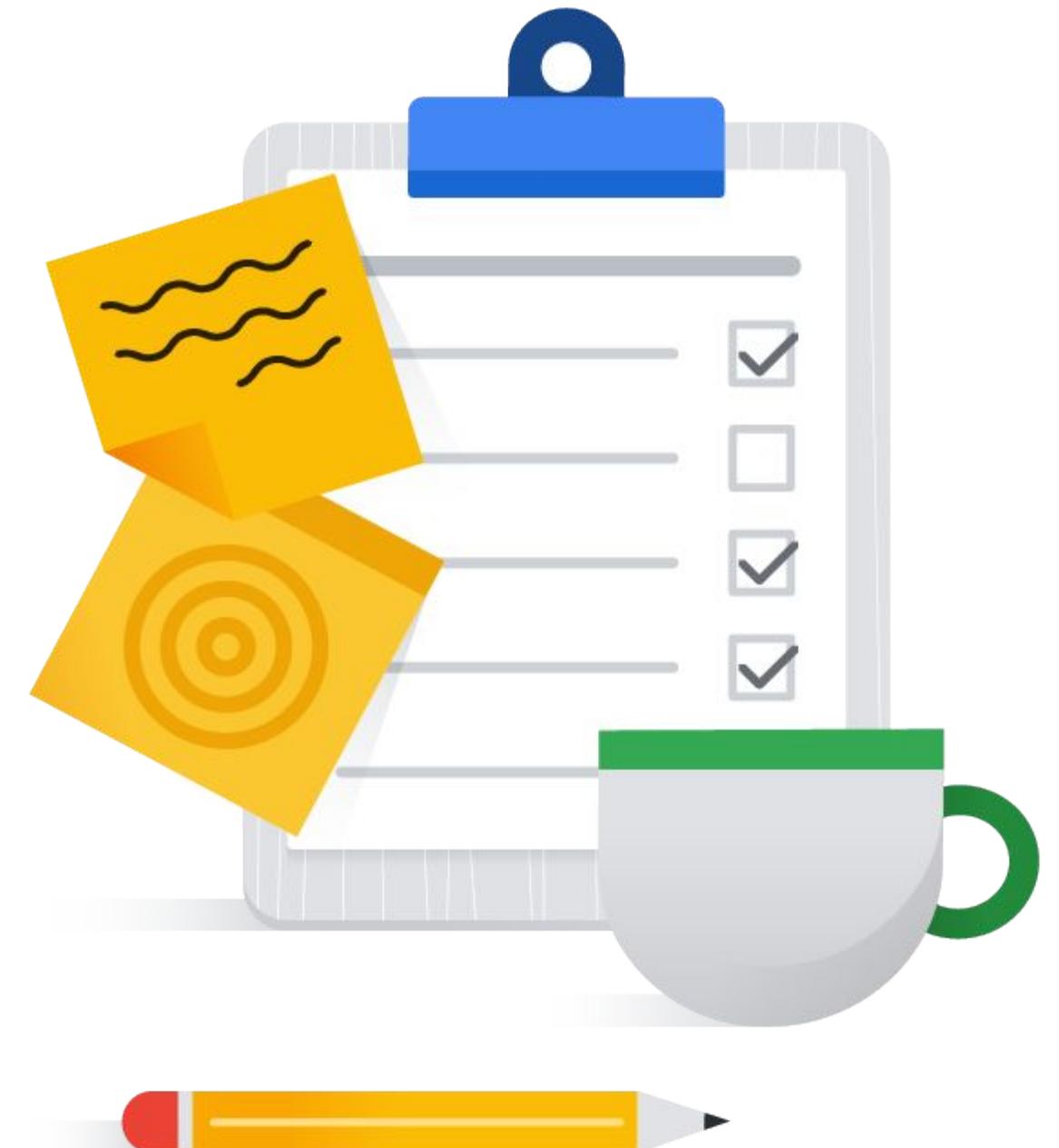
GenAI vs. AI/ML

02

GenAI Applications

03

Business Impact of GenAI



GenAI can solve different classes of problems



CREATE

Bring your thoughts and visions to life

Use cases

- Generate product descriptions from images
- Generate images from text



SUMMARIZE

Condense and summarize your knowledge base into a simple format

Use cases

- Content/video summarization
- Intra-knowledge Q&A



DISCOVER

Help your customers and employees find what they need at the right time

Use cases

- Search for a document
- Machine-generated event monitoring



AUTOMATE

Automate your customer service across multiple channels

Use cases

- Contract information extraction
- Feedback classification and ticket creation

GenAI can solve different classes of problems



CREATE

Bring your thoughts and visions to life

Use cases

- Generate product descriptions from images
- Generate images from text



SUMMARIZE

Condense and summarize your knowledge base into a simple format

Use cases

- Content/video summarization
- Intra-knowledge Q&A



DISCOVER

Help your customers and employees find what they need at the right time

Use cases

- Search for a document
- Machine-generated event monitoring



AUTOMATE

Automate your customer service across multiple channels

Use cases

- Contract information extraction
- Feedback classification and ticket creation

Where could **you** use GenAI in your job?

You can create a library of use cases

Document Search & Synthesis

Effectively find the most relevant documents and summarize their contents

Product / Content Catalog Discovery

Effectively find the most relevant Products / Content listings from a inventory catalogue

Regulatory Compliance Automation

Interpret regulatory policy / documents to identify potential violations relative to operating procedures

Business Process Automation

Automating the information retrieval and recommendation step of a recurring business process

Product / Content Recommendation

Recommend personalized Product / Content / Next Best Action from a catalogue

Research Acceleration

Find complex subject domain information across many disparate sources and synthesis the findings

Online interactions made conversational

Public Website Navigation

Effectively find information from a website via multi-modal inputs and conversational queries

Intra-Knowledge Q & A

Conversationally query questions for answers from internal knowledge sources

Documentation Generation

Write new documentation based on summarization of other documents & software code

Customer Service Automation

Effectively service customers requests for information and service provisioning

Creative Assistance

Empower creative teams to create bespoke images and creative content for campaigns and editorial content

Developer Efficiency

Complete and augment code to make your engineering team more efficient and effective

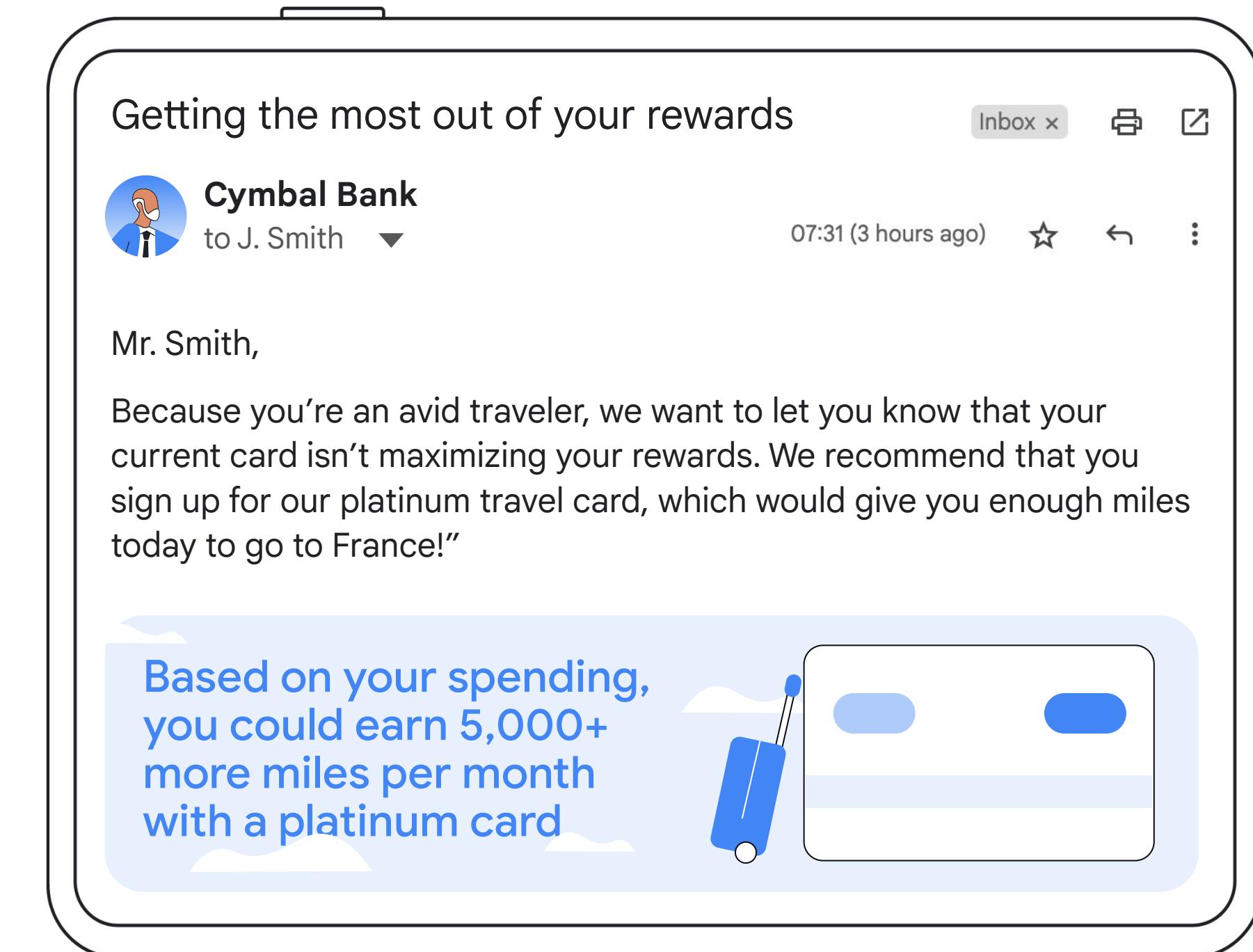
Complex data, intuitively accessible

Content generation at the click of a button

GenAI can provide hyper-individualized content

Improve cross-sell and retention strategies through one-to-one instead of one-to-many messaging.

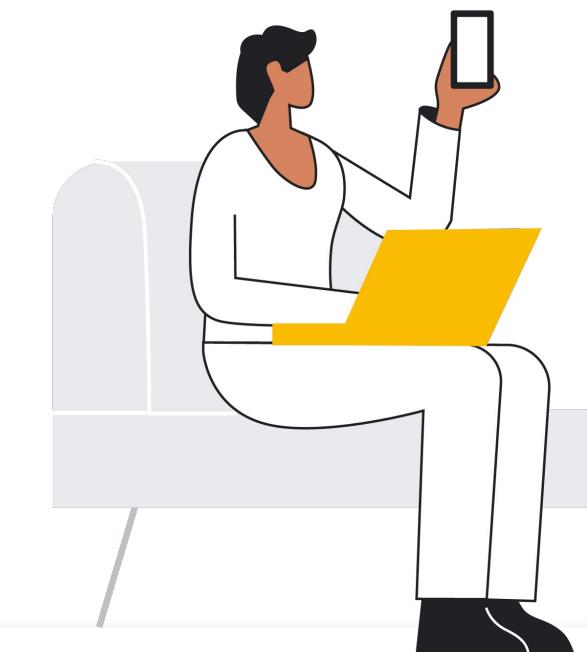
Tailor financial product recommendations via conversational language that are hyper-personalized and reinforces a “you know me” experience.



GenAI can elevate consumer and patient-centricity

Assist patients in explanation of plans & benefits, connect with doctors, book appointments and enhance consumer access to services in simplified language.

Example: A pregnant patient interacts with digital concierge to find gynecologists, sign up for classes, and find other relevant information.



On a hospital website, an expecting parent asks the digital concierge for information instead of browsing hospital website

The image shows a digital interface for a "Digital Concierge". At the top right are three small circular icons. Below them is a blue header bar with the text "Digital Concierge". The main area consists of several grey speech bubbles containing text. The first bubble says "I just found out I am pregnant. What do I need to do?". The second bubble responds with "Congratulations! You should make an appointment with your doctor to confirm your pregnancy and discuss your options. You may also want to consider taking a prenatal vitamin and eating a healthy diet.". The third bubble asks "Tell me female gynecologists at hospital". The fourth bubble lists names: "The female gynecologists at the hospital are: Dr. Jane Doe, Dr. Celine Doe, Dr. Jenny Doe, Dr. Andrea Doe, Dr. Amanda Doe, Dr. Janet Doe, Dr. Sheila Doe, Dr. Izzy Doe, Dr. Janine Doe, Dr. Sara Doe, and Dr. Charlene Doe.". The fifth bubble asks "How do I sign-up for childbirth and parenting classes". The sixth bubble provides instructions: "You can sign up for childbirth and parenting classes by calling the Cymbal New Parents office at 555.000.1234 or emailing them at newparents@cymbalparenting.com".

GenAI-enabled search drives faster and better insights

Query and extract insights from public datasets (medical literature) and private datasets (ELN). Quickly and accurately source insights from public and private datasets.

Summarize research into plain language

Example: Searching FDA website and public websites to search for secondary indication for existing pharmaceutical products

“Show me all FDA submissions related to Cymbal drug”



Here is a list of FDA submissions for drugs similar to Cymbal that have been filed for secondary indications:

Drug A: Secondary indication for the treatment of pain

Drug B: Secondary indication for the treatment of nausea

Drug C: Secondary indication for the treatment of vomiting



GenAI can drive developer efficiency and reduce risk

Help developers understand the underlying regulatory or business changes that will require them to change code, and assist in automating coding changes

Example: Gen AI can summarize a relevant section of the Basel III framework to help a developer understand the section's intent, identify the related parts of a product document and code repository.

"Check our data transformation repo for SQL queries that need to be updated for Basel III"

```
SELECT
  *,
CASE
    WHEN deal_type IN ('New client', 'Existing client, new transaction') THEN 'New_client_or_transaction'
        WHEN deal_type = 'Existing client, refinance of existing transaction' THEN 'Refinance_of_existing_transaction'
    END AS table_A_category,
CASE
    WHEN transaction_type = 'CRE Investment' THEN
        CASE
            WHEN ltv_at_origination < 0.50 THEN 'A. <50%'
                WHEN ltv_at_origination >= 0.50 AND ltv_at_origination < 0.60 THEN 'B. 50.00 - 59.99%'
                    WHEN ltv_at_origination >= 0.60 AND ltv_at_origination < 0.65 THEN 'C. 60.00 - 64.99%'

```

Topics

01

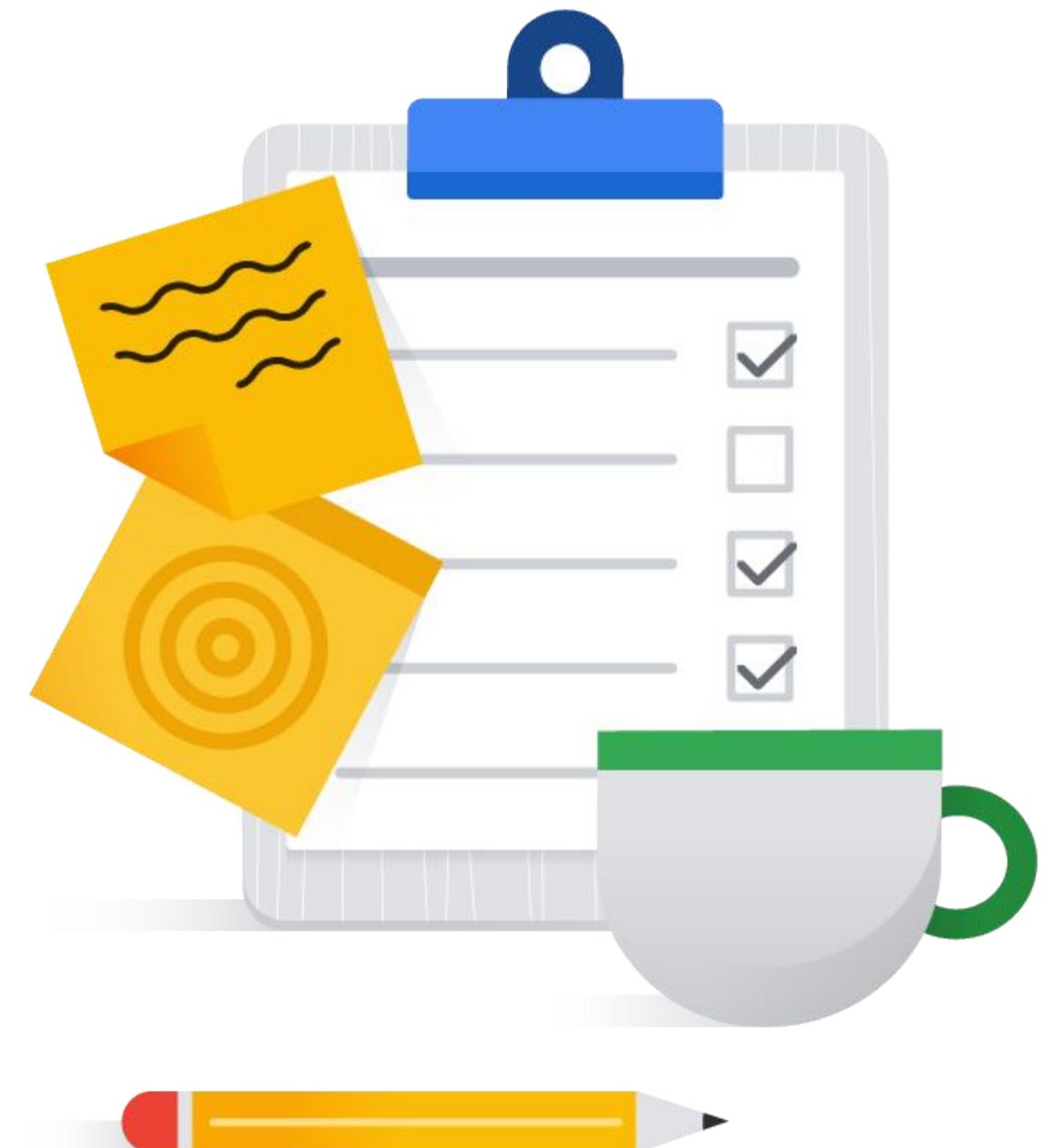
GenAI vs. AI/ML

02

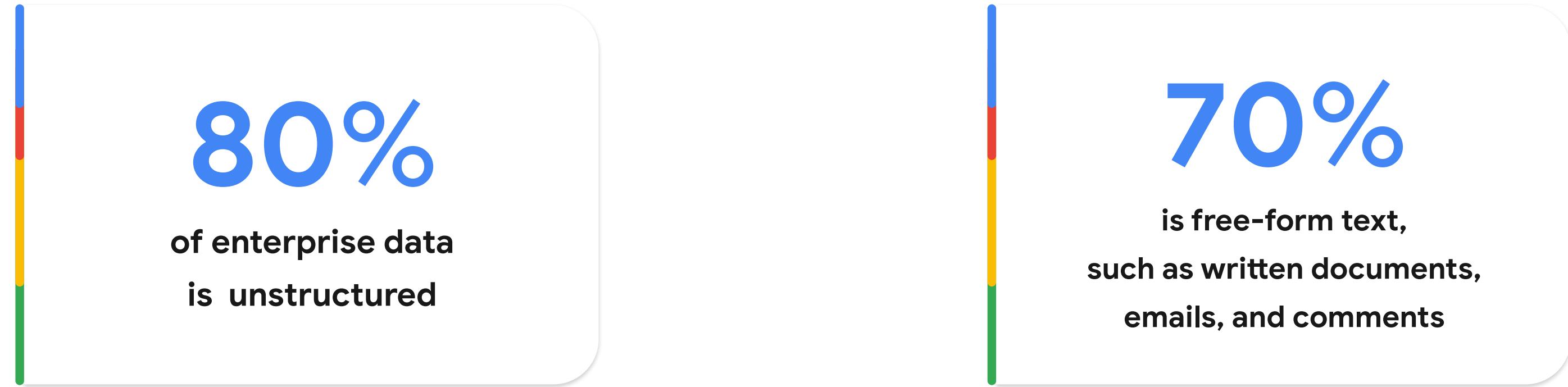
GenAI Applications

03

Business Impact of GenAI



Most enterprise workflows deal with unstructured data



- Customer emails
- Call transcripts
- Multimedia content (digital photo, audio, video)
- Knowledgebase to address customer queries

- Product reviews, campaign reviews
- PDFs, presentations, documents
- Machine-generated data

Here are some examples of where and how



Complex data,
intuitively accessible

Improve time-to-value to search, navigate, and extract insights and understanding from large amounts of complex data



Analyst



Online interactions
made conversational

Improve customer experience, reaching larger client bases by making online interactions more natural, conversational, and rewarding



Customer Service



Content generation at
the click of a button



Creative



Customize Foundational
models

Customize large models and incorporate state of the art generative capabilities natively into your own internal ML operational platforms



AI
Practitioner

AI can improve the product discovery experience, increase basket size, and reduce cart abandonment.

72% of consumers'
favorite retail websites are ones where they
can quickly find what they are looking for.

72% of consumers
buy the product they're looking for
when they can find it quickly and easily.

Reference: [GOOGLE CLOUD: New Research on Search Abandonment in Retail \(March '23\)](#)

The global healthcare industry
is expected to reach
\$21.06 trillion by 2030.

Customer service will be key
to determining who capitalizes
on this opportunity.

32% said

they switched providers in 2021 did so
because of poor healthcare navigation.

Reference: Accenture 2022

32% said

customer service was a top factor
in choosing a provider.

Reference: Press Ganey 2022

3% said

COVID-19 made them more aware of
the need for responsiveness in healthcare

Reference: Press Ganey 2021

30%
of US Millennials + Gen Z
canceled a streaming service in
the last 6 months.

Nearly $\frac{1}{2}$ of viewers feel that
“*finding content is too difficult*”

Almost $\frac{1}{3}$ of viewers said personalized
content recommendations would be a
key reason to stay

Source: [PWC: Global Entertainment & Media Outlook 2022 - 2026](#)

In this module, you learned to ...

01

Differentiate Generative AI (GenAI) from other Artificial Intelligence and Machine Learning solutions

02

Recognize how GenAI uniquely adds value to key business processes

03

Assess the impact GenAI can have on your organization





Generative AI on Google Cloud

In this module, you learn to ...

01

Differentiate between Google Cloud Generative AI tools

02

Consider how you can help ensure your organization is building and using AI solutions responsibly

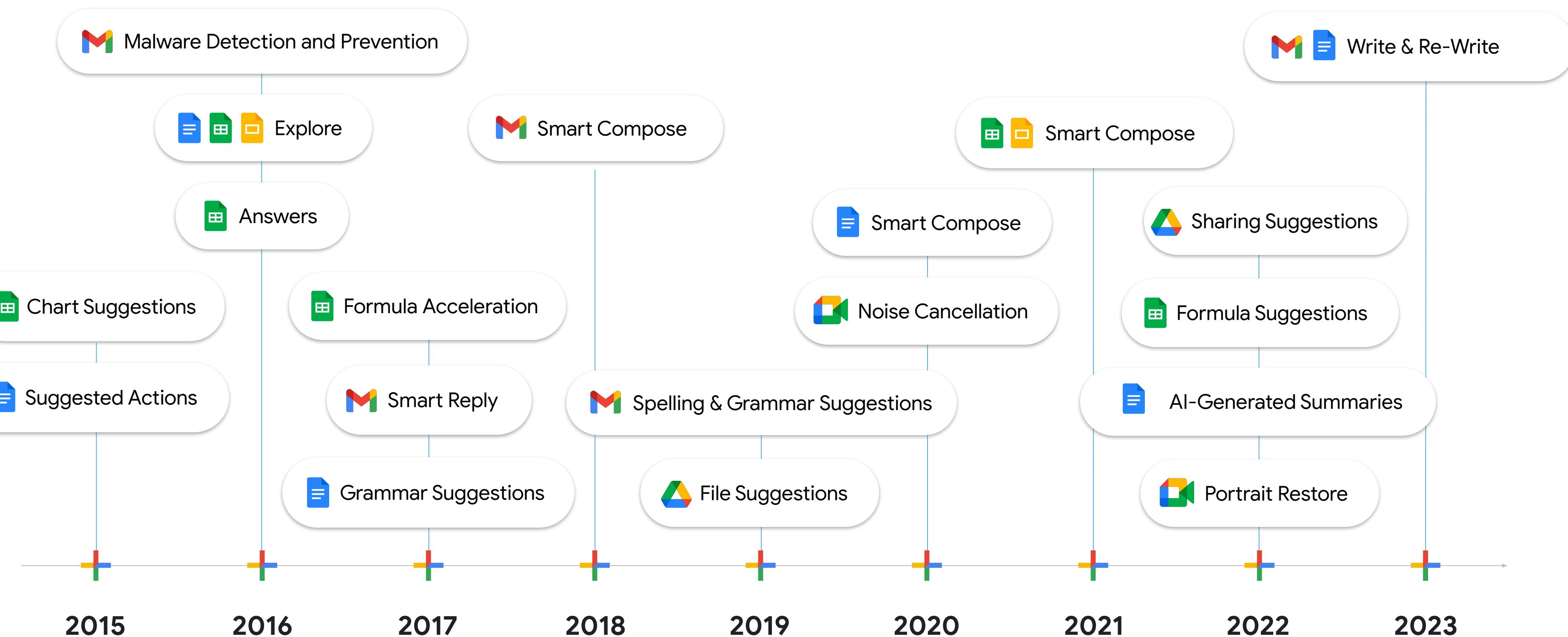


Topics

- 01 Duet AI for Google Workspace
- 02 Enterprise AI - APIs, and models
- 03 GenAI Studio
- 04 Model Garden
- 05 Gen App Builder
- 06 Responsible AI



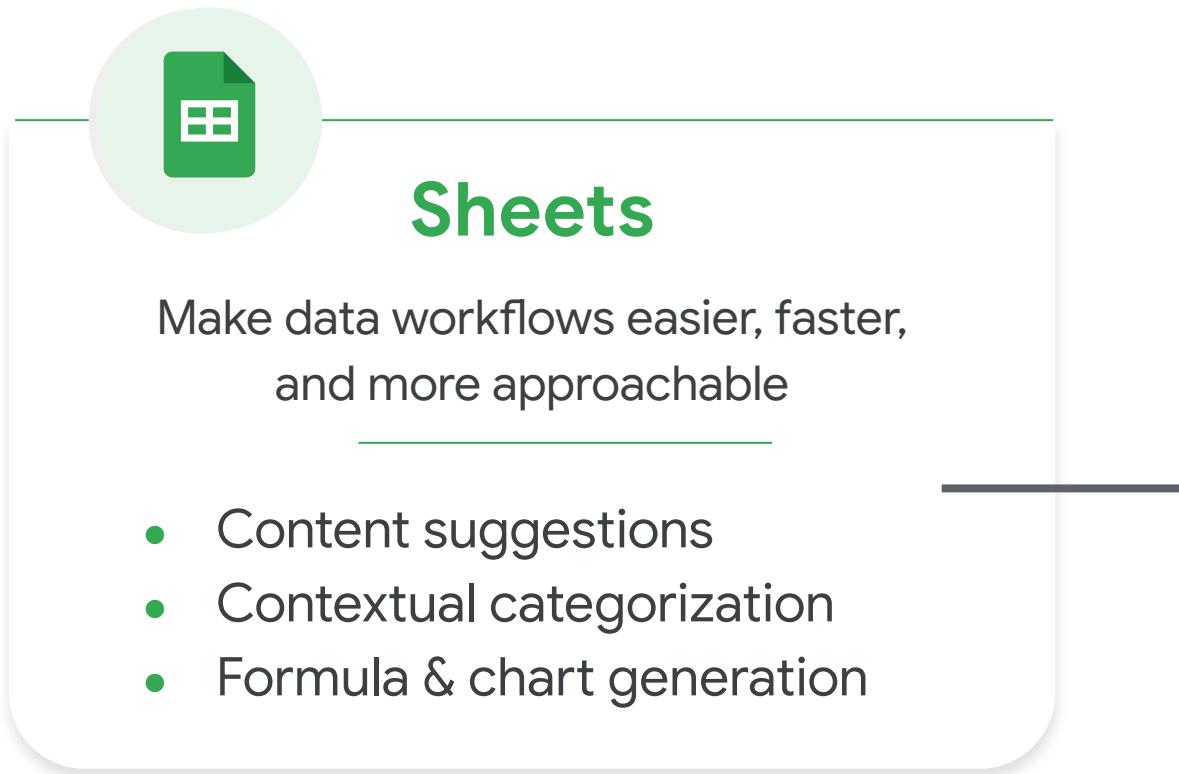
Google has put AI to work for 3 billion users in Workspace



Google Docs will provide features like “Help me write”

The diagram illustrates the integration of AI writing assistance into the Google Docs interface. On the left, a white rounded rectangle contains the Google Docs logo (a blue document icon) and the word "Docs". Below this, the tagline "Think bigger, work faster, and supercharge your imagination" is followed by a horizontal line and a bulleted list: "Help write", "Proofread", and "Summarize". A thick black arrow points from this box to the right side of the image. The right side shows a screenshot of the Google Docs editor for a document titled "Sales Representative Job Description". The toolbar at the top includes standard options like File, Edit, View, Insert, Format, Tools, Extensions, and Help. Below the toolbar is a set of text and style controls. In the bottom-left corner of the document area, there is a small blue button with a pencil icon and the text "Help me write".

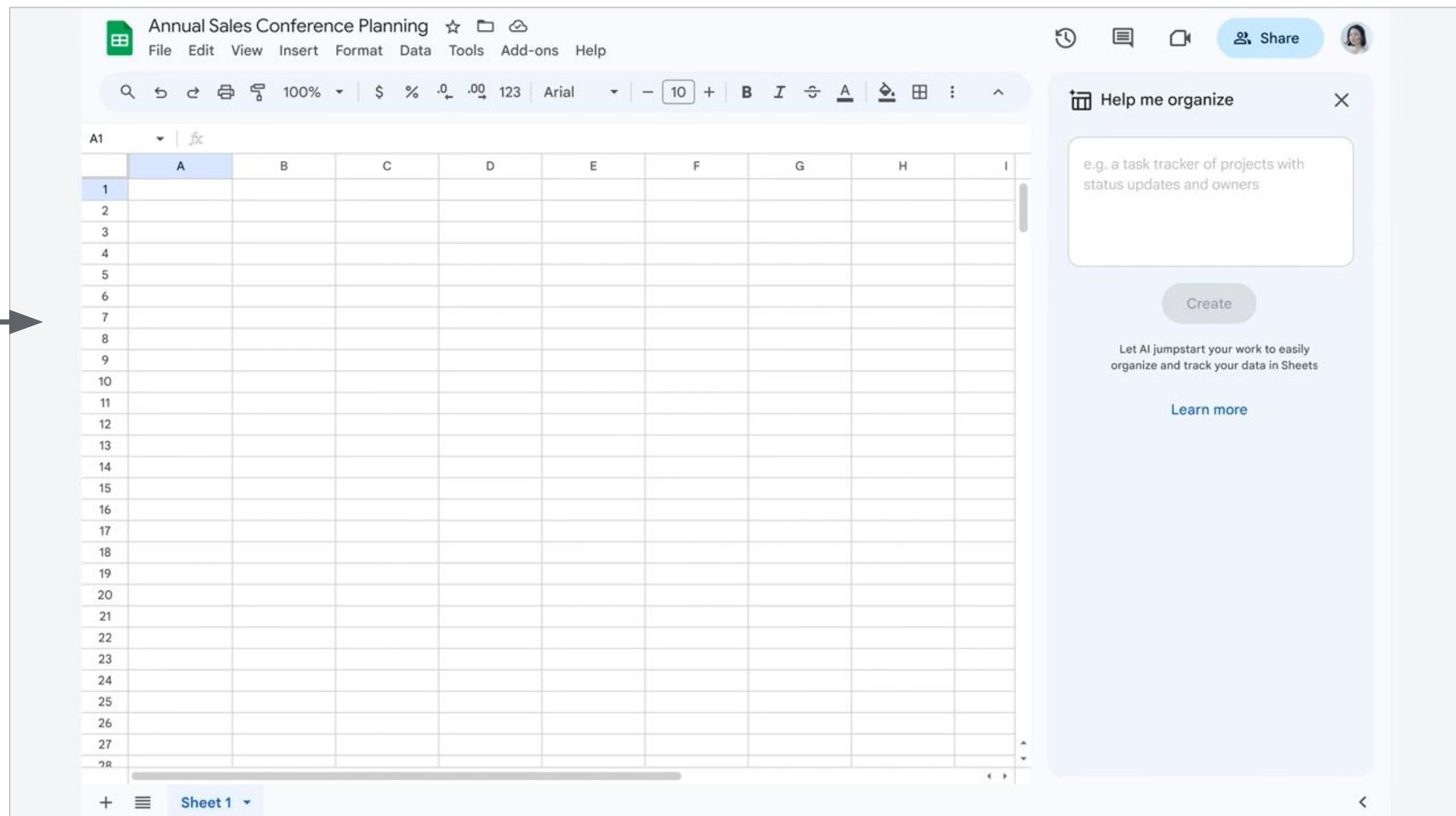
Google Sheets will provide features like “Help me organize”



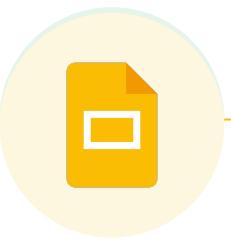
Sheets

Make data workflows easier, faster, and more approachable

- Content suggestions
- Contextual categorization
- Formula & chart generation



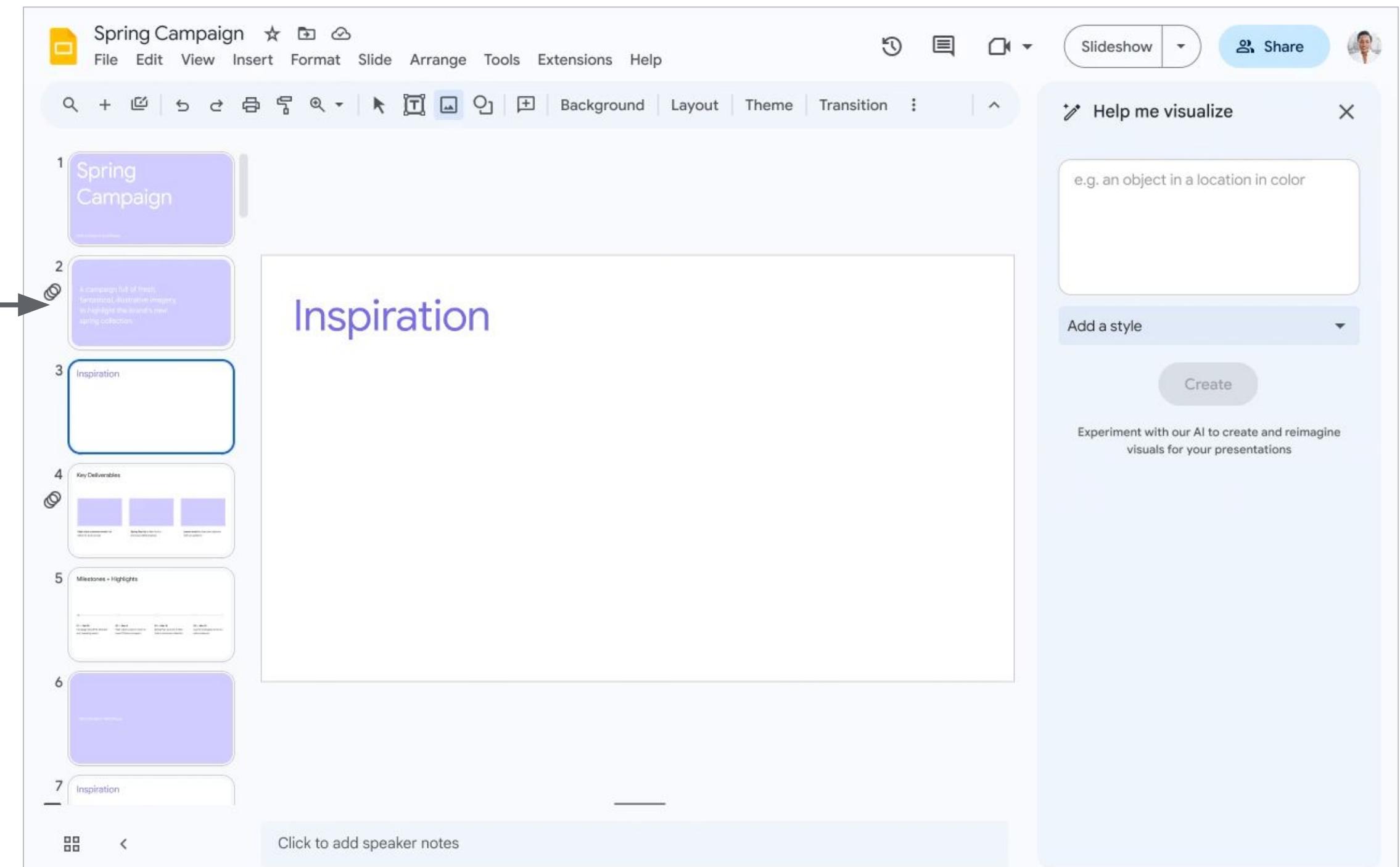
Google Slides will provide features like “Help me visualize”



Slides

Realize your creative vision easier and show up like a pro

- Generate images from text
- Rich content libraries
- Generate soundtracks

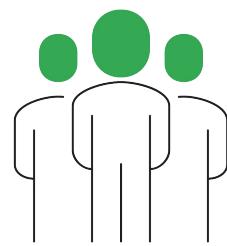


Topics

- 01 Duet AI for Google Workspace
- 02 Enterprise AI - APIs, and models
- 03 GenAI Studio
- 04 Model Garden
- 05 Gen App Builder
- 06 Responsible AI



Consumers & enterprises have different needs...



Consumers and enthusiasts

Bard, MakerSuite

- Help me plan a neighborhood block party
- Outline my blog post about summer mocktail recipes
- I want to write a novel. How do I get started?
- Draft a packing list for my weekend fishing and camping trip
- Give me a list of idioms for “let’s circle back” that aren’t cringe



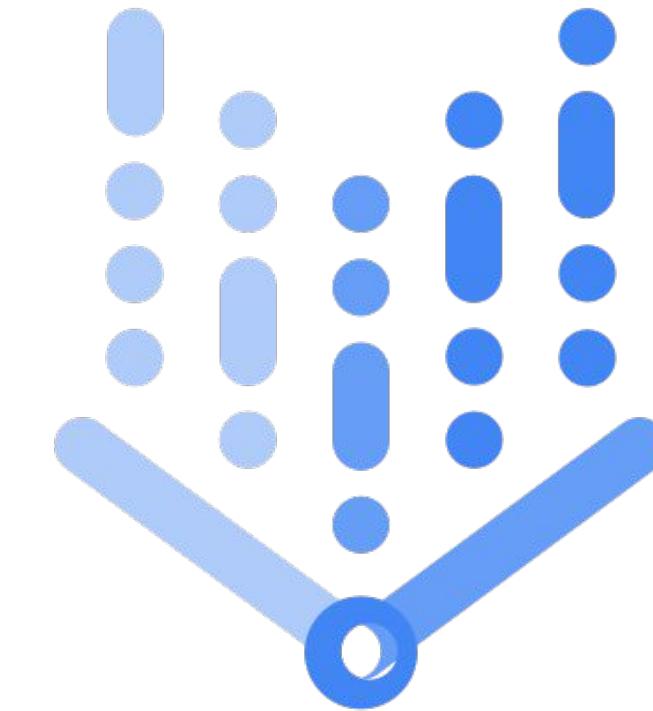
Enterprise

Vertex AI, GenApp Builder, AI APIs

- Allow data analysts to search and summarize market reports while **controlling our data**
- Help my customers understand my financial products while being **safe, explainable, and regulatory compliant**
- Is it easy to **integrate into our existing apps and platform?**
- Handle a customer service interaction with **accurate info**
- Can you guarantee we have **access to state-of-the-art models?**
- Can we generate content while **controlling costs?**

Vertex AI is a machine learning platform that helps you build, deploy, and manage ML models

- Combines data engineering, data science, and ML engineering workflows
- Provides several options for model training and deployment
 - AutoML
 - Custom Training
 - Model Garden
 - Generative AI
- Uses fully-managed infrastructure that you can customize based on your performance and budget needs
- Supports Python, the Google Cloud Console, the gcloud command line tool, client libraries, and Terraform



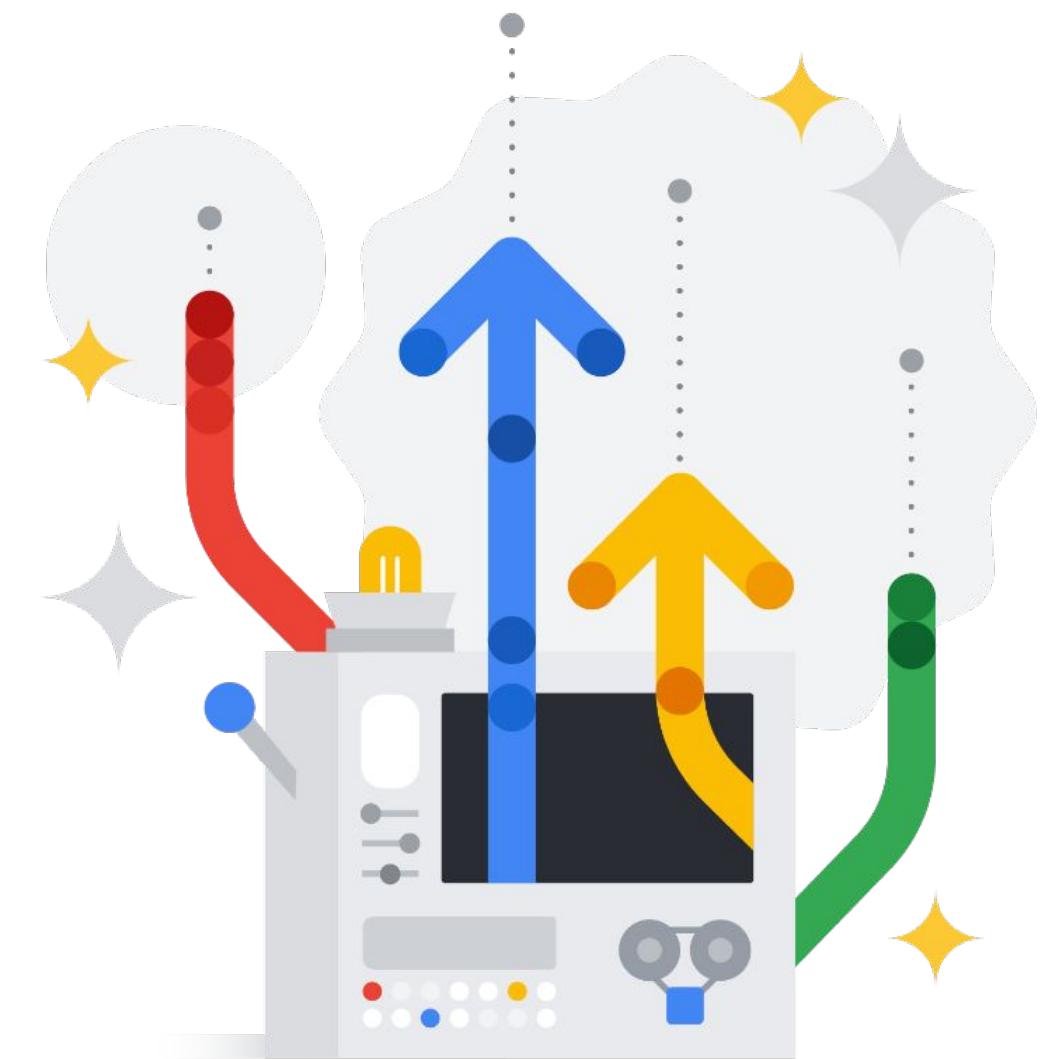
Developers access Google's Large Language Models via multiple APIs

- The **Vertex AI PaLM APIs** offers access to Google's Pathways Language Model 2 (PaLM 2) large language models
 - **PaLM API for text** is fine-tuned for tasks like classification, summarization, text generation, and entity extraction
 - **PaLM API for chat** is fine-tuned for multi-turn chat
 - The **Codey APIs** provide code generation, code completion, and code-related chat functionality
 - The **Text Embedding API** generates text embeddings that can be used for semantic search, recommendation, etc.
- Applications written using these APIs can choose between different versions of the underlying models



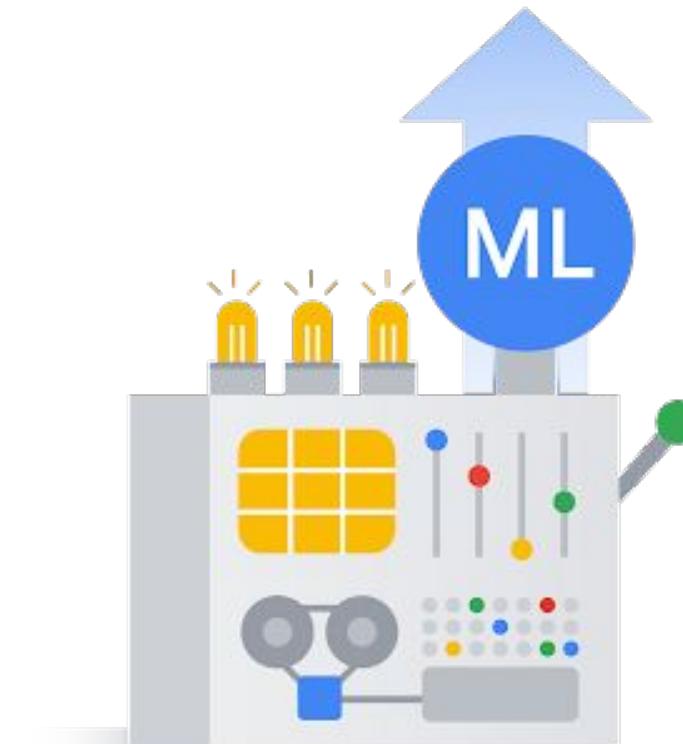
Google also offers multiple APIs for non-text Generative AI functionality

- The Vertex AI image APIs offer access to Google's large AI models for images
 - **Imagen** is a large AI model, specifically a text-to-image diffusion model, that can generate or edit images based on text prompts
 - The **Google Cloud Vision API** can generate embeddings which can be consumed by downstream tasks like image search or classification
- The Google Cloud Speech API offers access to Google's large language synthesis model - **Chirp**



Google refers to large-scale pre-trained models, which can later be tuned, as foundation models

- Foundation models, trained on extensive datasets, learn general patterns that can be applied across domains and tasks
- All the models previously described are considered foundation models by Google
- The APIs described earlier use these models as is, without fine tuning
- Some of these models, though, can be fine-tuned using Google Cloud tools
 - This tuning can improve performance and security
 - More on this soon!



Topics

01 Duet AI for Google Workspace

02 Enterprise AI - APIs, and models

03 GenAI Studio

04 Model Garden

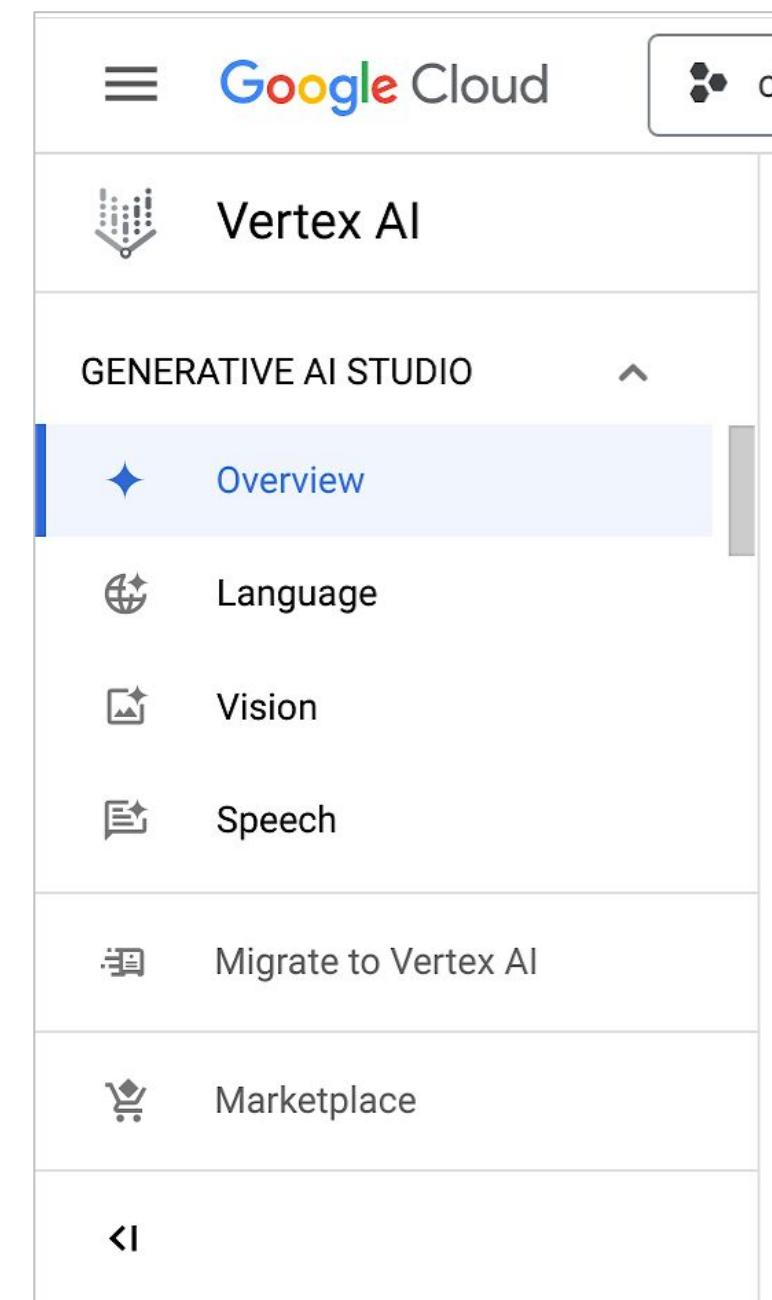
05 Gen App Builder

06 Responsible AI



Generative AI Studio simplifies prompting, tuning, and deploying Google foundation models

- Prompt design tools allow you to craft high-performing prompts
- To improve model performance, you can tune the model
- Generative AI studio offers tooling for three types of generation
 - Language
 - Vision
 - Speech



The prompt UI makes it easy to develop effective prompts

Context [?](#)

Multi-choice problem: Define the category of the ticket?
Categories:
- Credit card
- Bank account services
- Loans and Mortgages

Examples [?](#)

INPUT	OUTPUT	Category
I lost my credit card numbered 12345. Can you help with deactivating the card?	Credit card	X
I would like to change the address associated with my account. I have been calling the bank multiple times but couldn't get through. Please help me.	Bank account services	X
good morning my name is xxxx xxxx and i appreciate it if you could help me put a stop to chase bank cardmember services. I wrote to chase asking for debt verification and what they sent me a statement which is not acceptable i am asking the bank to validate the debt instead i been receiving mail every month from them attempting to collect a debt i have a right to know this information as a consumer chase account xxxx xxxx xxxx xxxx thanks in advance for your help	Loans and Mortgages	X
+ Write an input to add a new row	+ Write an output to add a new row	

Test [?](#)

INPUT	OUTPUT	Category
my grand son give me check for i deposit it into my chase account after fund clear my chase bank closed my account never paid me my money they said they need to speak with my grand son check was clear money was taking by my chase bank refuse to pay me my money my grand son called chase times they told him i should call not him to verify the check owner he is out the country most the time date happen check number xxxx claim number is xxxx with chase	Model responses will appear here	- Markdown

We want your [feedback](#).

Model [text-bison@001](#)

Temperature [?](#) 0.2

Token limit [?](#) 256

Top-K [?](#) 40

Top-P [?](#) 0.8

Safety filter threshold [?](#) Block few

SUBMIT **RESET PARAMETERS**

Google provides a library of prompt examples

Extraction

Structured

Pixel Technical Specification

Generate technical specification from text of a Pixel phone into JSON, one-shot.

[OPEN](#)

Freeform

Wifi troubleshooting

Given description of the different status lights on the Google WiFi router, what should be the troubleshooting step.

[OPEN](#)

Freeform

Contract analysis

You are a partner of a law firm. Your associates are bored of reading contracts to find specific provisions when they can work on more...

[OPEN](#)

Structured

Extractive Question Answering

Answer questions from given background texts.

[OPEN](#)

Writing

Freeform

Marketing generation Pixel

You work in Google's device marketing team and you need to create marketing pitch for the new Pixel 7 Pro. You have writers block...

[OPEN](#)

Freeform

Ad copy generation

You are a marketer and want to create different versions of the same ad to target different audiences. You would like some suggestions.

[OPEN](#)

Freeform

Essay outline

Generate an outline for an essay on a particular topic.

[OPEN](#)

Freeform

Correct grammar

Correct grammar in the text.

[OPEN](#)

Vision tools allow you to create images or image-related text content

The screenshot displays the Vertex AI Vision interface. On the left, a sidebar lists various tools: Dashboard, Model Garden, Workbench, Pipelines, Overview, Language, **Vision** (which is selected), Speech, Feature Store, Datasets, Migrate to Vertex AI, and Marketplace. The main area shows a grid of eight generated images. Below the images is a text input field containing "Material design style icon for video conferencing appliance". The top right features buttons for EXPORT, RESET, and HISTORY. The right side contains a Parameters panel with fields for Model options (set to imagegeneration@001), Images (Aspect ratio: Square, Number of results: 8), Negative prompt (Define what you don't want to see), ADVANCED OPTIONS (Enable fine-tuned model), and Model Properties (checkbox for Enable fine-tuned model). At the bottom are buttons for GENERATE, EDIT, CAPTION, and VISUAL Q & A.

Vertex AI

Vision

EXPORT RESET HISTORY

Model options: imagegeneration@001

Images

Aspect ratio: Square

The image resolution is 1024 by default

Number of results *: 8

Negative prompt

Define what you don't want to see

ADVANCED OPTIONS

Model Properties

Enable fine-tuned model

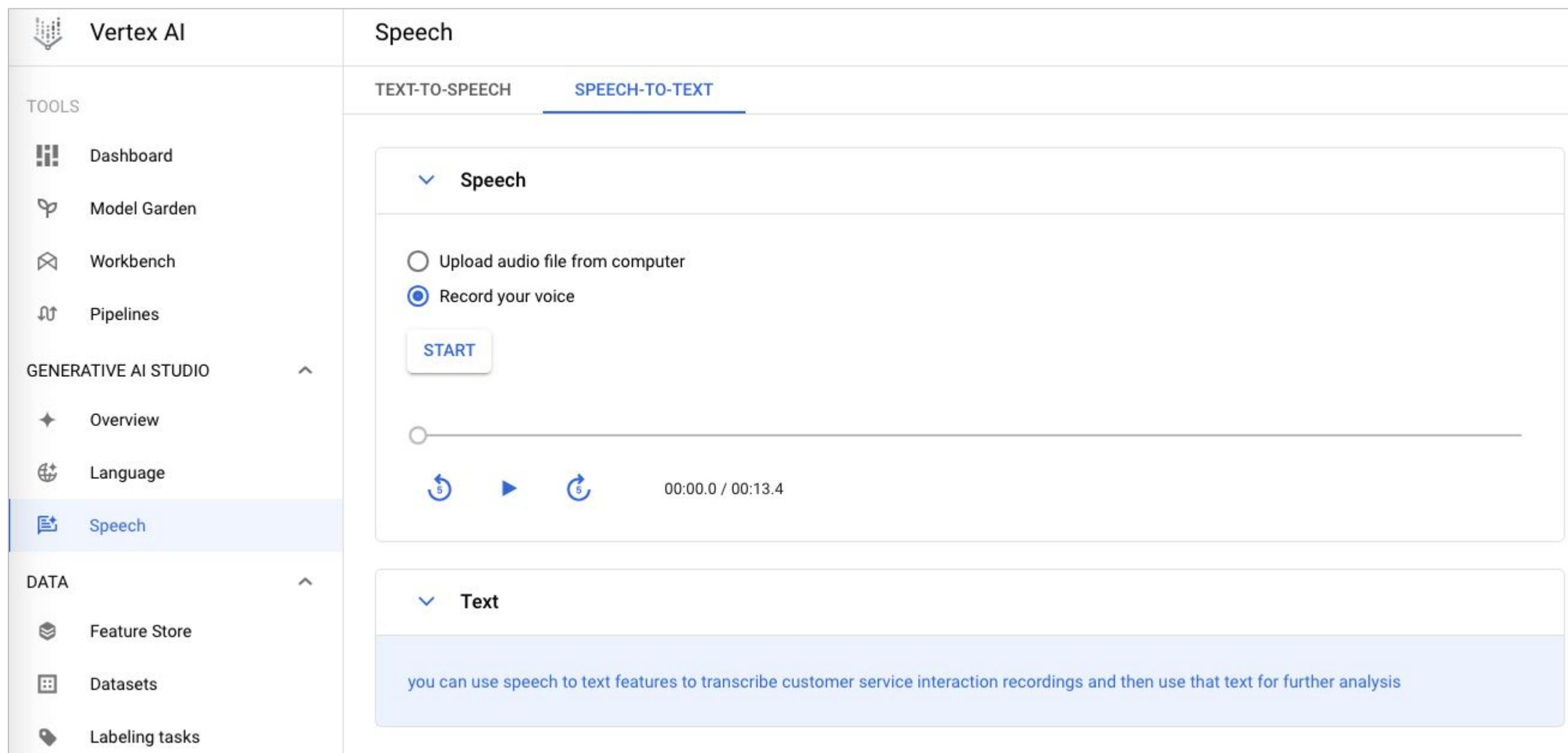
Customize the model with your own custom subjects

GENERATE

EDIT CAPTION VISUAL Q & A

Material design style icon for video conferencing appliance

Speech tools allow users to easily convert text to speech, or speech-to-text



GenAI Studio allows you to train and deploy fine-tuned (customized) models

- Allows you to teach the model more about what your expected output should be
- You specify a custom dataset which includes prompts along with the expected output
- The result of tuning a model is a set of small add-in layers that is used in conjunction with a foundation model
 - The adapters are what enable the foundation model to perform better for certain tasks and domains



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 Duet AI for Google Workspace
- 02 Enterprise AI - APIs, and models
- 03 GenAI Studio
- 04 Model Garden
- 05 Gen App Builder
- 06 Responsible AI



Model Garden helps you discover, customize, and deploy models from many sources

The screenshot shows the Vertex AI Model Garden interface. On the left, there's a sidebar with sections like TOOLS (Dashboard, Model Garden, Workbench, Pipelines), GENERATIVE AI STUDIO (Overview, Language, Vision, Speech), DATA (Feature Store, Datasets, Labeling tasks), and MODEL DEVELOPMENT (Training, Experiments). The main area is titled "Model Garden" with "EXPLORER GENERATIVE AI" and "VIEW MY MODELS" buttons. It features a search bar and a modalities section with categories: Language (23), Vision (46), Tabular (2), Documents (2), and Speech (1). Below these are tasks: Generation (22), Classification (26), Detection (17), Extraction (7), Recognition (6), Translation (2), Embedding (2), Segmentation (3), and Features (9). A central grid displays various models with their names, descriptions, and "VIEW DETAILS" buttons. The models shown include:

Model	Description	Type
Codey for Code Completion	Generates code based on code prompts. Good for code suggestions and minimizing bugs in code.	Foundation, Language
Codey for Code Generation	Generates code based on natural language input. Good for writing functions, classes, unit tests, and more.	Foundation, Language
Codey for Code Chat	Get code-related assistance through natural conversation. Good for questions about an API, syntax in a supported language, and more.	Foundation, Language
Stable Diffusion v1-5	Latent text-to-image diffusion model capable of generating photo-realistic images given a text input.	Foundation, Vision
BERT	Neural network-based technique for natural language processing. Use it to train your own question answering system and more.	Foundation, Language
InstructPix2Pix	Given an input image and a text prompt that tells the model what to do, the instruct-pix2pix model follows the prompt to edit the image by generating a new one.	Foundation, Language, Vision

Model Garden helps you discover, customize, and deploy models from many sources

PaLM 2 for Text

Fine-tuned to follow natural language instructions and is suitable for a variety of language tasks, such as: classification, extraction, summarization and content generation.

[OPEN PROMPT DESIGN](#) [VIEW API CODE](#)

[OVERVIEW](#) [USE CASES](#) [DOCUMENTATION](#)

Overview

text-bison is the name of the PaLM 2 for text large language model that understands and generates language. It's a foundation model that performs well at a variety of natural language tasks such as sentiment analysis, entity extraction, and content creation. The type of content that **text-bison** can create includes document summaries, answers to questions, and labels that classify content.

The PaLM 2 for text is ideal for tasks that can be completed with one API response, without the need for continuous conversation. For text tasks that require back-and-forth interactions, use the [PaLM 2 for chat](#).

Use cases

- **Summarization:** Create a shorter version of a document that incorporates pertinent information from the original text. For example, you might want to summarize a chapter from a textbook. Or, you could create a succinct product description from a long paragraph that describes the product in detail.

Stable Diffusion v1-5

Latent text-to-image diffusion model capable of generating photo-realistic images given a text input.

[DEPLOY](#) [OPEN NOTEBOOK](#)

Save to Model Registry and deploy to an endpoint

[OVERVIEW](#) [USE CASES](#) [DOCUMENTATION](#) [PRICING](#)

Resource ID	text-bison@001
Tags	Classification, Detection, Extraction, Generation, Recognition, Translation

Topics

- 01 Duet AI for Google Workspace
- 02 Enterprise AI - APIs, and models
- 03 GenAI Studio
- 04 Model Garden
- 05 Gen App Builder
- 06 Responsible AI



Gen App Builder let's developers and power users build no-code apps that tap into Google's foundation models



Enterprise Search

1. Quickly build custom search engines on your proprietary data
2. Deliver natural language understanding and semantic search
3. Automatically handle synonyms & misspellings
4. Include AI-generated summaries and enable conversational search
5. Incorporate search across structured and unstructured data



Infobot

1. Easily create virtual agents that use GenAI
2. Agents can answer questions and have conversations with users based on information you provide
3. Infobot applications are built on top of Dialogflow CX technologies, but are easy to bootstrap
4. Infobot features are not currently available to the public

Solutions typically require more than one tool

- Each tool, in itself, provides valuable features
- However, most enterprise-grade solutions will require careful integration of multiple products and processes
- Architects and developers will need to...
 - Break problems down
 - Identify the tools and processes required for each phase
 - Integrate the pieces
 - Implement operational best practices to make sure everything runs as expected



GenAI-powered content discovery

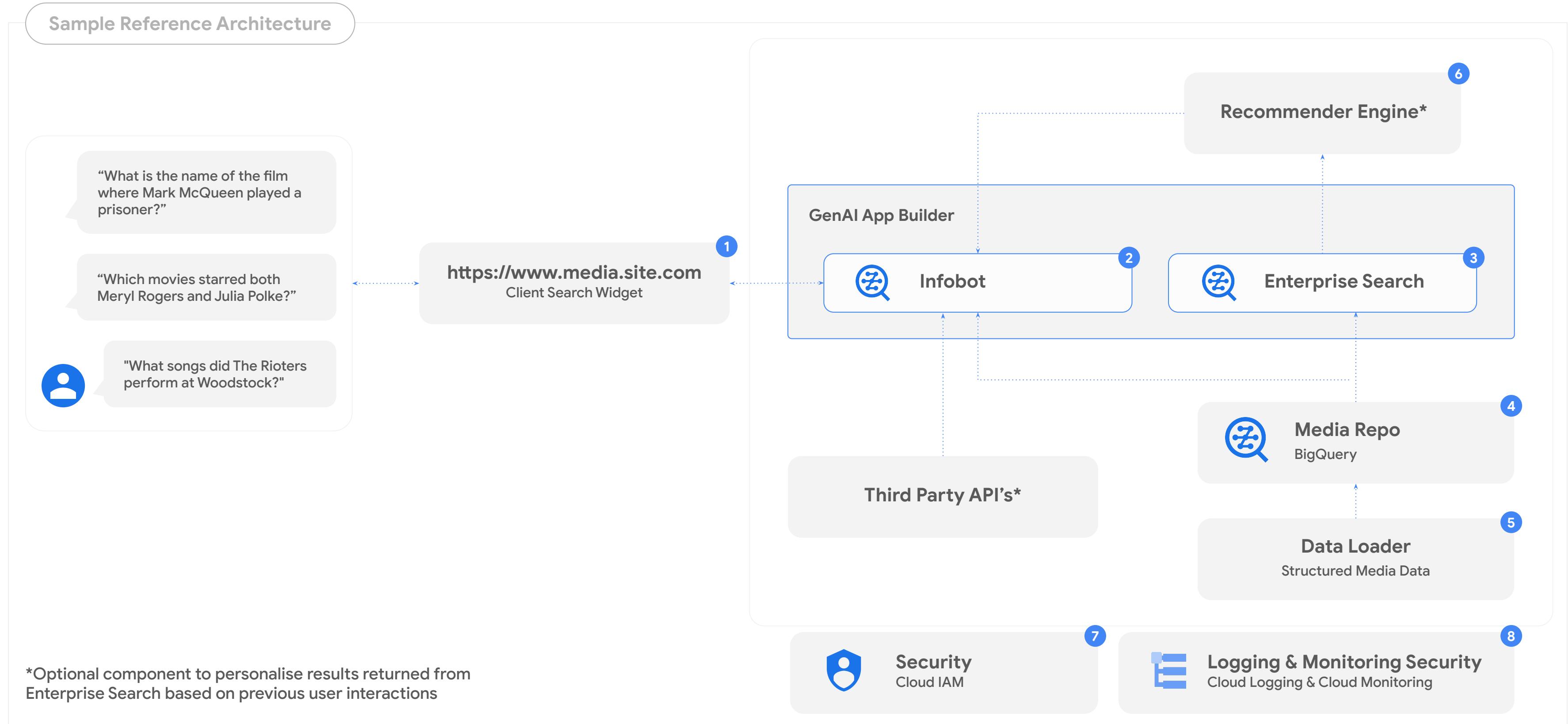


Build personalized content discovery experiences with generative AI



Google Cloud

GenAI-powered content discovery



Topics

- 01 Duet AI for Google Workspace
- 02 Enterprise AI - APIs, and models
- 03 GenAI Studio
- 04 Model Garden
- 05 Gen App Builder
- 06 Responsible AI

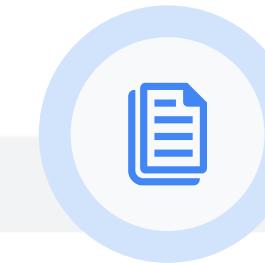


Google strives to develop AI solutions responsibly



Principles

1. Be socially beneficial
2. Avoid creating or reinforcing bias
3. Be built and tested for safety
4. Be accountable to people
5. Incorporate privacy design principles
6. Uphold high standards for scientific excellent
7. Be made available for uses that accord with these principles

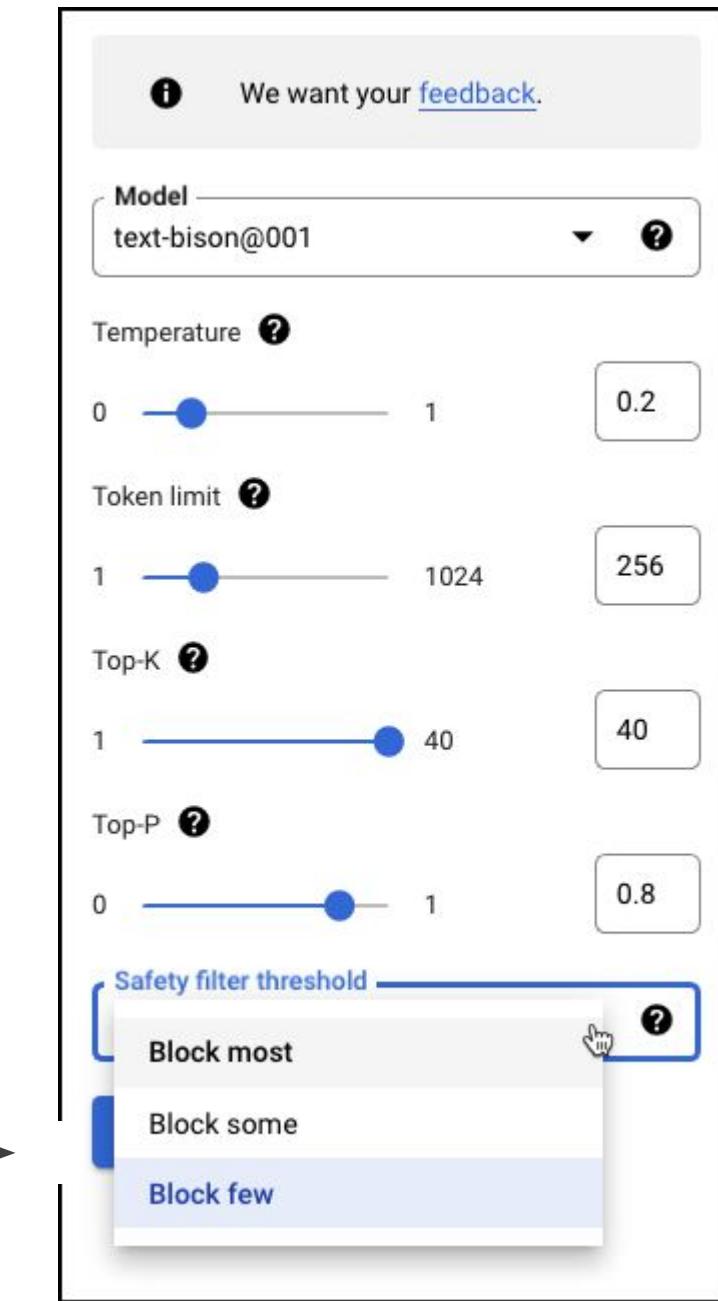
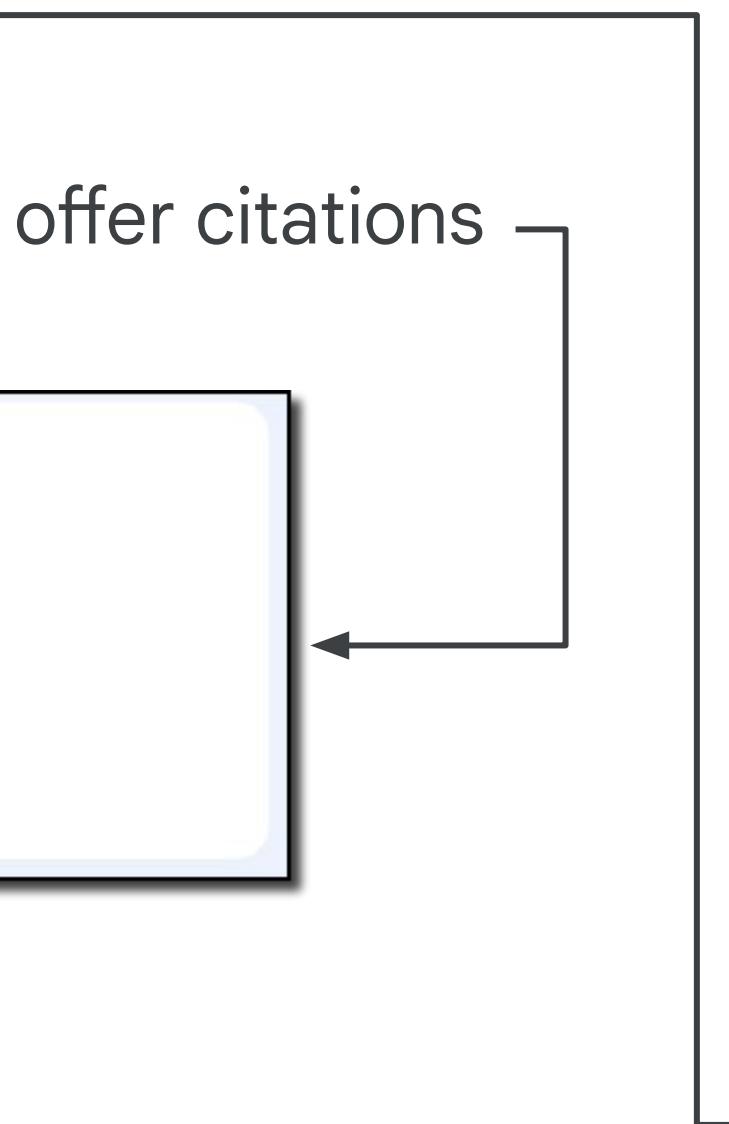
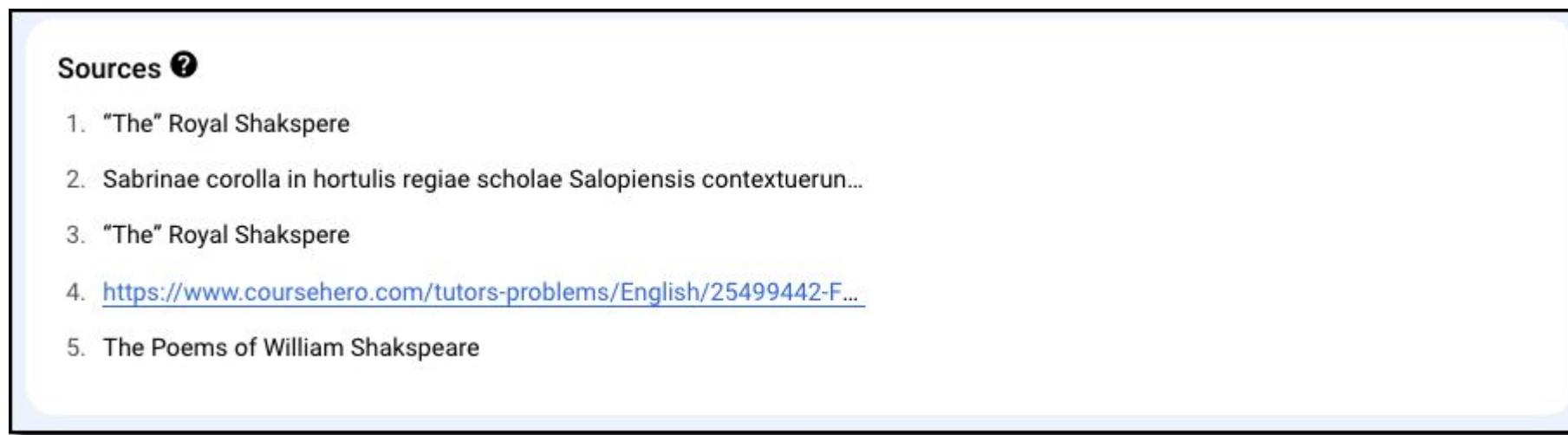


Practice

1. Examine your raw data and test (General)
2. Check the system for unfair bias (Fairness)
3. Communicate explanations to users (Interpretability)
4. Collect and handle data responsibly (Privacy)
5. Develop an approach to combat threats (Safety)

Generative AI Studio includes responsibility features

- Generative AI Studio actively filters responses
- Generative AI Studio and the PaLM API both can offer citations



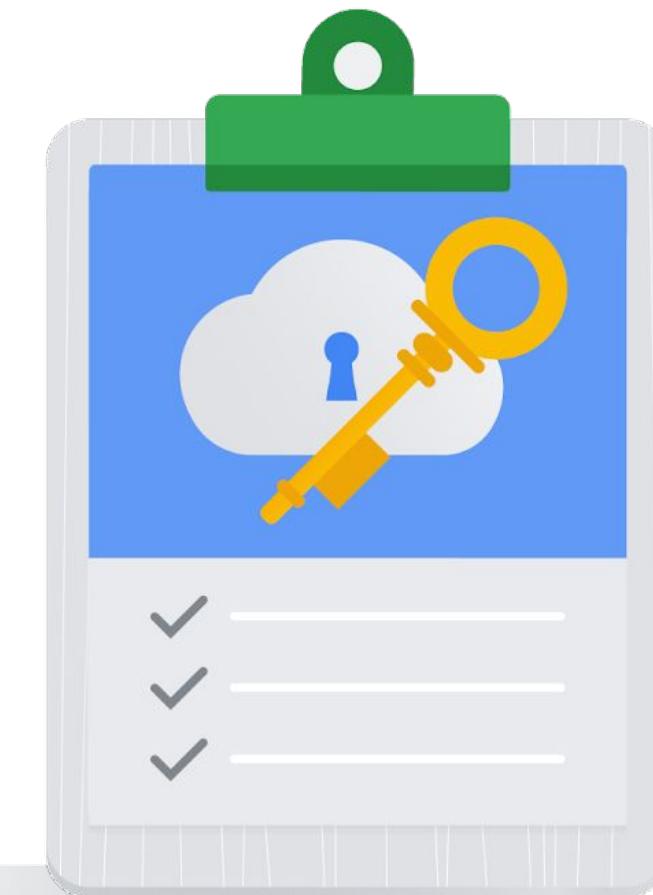
Vertex AI PaLM API has additional safety features

- Developers can filter responses by looking at safety feature ratings

```
{  
  "predictions": [  
    {  
      "safetyAttributes": {  
        "categories": [  
          "Derogatory",  
          "Toxic",  
          "Violent",  
          "Sexual",  
          "Insult",  
          "Obscene",  
          "Death, Harm & Tragedy",  
          "Firearms & Weapons",  
          "Public Safety",  
          "Health",  
          "Religion & Belief",  
          "Drugs",  
          "War & Conflict",  
          "Politics",  
          "Finance",  
          "Legal"  
        ],  
        "scores": [  
          0.1,  
          0.1,  
          0.1,  
          0.1,  
          0.1,  
          0.1,  
          0.1,  
          0.1,  
          0.1  
        ]  
      }  
    }  
  ]  
}
```

Google implements security and privacy features throughout the AI operations flow

- As noted earlier, customer data such as prompts and training data aren't logged or used for training a foundation model
- If the customer does fine-tuning, adapter layers definition and weights are owned by the customer and stored securely
- Customers can even use Customer Managed Encryption Keys to secure the adapter data, and secure access to the data using VPC service controls
- Googler access to this data requires a business justification and is logged in Access Transparency logs.



Customers also have responsibilities when using GenAI

01

Avoid hallucinations

Models can generate plausible but incorrect content. You can reduce a model's "creativity" with carefully crafted prompts and setting the parameters of requests

02

Fact check and test

Users should always check factual assertions provided by GenAI tools. Organizations should have standards and processes for testing results over time.

03

Use GenAI as a starting point

In some cases, using content generated entirely by AI is inappropriate. In many cases, the results can be improved by human fine-tuning.

04

Be careful with prompts

Avoid inclusion of sensitive information. Also, avoid using brand or celebrity names in prompts as these might result in answers that are subject to infringement allegations.

05

Keep records

In situations where you might need to demonstrate the source of your work, keeping track of your prompts, results, and final products can be important.

06

Disclose AI use

In many cases, it's a good idea to acknowledge when content has been generated by an AI tool.

When building solutions, put guardrails in place

- Document and promote user best practices
- Don't assume the LLM safety controls provide guarantee appropriate levels of safety
 - Different use cases have different requirements
 - Leverage Palm2 in built safety scores to validate outputs
 - Use other mechanisms to vet the behavior of GenAI-backed solutions
- Have developers tune prompts used in apps to reduce hallucinations
 - Example: “You are truthful and never lie. Never make up facts and if you are not 100% sure, reply with why you can not answer in a truthful way.”
- Have apps use fallback protocols when receiving a non-answer
- Validate and verify specific data as part of your application logic
- Use Reinforcement Learning from Human Feedback (RLHF) to improve model accuracy



GenAI in Your Organization

In this module, you learn to ...

- 01 Identify different types of GenAI consumers in your organization
- 02 Map Google's AI tools to these different audiences
- 03 Begin developing a process for identifying priority use cases across your company



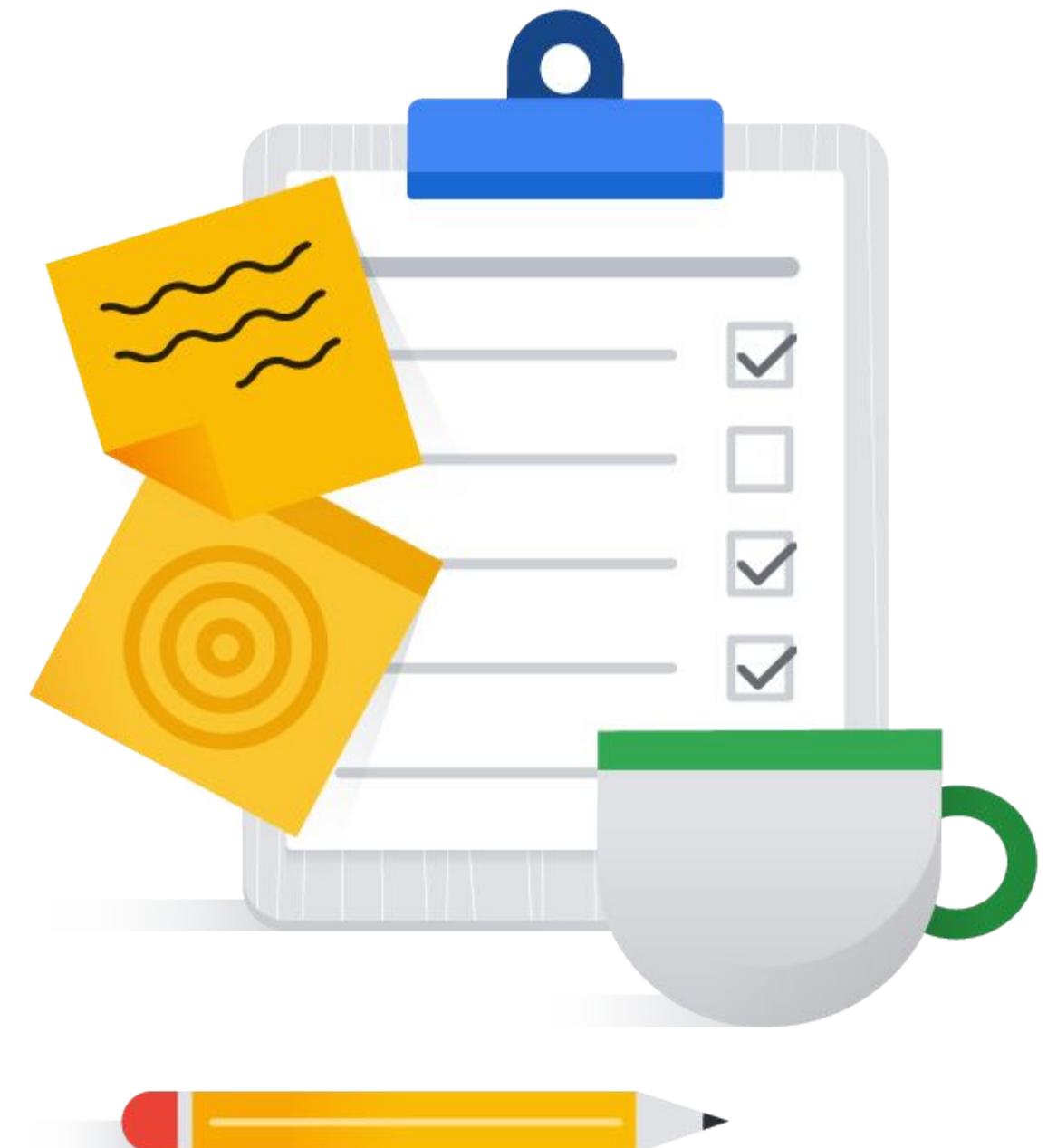
Topics

01

Personas and products

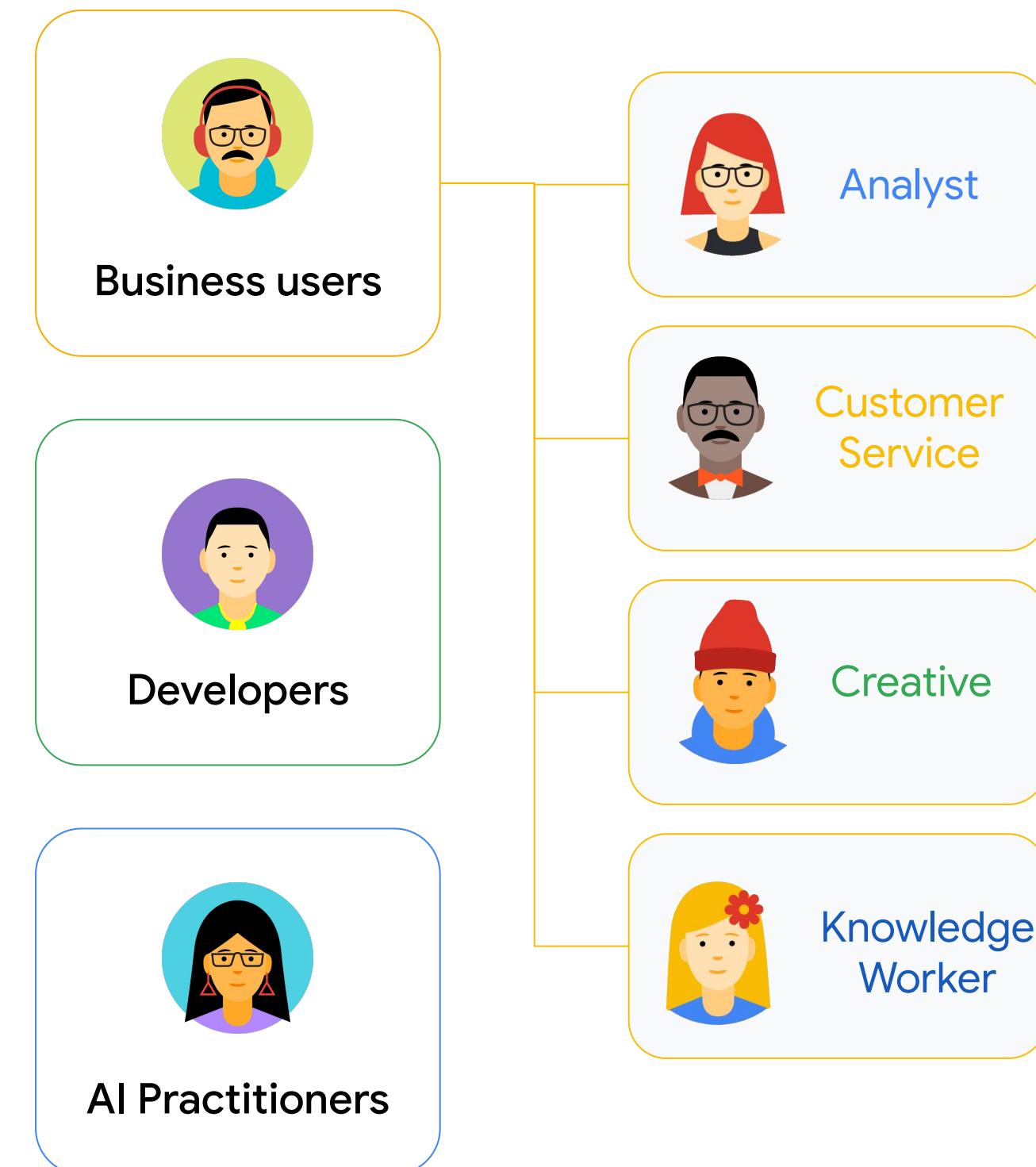
02

Finding target use cases

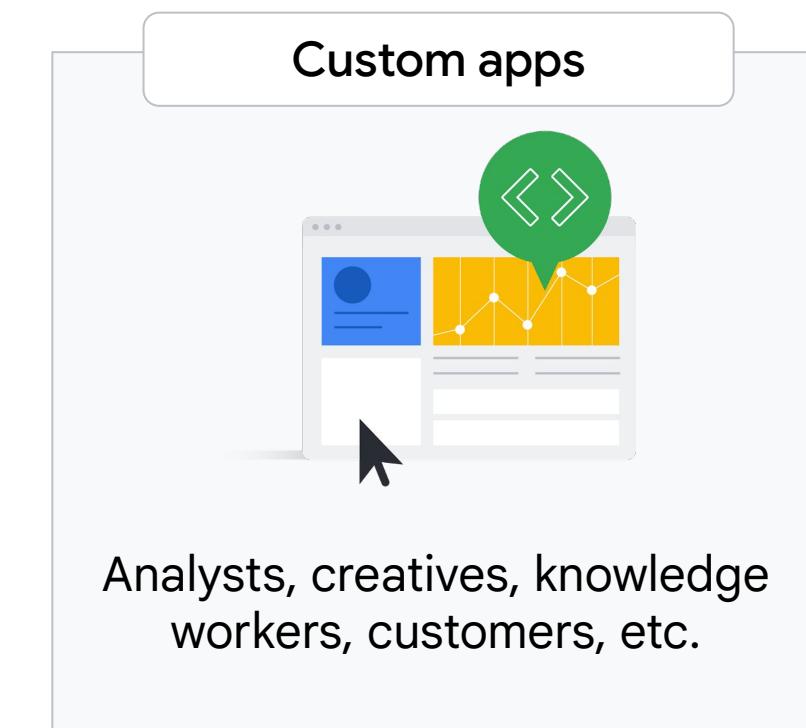
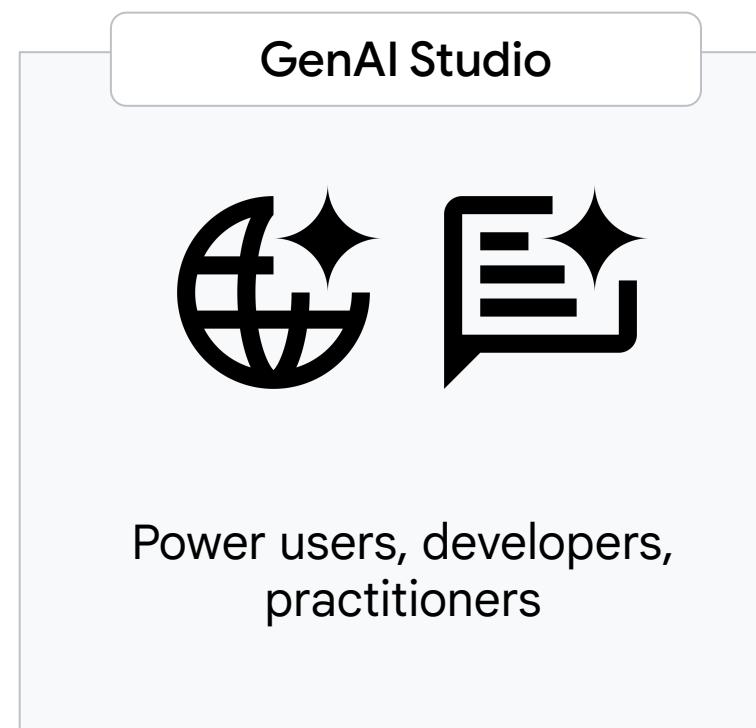
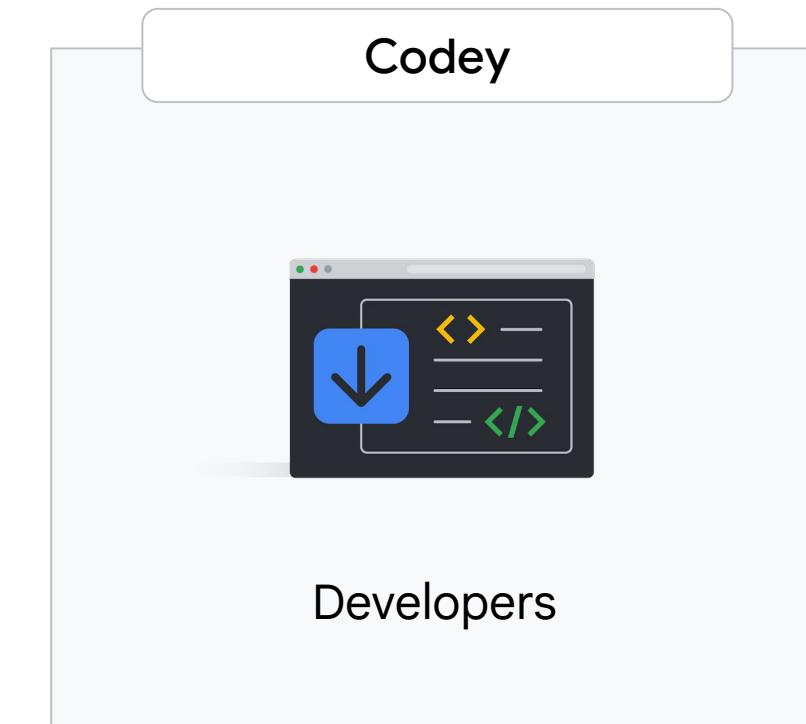
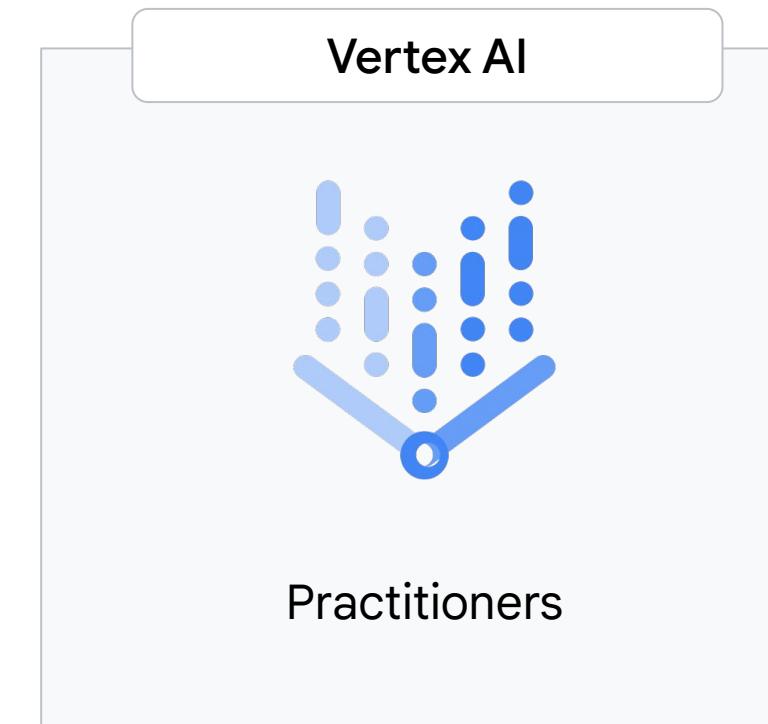
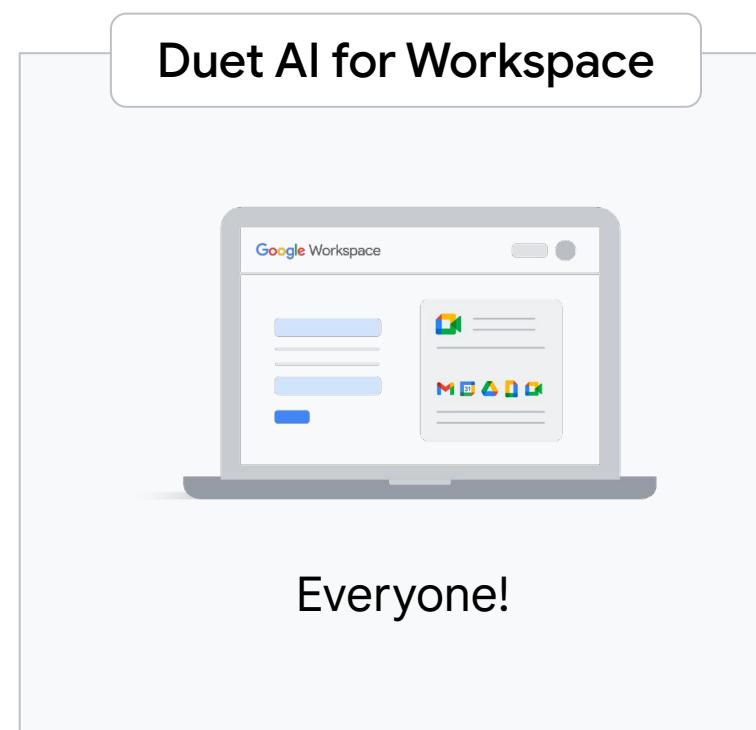


Within any enterprise, there are different audiences doing different things with GenAI

- Business end users will use GenAI-enabled tools to perform their jobs with greater efficiency and better outcomes
- Developers and architects will be responsible for building applications that deliver GenAI features in solutions that are optimized for the business
- AI practitioners will be those who help create and deploy business-specific models, as well as implement responsible AI practices



Different personas will use different GenAI tools



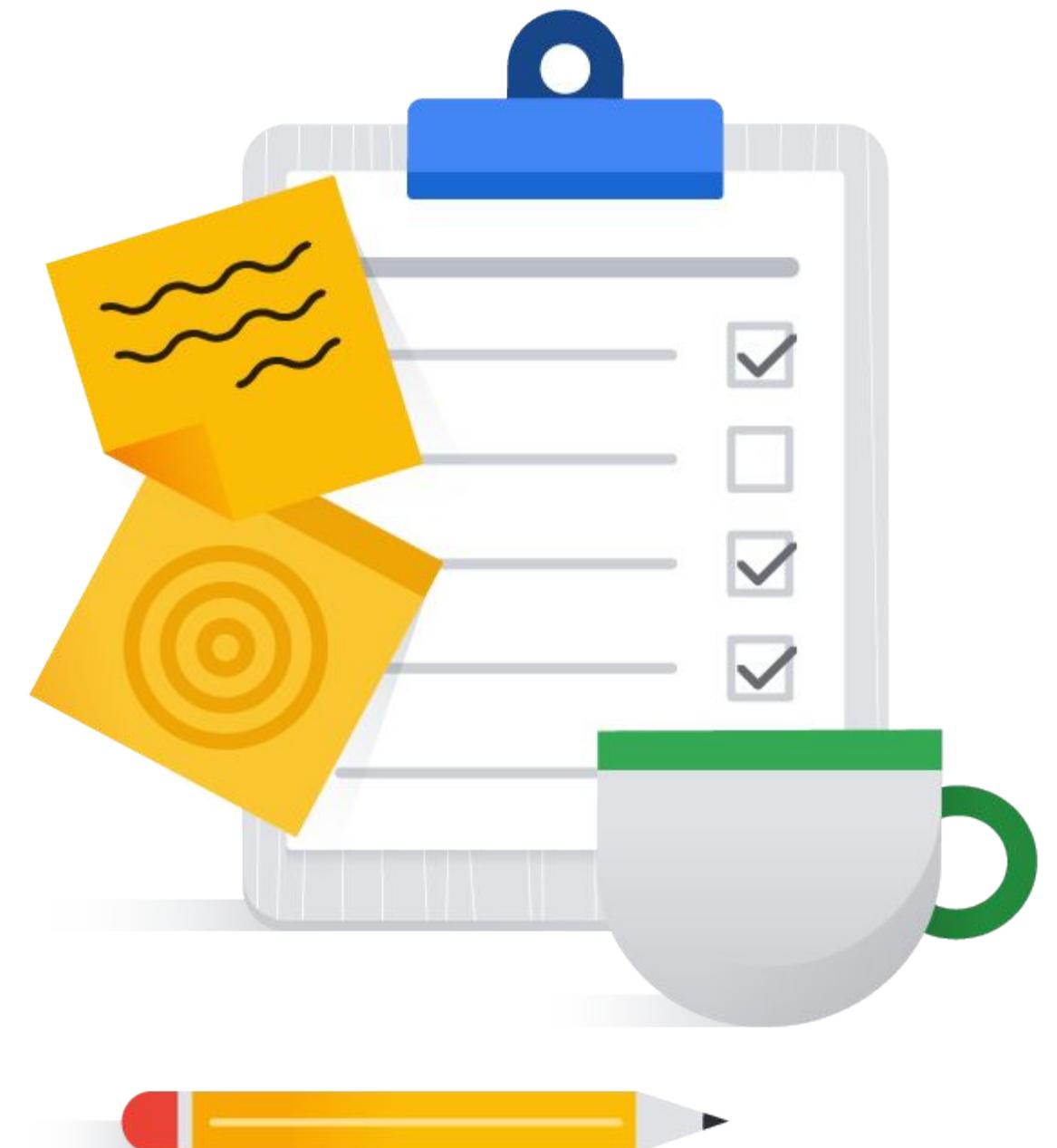
Topics

01

Personas and products

02

Finding target use cases



Start with the solution classes discussed earlier



CREATE

Bring your thoughts and visions to life

Use cases

- Generate product descriptions from images
- Generate images from text



SUMMARIZE

Condense and summarize your knowledge base into a simple format

Use cases

- Content/video summarization
- Intra-knowledge Q&A



DISCOVER

Help your customers and employees find what they need at the right time

Use cases

- Search for a document
- Machine-generated event monitoring



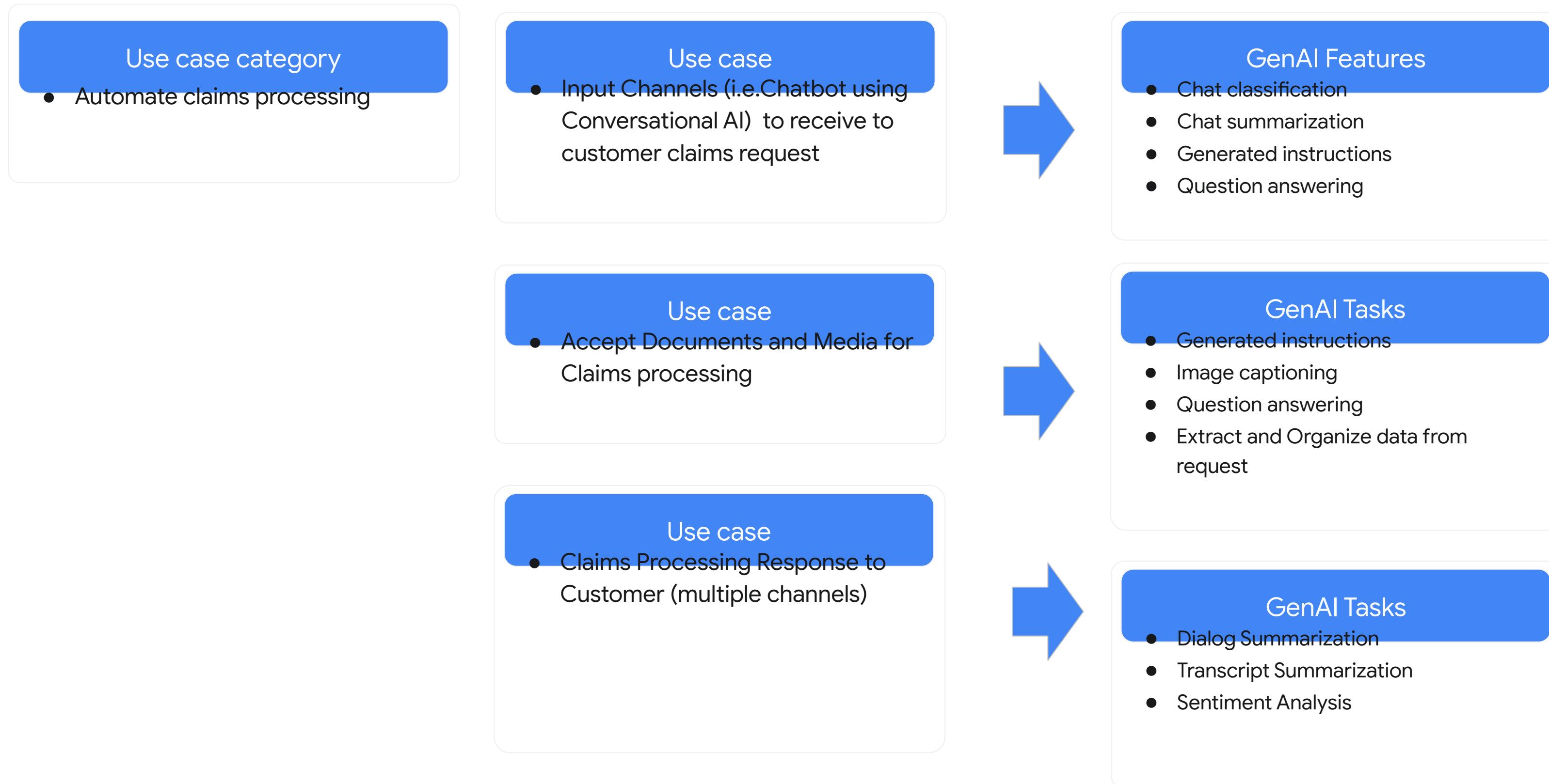
AUTOMATE

Automate your customer service across multiple channels

Use cases

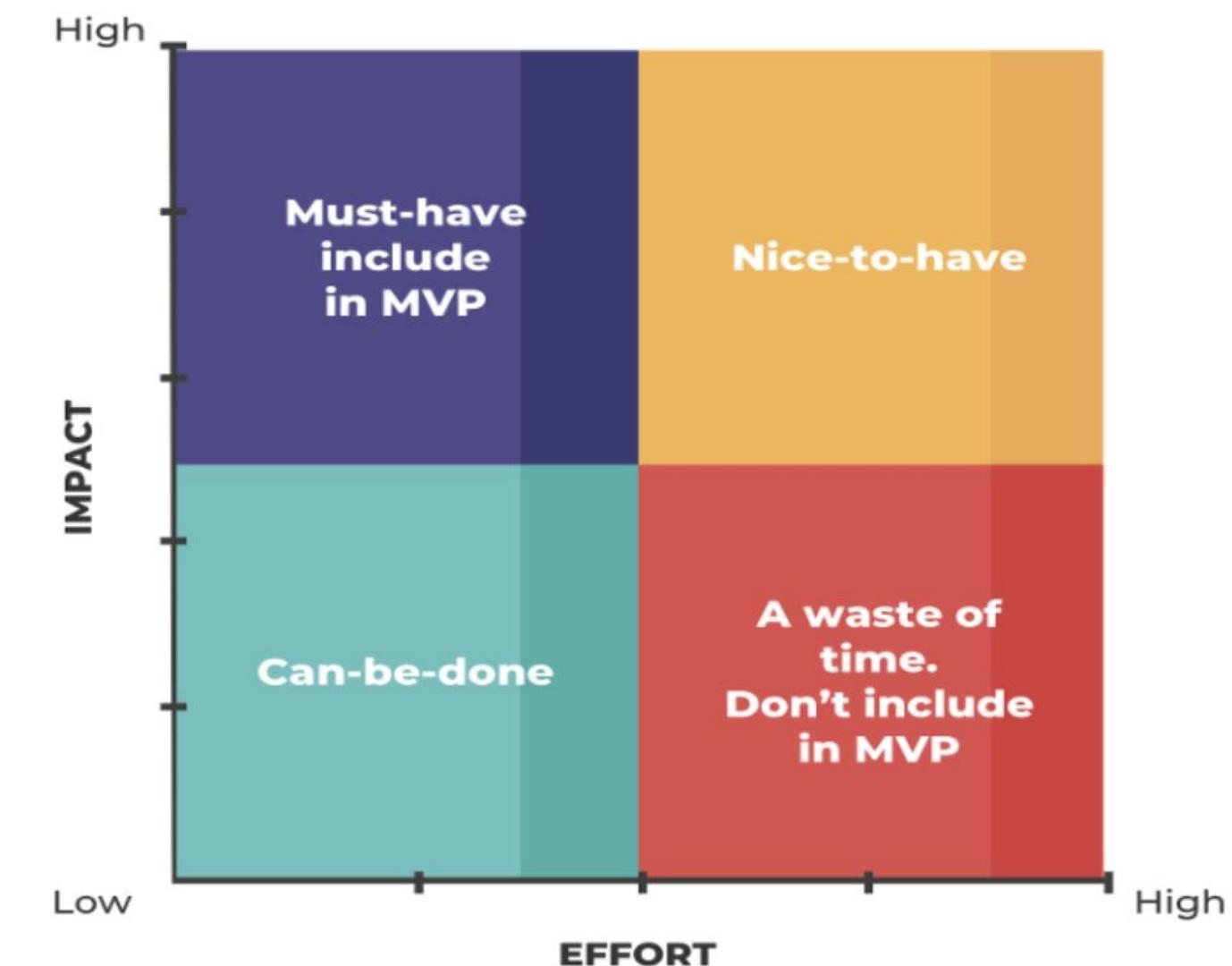
- Contract information extraction
- Feedback classification and ticket creation

Category > Use case > GenAI features



Identify priorities based on impact and effort

1. **Business Impact:** What GenAI applications would have the highest impact on a given business area?
 - a. Complex search
 - b. Infobots
 - c. Content generation
2. **Data Availability:** What data would be available?
 - a. Sample prompts and responses
 - b. Unstructured data sources
3. **Complexity:** What are the key challenges? (business/technical requirements, limitations, risks)
4. **Expectation:** How would you define your success criteria?



A standard worksheet can help with prioritization

Criteria	Description	Score
Organization's interest/ readiness to test (flexibility to test a nascent api)	What is the level of relevance of Google GenAI capabilities to customer use cases. Willingness to engage in Trust Tester Program (TTP)?	
Product availability	Are there detailed APIs, or other GenAI product assets available within the required timeline?	
Product maturity	Does the product have the requisite level of maturity?	
Product depth	What is the level of detail and depth that the product has that is currently available for use?	
Tokens generated	Are there any challenges with token limitations?	
Input/Output token comparison / ratio	What is the ratio of tokens used for input prompts and output response? Any challenges or impact on prompt design, prompt tuning, quality of response?	
Customization requirements	Can we solve this with OOTB components, foundational models as is? Or requires detailed customization?	

In this module, you learned to ...

- 01 Identify different types of GenAI consumers in your organization
- 02 Map Google's AI tools to these different audiences
- 03 Begin developing a process for identifying priority use cases across your company





Q&A and Next Steps

Topics

01	Next Steps
02	Q&A



Take a structured approach to GenAI transformation



Ramp your skills

Continue your learning journey and complete our [Generative AI Learning Path](#), free of charge, on Google Cloud Skills Boost



Organizational readiness

Assess your organization's current status and business needs for generative & traditional AI capabilities



Identify use cases

Select from one of our [Jumpstart GenAI offers](#), and work with Google Cloud to develop a technical design doc and sample code to solve the use case



Test and scale

Purchase and implement generative AI solutions. Not all AI is built equal. POC often and fail fast to identify what works for your business.

Contact your Google Cloud Representative to learn more

4 GenAI Jumpstart offers - \$25k & 2 weeks per use case



CREATE

Bring your thoughts and visions to life



SUMMARIZE

Condense and summarize your knowledge base into a simple format



DISCOVER

Help your customers and employees find what they need at the right time



AUTOMATE

Automate your customer service across multiple channels

Use cases:

- Images from text
- Product descriptions from images
- Blog post from content*
- Email from content*
- Release notes from content*
- Report from content*
- Press releases from content*
- Personalized ads*

Use cases:

- Content/video summarization
- Intra-knowledge Q&A
- Explanations of code content*
- External chatbot using internal data*
- External chatbot using website data*

Use cases:

- Search for a document
- Machine-generated event monitoring
- File organization based on content*
- Exam questions from content*

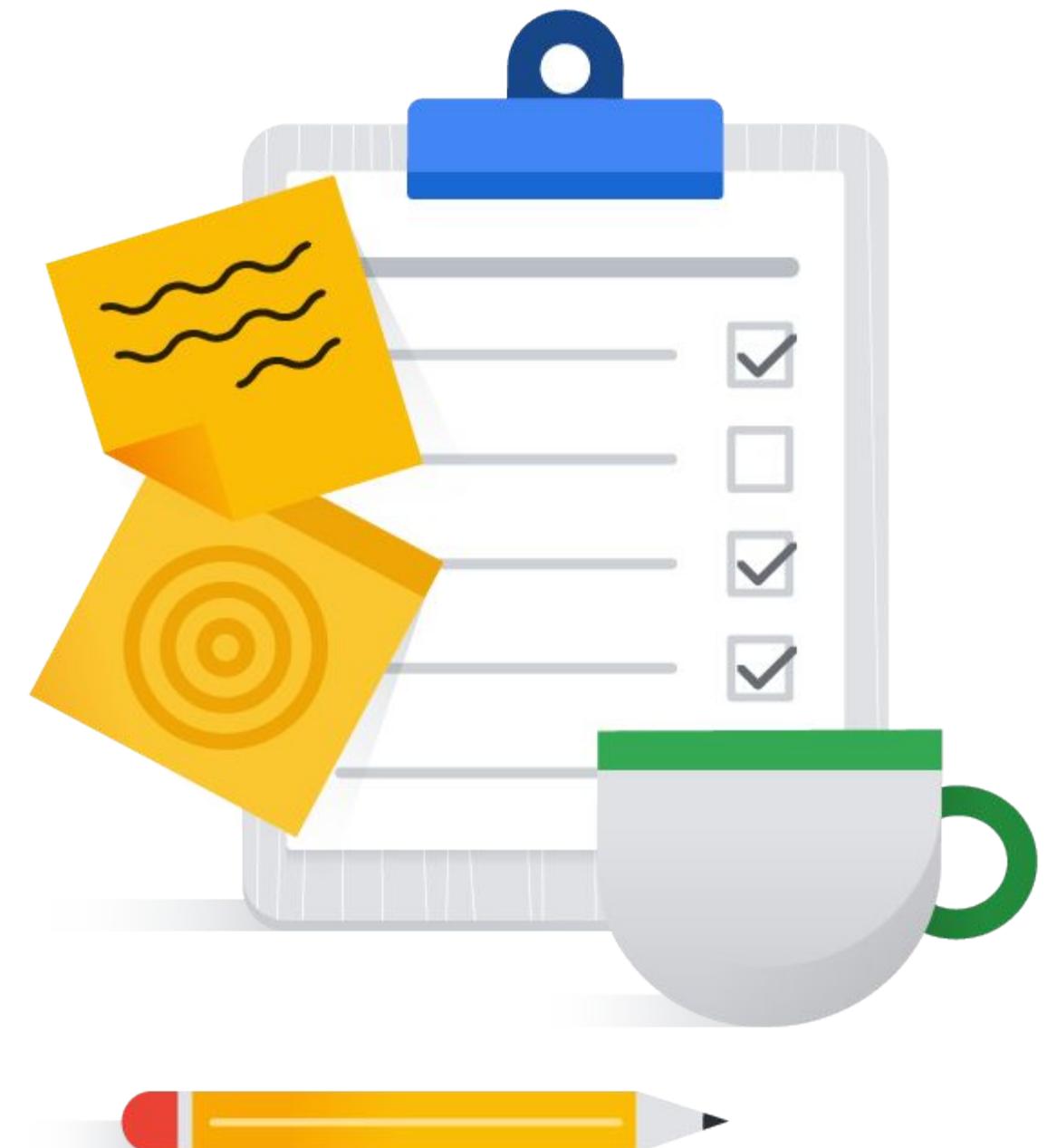
Use cases:

- Contract information extraction
- Feedback classification and ticket creation
- Sentiment analysis*
- Content translation*
- Structured data extraction from file*
- Media tagging*
- Product tagging*
- Content moderation *

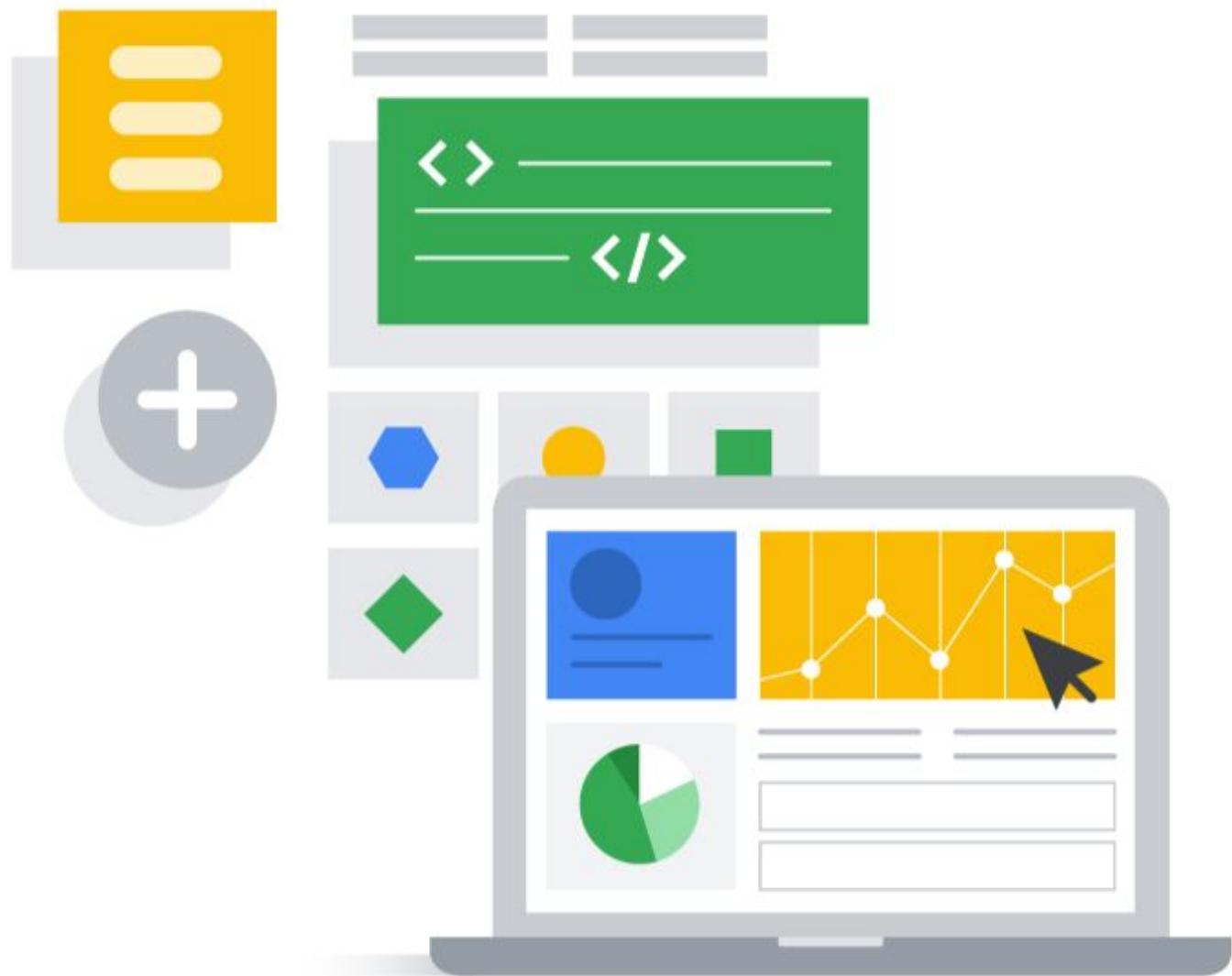
*may require Responsible AI Review

Topics

01	Next Steps
02	Q&A



Is there anything else you want to see?



Is there anything else you want to ask?



