

வணக்கம்
Welcome

GDG Cloud Chennai

Google Cloud GenAI 101



- ✓ Continuous learner
- ✓ Community Volunteer
- ✓ Academic Stimulator



Kumaresan M K
Cloud Integration Architect

Google

🤖 AI Conversation

CISCO

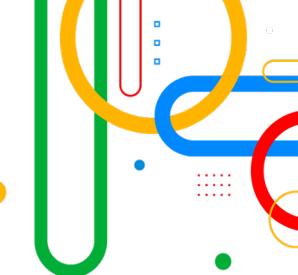
📞 VoIP Conversation

Capgemini

🌐 Packet Conversation

Agenda:

- ✓ Introduction
- ✓ Buzzwords **LLM ,GenAI, Prompt**
- ✓ How GenAI works?
- ✓ Google's offering
- ✓ GCP Cloud AI Portfolio
- ✓ GenApp Builder demo
- ✓ Summary



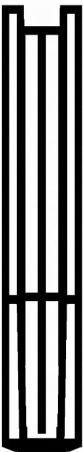
Let's recall some basics



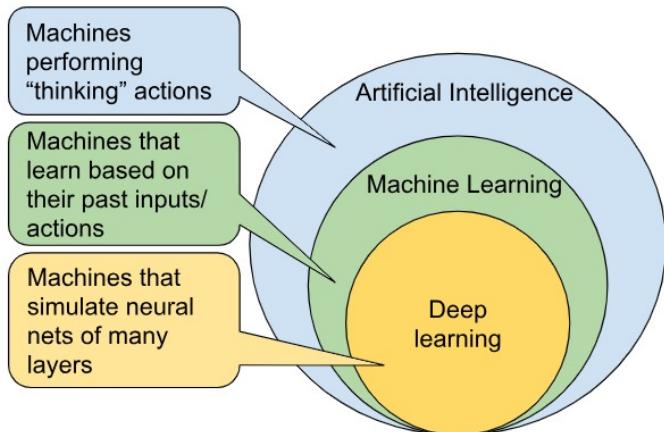
Before we go into **GenAI**
We need to have a good understanding **on AI**

- Artificial Intelligence ?
- Machine Learning ?
- Deep Learning ?
- Neural Networks?

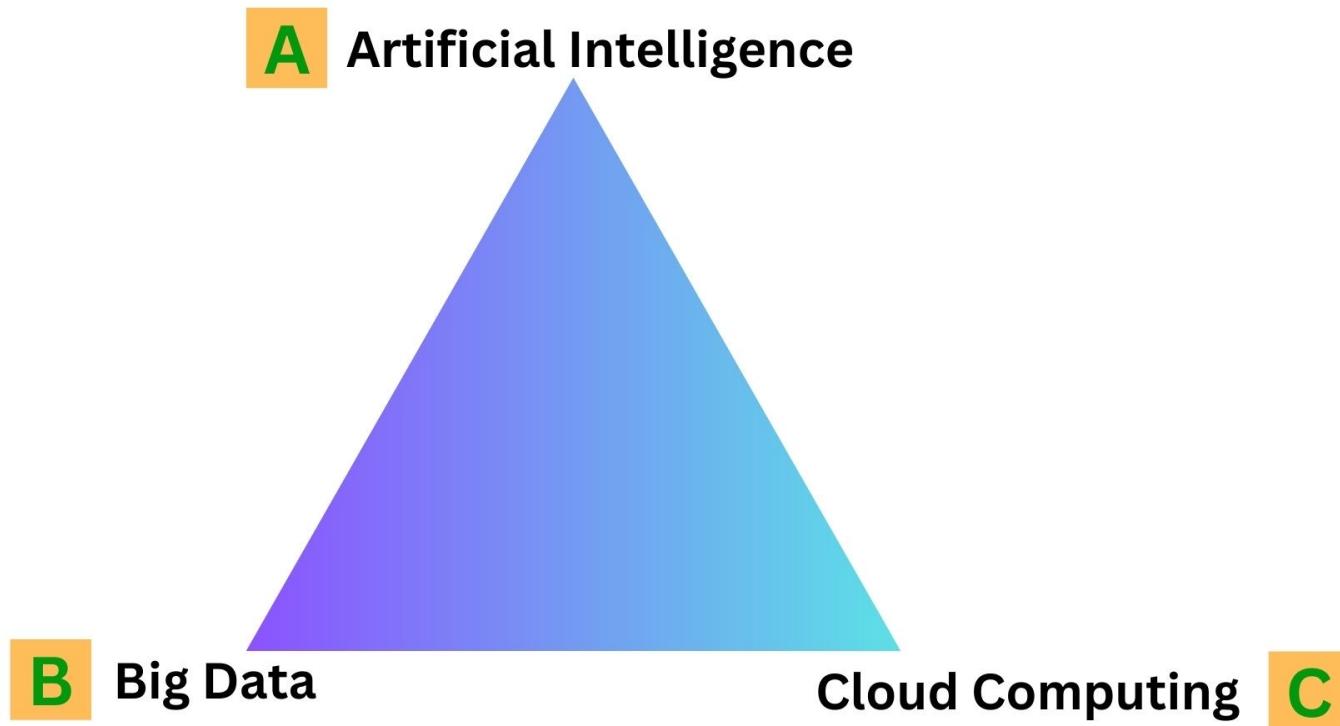
Let's recall some basics



Before we go into **GenAI**
We need to have a good understanding **on AI**



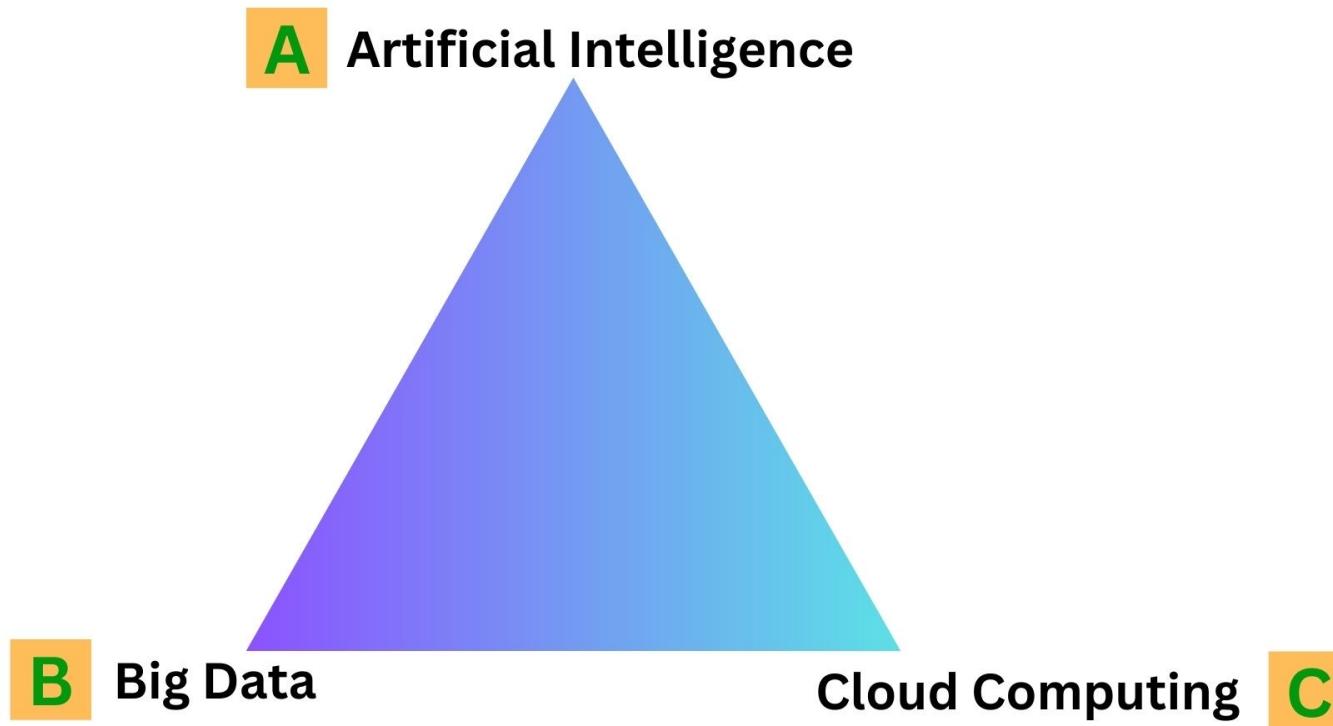
ABC Triangle



ABC Triangle

Year 1956:

The word "**Artificial Intelligence**" first adopted by American Computer scientist John McCarthy

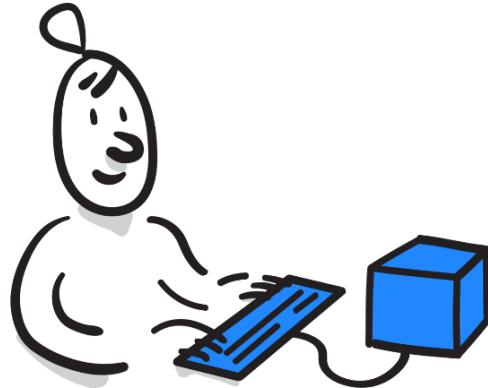


Cat:

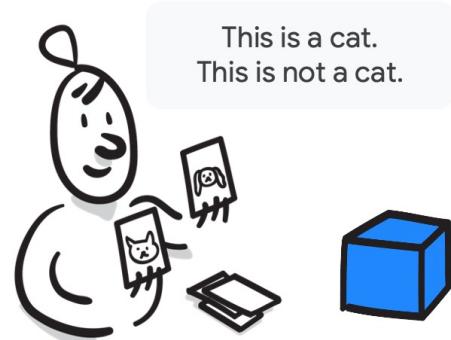
```
type: animal  
legs: 4  
ears: 2  
fur: yes  
likes: yarn, catnip
```

(etc ...)

how we train the computer ?



Traditional
programming

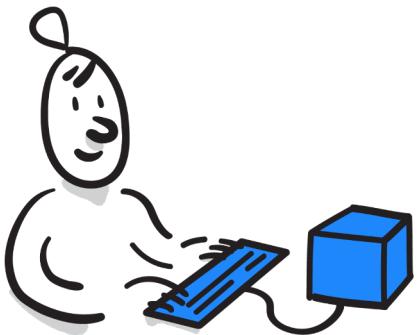


Wave of
neural networks

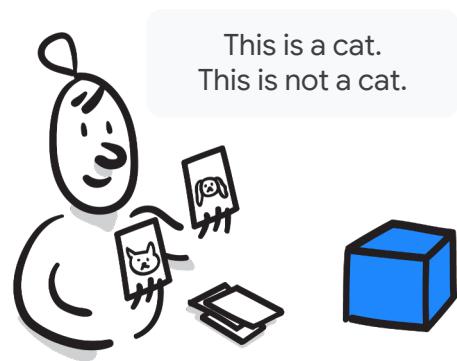
~2012

Cat:
 type: animal
 legs: 4
 ears: 2
 fur: yes
 likes: yarn, catnip

(etc ...)



Traditional
programming

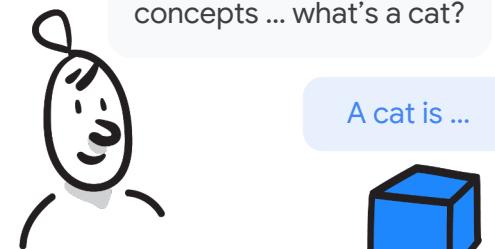


Wave of
neural networks

~2012



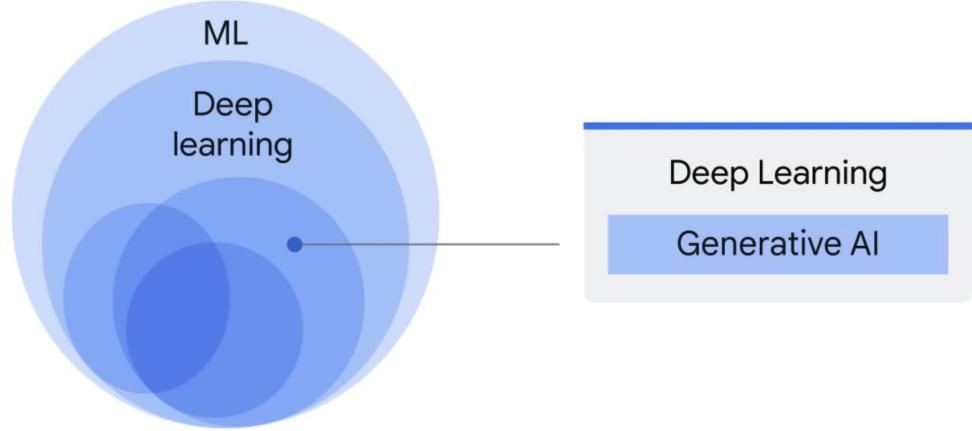
So, you've learned about
cats and millions of other
concepts ... what's a cat?



Generative
language models

LaMDA, PaLM, GPT-3, etc.

Generative AI
is a **subset of**
Deep Learning



First, some vocab

Generative Model: An ML model that **produces content**, including language, code, imagery, music, video, etc.

Large Language Model: machine learning models that are really good at **understanding** and **generating** human language. Large refers to both size and complexity of the neural network itself, as well as the size of the dataset it was trained on. Can also be used for non-generative language tasks (eg classification)

Transformer: The **neural network architecture** underlying most of today's generative models (and, more widely, most SOTA ML models). Published by Google in 2017.

Prompt : In the world of generative AI, a prompt is just a fancy name for the input text that you feed to your model.

Generative AI can able to create below :



Text



Code



Image



Speech



Video



3D



PALM



Google's Foundation Models

Generative AI describes algorithms that generate new outputs based on data they have been trained on rather than simply perceiving and classifying

Generative AI is multi-modal using text input to create new outputs as text, code, audio, images, video and simulations

Foundation models of different sizes offer different benefits in latency, cost, and fine tuning optimizations for certain data sets

Generative AI is driving new opportunities

Operational efficiencies, cost savings, and value creation



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

Generate text, code, image, video, or music quickly and multi-modally, speeding up every business processes and **maximizing employee productivity**



Creative

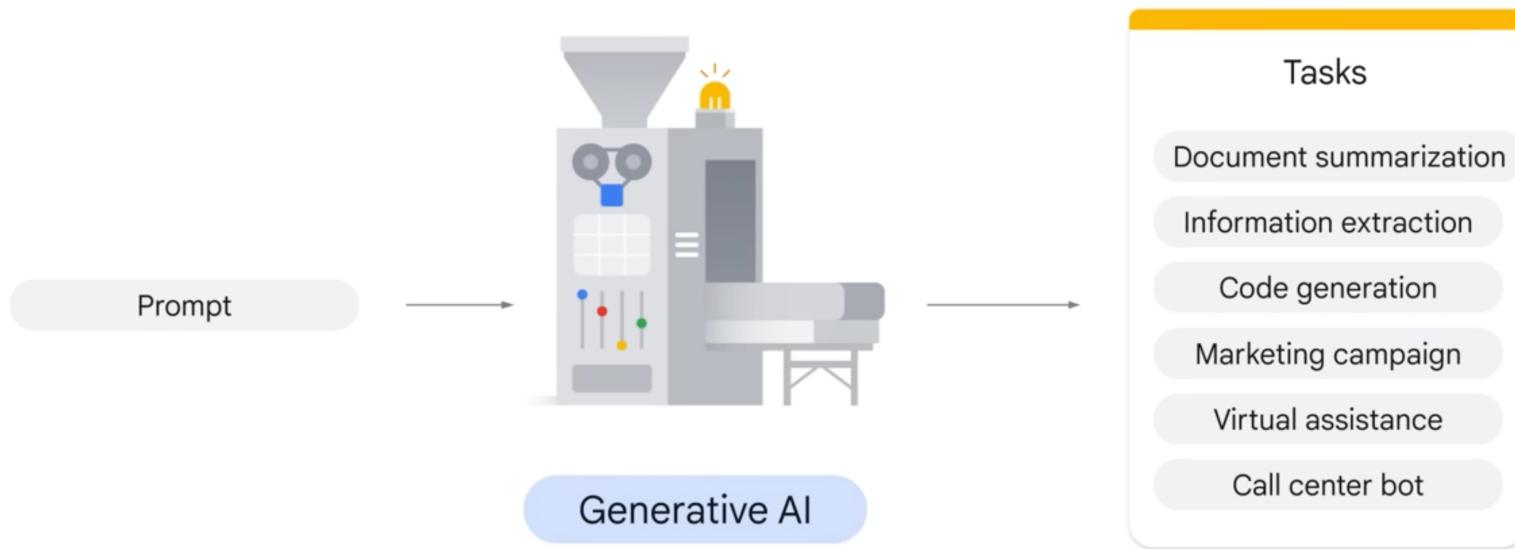


Customize Foundational models

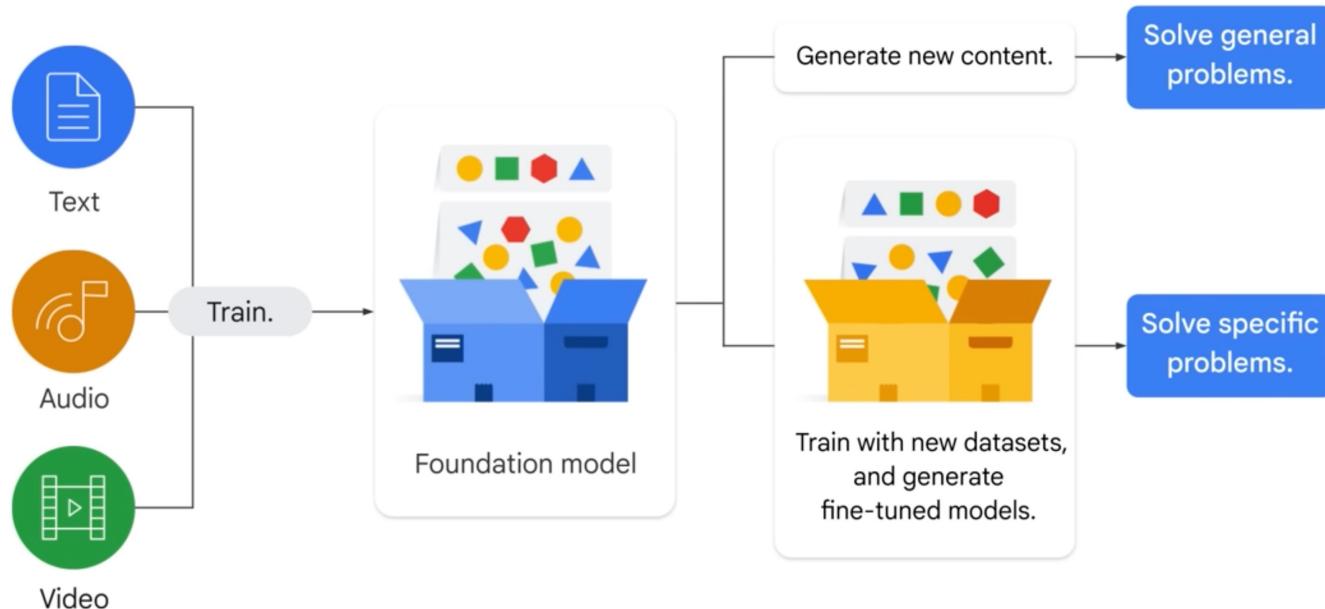
Customize large models and incorporate state of the art generative capabilities natively into **your own internal ML operational platforms**



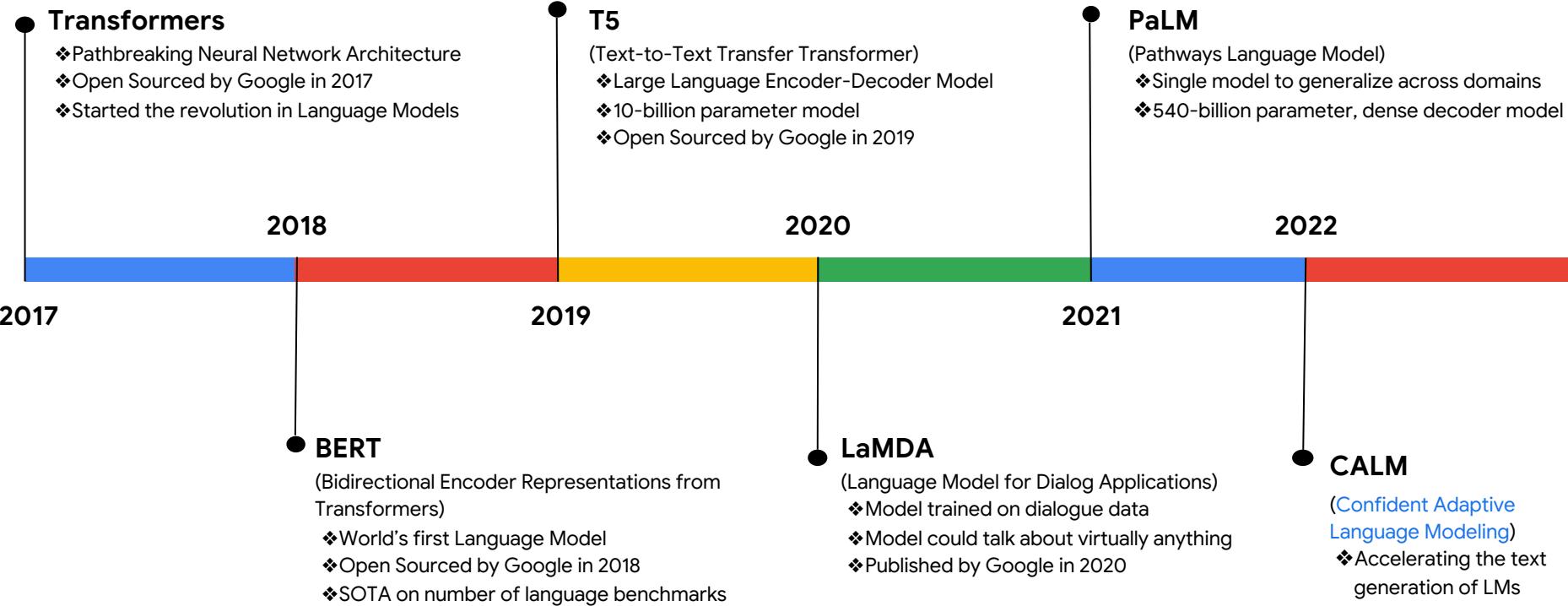
AI Practitioner



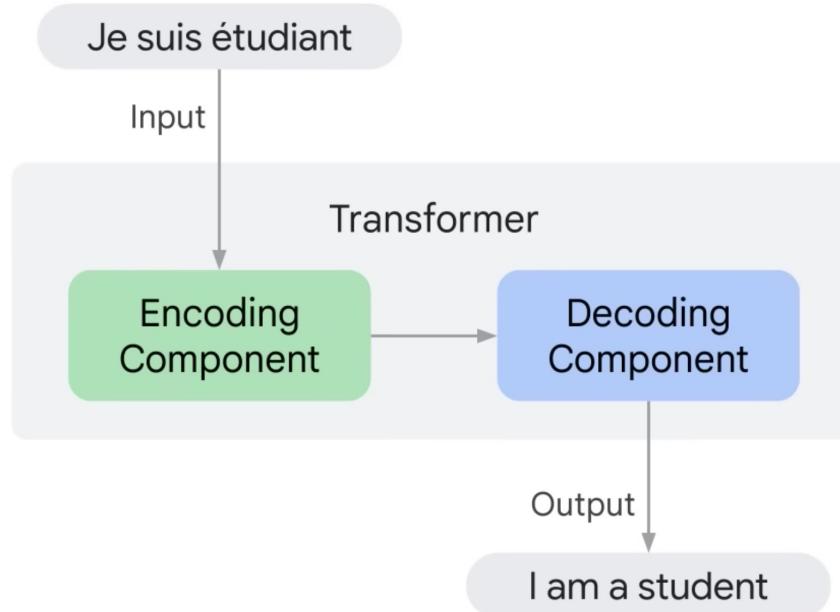
How does AI generate new content?

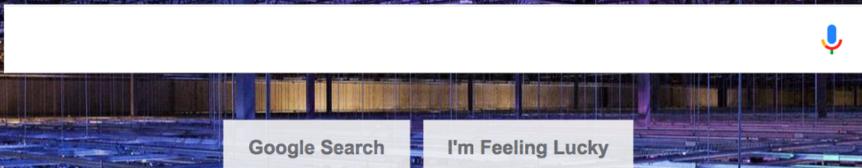


LLM Research at Google



Transformer model





Expanding the AI frontier of performance

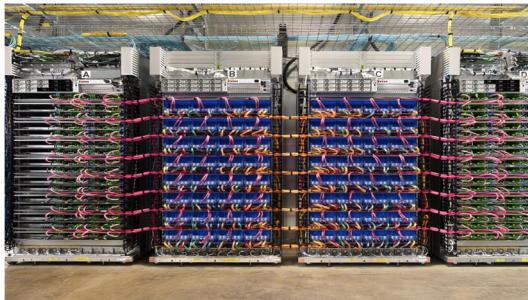


Cloud TPU v2

180 teraflops

64 GB HBM

Training and inference



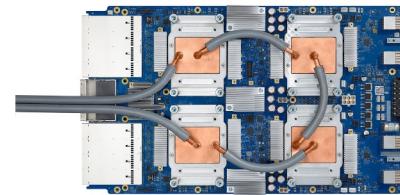
Cloud TPU v2 Pod

11.5 petaflops

4 TB HBM

2-D toroidal mesh network

Training and inference



Cloud TPU v3

420 teraflops

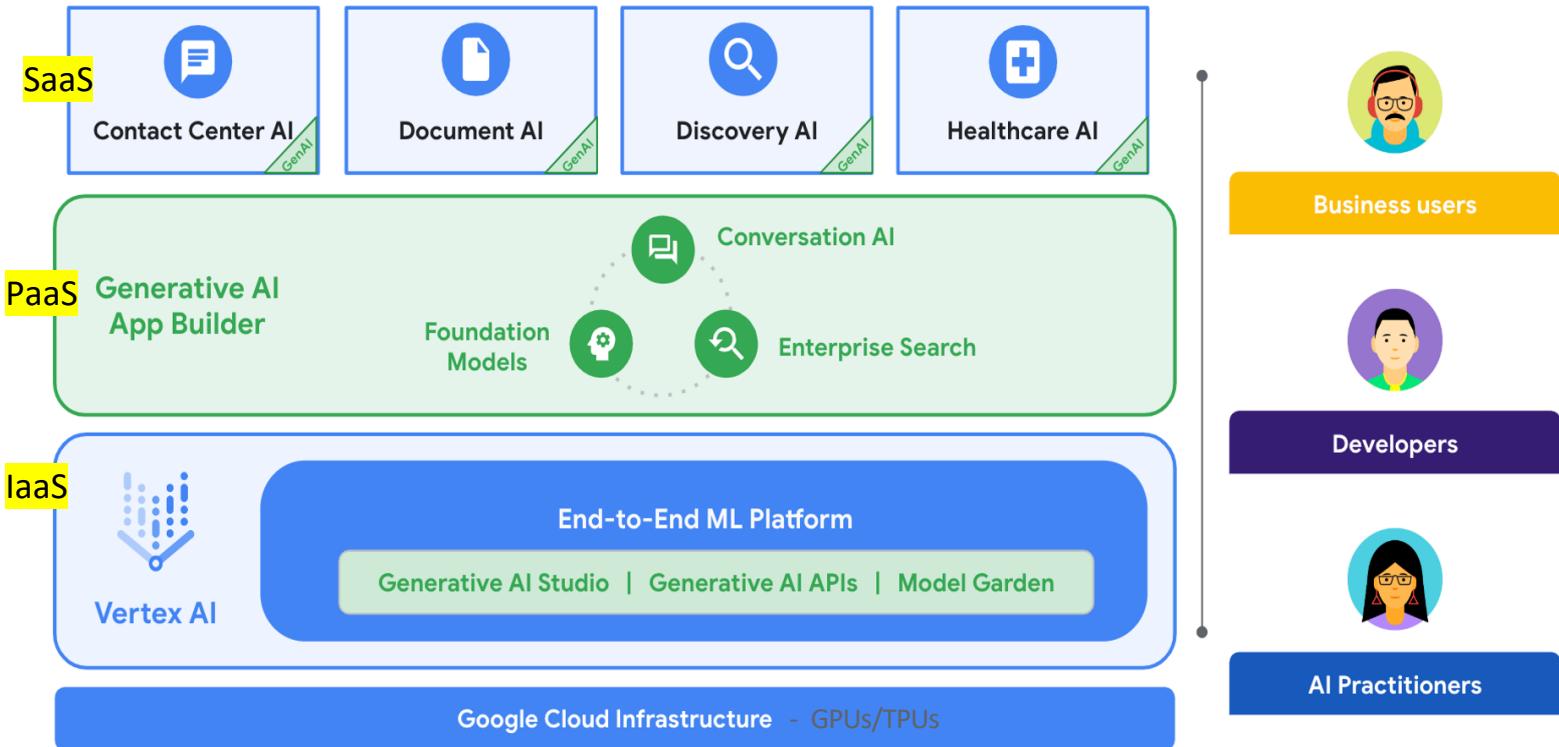
128 GB HBM

Training and inference

(**11.5 petaflops = 100,000 times**
more powerful than a typical
personal computer)

Cloud AI Portfolio

To support the needs of **Generative AI** centric enterprise development.

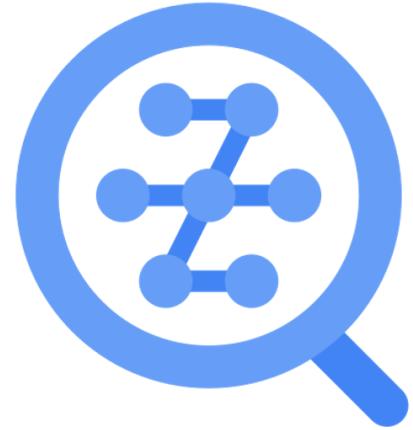


Generative AI App Builder (Gen App Builder)

- Cloud Documentation: <https://cloud.google.com/generative-ai-app-builder/docs/introduction>
- Generative AI App Builder (Gen App Builder) lets developers, even those with limited machine learning skills, tap into the power of Google's foundation models, search expertise, and conversational AI technologies to create enterprise-grade generative AI applications.
- Gen App Builder is comprised of the following 3 features:
 - Generative AI agent (aka. Infobot)
 - Enterprise Search (aka. ES, Unified Cloud Search, UCS)
 - Personalize (Recommendation)

Harnessing the power of AI for next-gen search experience

Generative AI App Builder: Enterprise Search



Data Ingestion Capabilities Overview



Internet



On-prem Data



Bigquery



GCS

Public URLs

CRUD API

Bigquery Import

GCS Import



Enterprise
Search

Website Search

- Instant search out-of-the-box
- Advanced site search with full fledged features
- URL-based crawling

Structured Data Support

- JSON-schema based format
- Client-specified schema
- Dynamic mapping
- Streaming and batch import

Unstructured Data Support

- PDF, HTML, Docx, PPTX, Text etc
- Semantic understanding & parsing for extractive/generative answers
- CRUD and GCS import

Tutorial: Getting Started

Gen App Builder [Create Engine](#)

Create your new engine

Select the type of engine you want to create

Search
Get quality results out-of-box and easily customize the engine

Personalize PREVIEW
Build a similar documents personalization engine

Name your engine

Engine Name *

ID: my-test-engine_1687920342517. It cannot be changed later. [EDIT](#)

Want to use advanced LLM features in your app?

Learn more about [summarization](#) and [multi-turn search](#)

Advanced LLM features PREVIEW
Advanced LLM features enable you to summarize results from unstructured search queries and perform multi-turn searches. This feature is currently in preview. After it is generally available, additional charges will apply.

[CONTINUE](#)

Gen App Builder [Create Engine](#)

Create Engine

Create Data Store

Create a data store for your engine

Select the type of data your engine will index

Website
Create a data store from your public website

Unstructured data
Data that doesn't conform to a database schema. Fully supported formats: HTML & PDF, in preview supported formats: DOCX, PPTX, TXT.

Structured data PREVIEW
Data that is organized in a specific format and follows a schema

[CREATE](#)

Tutorial: Configuring Basic Website Search (GA)

Specify the websites for your data store

Sites to add *
cloud.google.com/*

One site per line, https not needed 1/10

Sites to exclude
cloud.google.com/private/*

One site per line, https not needed 1/20

You can use the operations listed below

Individual pages: www.mysite.com/faq.html
Entire site: www.mysite.com/*
Parts of site: www.mysite.com/faq/*
Entire domain: *.mysite.com

SAVE Data Import Experience   

Engines > test-chlei-website > Search Preview

Generative AI App Builder

Showing you 10 out of 456 results

<https://cloud.google.com/.../ai.../create-generative-apps-in-minutes-with-gen-app-builder>
[Create generative apps in minutes with Gen App Builder | Google ...](#)
Mar 29, 2023 ... Gen App Builder is exciting because unlike most existing generative AI offerings for developers, it offers an orchestration layer that abstracts ...

<https://cloud.google.com/generative-ai-app-builder>
[Generative AI App Builder | Google Cloud](#)
Gen App Builder makes it simple to create digital assistants and bots that not only serve content, but also connect to purchasing and provisioning systems to ...

<https://cloud.google.com/.../ai.../generative-ai-for-businesses-and-governments>
[Google Cloud brings generative AI to developers, businesses, and ...](#)
Mar 14, 2023 ... Generative AI App Builder allows developers to quickly ship new experiences including bots, chat interfaces, custom search engines, ...

<https://cloud.google.com/blog/.../ai.../enterprise-search-on-gen-app-builder>
[Exploring Enterprise Search in Generative AI App Builder | Google ...](#)

Note: Advanced Website Search is under limited preview

Tutorial: Configuring Unstructured Search (GA)

The screenshot shows the Google Cloud Search interface. On the left, there's a sidebar with a search bar at the top and three main sections: 'Add data to your data store', 'Cloud Storage (without metadata)', and 'Cloud Storage (with metadata)'. The 'Cloud Storage (without metadata)' section is expanded, showing 'Your documents in a GCS bucket'. The 'Cloud Storage (with metadata)' section is also expanded, showing 'JSONs in a GCS bucket with links to your documents and metadata'. The 'BigQuery (with metadata)' section is collapsed. On the right, there's a 'Import data from Cloud Storage' panel. It has a 'gs://*' input field containing 'cloud-samples-data/gen-app-builder/search/alphabet-investor-pdfs', a 'BROWSE' button, an 'IMPORT' button, and a 'Data Import Experience' link.

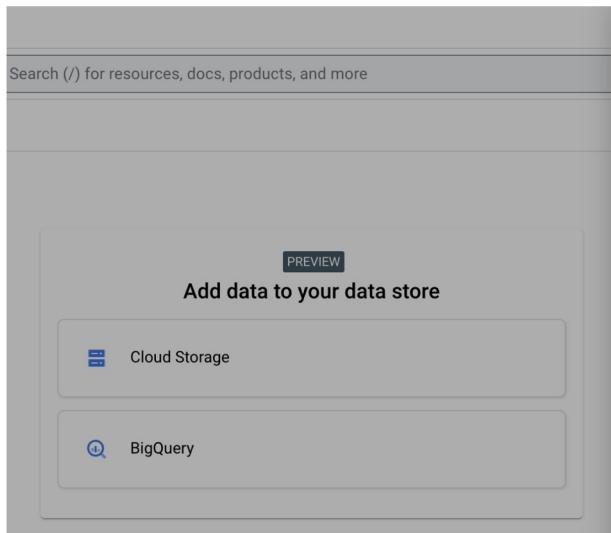
Supported sources:

- GCS Bucket
- GCS (with metadata)
- BigQuery (with metadata)

Supported formats:

- PDF
- HTML
- TXT
- DOCX
- PPTX
-

Tutorial: Configuring Structured Search (Preview)



Import data from BigQuery

X LEARN

Import the data you have stored in BigQuery. To learn how to upload your data to BigQuery, see the [BigQuery](#) documentation.

BigQuery path

bigquery-public-data.google_ads.geotargets

BROWSE

Search by table name or path using the format: projectId datasetId tableId.



We will detect the schema automatically for you

IMPORT

Data Import Experience



Supported sources:

- BigQuery
- GCS

Supported formats:

- JSONL

Tutorial: Data Import Tracking

Engines > test-chlei-unstructured > Data

LEARN

DETAILS

ACTIVITY

Activity

+ IMPORT DATA

REFRESH

Filter

?

Status	Detail	Items succeeded	Operation name	Last updated
<input checked="" type="checkbox"/> Import completed	No errors	173	import-documents-15004784485356394957	Jun 27, 2023, 8:00:54PM

Engines > test-chlei-structured > Data

LEARN

DETAILS

SCHEMA

PREVIEW

ACTIVITY

Activity

+ IMPORT DATA

REFRESH

Filter

?

Status	Detail	Items succeeded	Operation name	Last updated
⚠ Import completed with issues	829875 items failed to integrate View full error logs	0	import-documents-146199208207020188	Jun 27, 2023, 8:10:52PM

Tutorial: Preview Search Results

Engines > test-chlei-unstructured > Search Preview



What is Alphabet's revenue in 2022?

X

Alphabet's revenue in 2022 was \$283 billion, up 10%, or up 14% in constant currency.

Showing you 10 results



[1] 2022Q4_alphabet_earnings_release.pdf

We're on an important journey to re-engineer our cost structure in a durable way and to build financially sustainable, vibrant, growing businesses across Alphabet." Ruth Porat, CFO of Alphabet and Google, said: "Our Q4 consolidated revenues were \$76 billion, up 1% year over year, or up 7% in constant currency, and \$283 billion for the full year 2022, up 10%, or up 14% in constant currency.

Page 1

[2] 2022Q1_alphabet_earnings_release.pdf

Quarter Ended March 31, 2021 2022 Revenues \$ 55314 \$ 68011 Change in revenues year over year 34 % 23 % Change in constant currency revenues year over year(1) 32 % 26 % Operating income \$ 16437 \$ 20094 Operating margin 30 % 30 % Other income (expense), net \$ 4846 \$ (1160) Net income \$ 17930 \$ 16436 Diluted EPS \$ 26.29 \$ 24.62 (1) Non-GAAP measure.

Page 1

[3] 2022Q2_alphabet_earnings_release.pdf

Quarter Ended June 30, 2021 2022 Revenues \$ 61880 \$ 60685 Change in revenues year over year (1) 62 % 13 %

Traditional ML Dev

- Needs 1000+ training examples to get started
- Needs ML expertise (probably)
- Needs compute time + hardware
- Thinks about minimizing a loss function

LLM Dev

- Needs 0 training examples*
- Does not need ML expertise*
- Does not need to train a model*
- Thinks about prompt design

Summary



For Reference documents being used for demo

Generative AI App Builder

<https://cloud.google.com/generative-ai-app-builder/docs/introduction>

Other Git hub Resources

<https://github.com/GoogleCloudPlatform/generative-ai/tree/main/gen-app-builder>

<https://github.com/GoogleCloudPlatform/generative-ai>

<https://github.com/orgs/GdgCloudChennai/repositories>

[GDG Cloud Chennai](#)



Thank You

To continue the conversation with speaker 

