Google Developer Groups

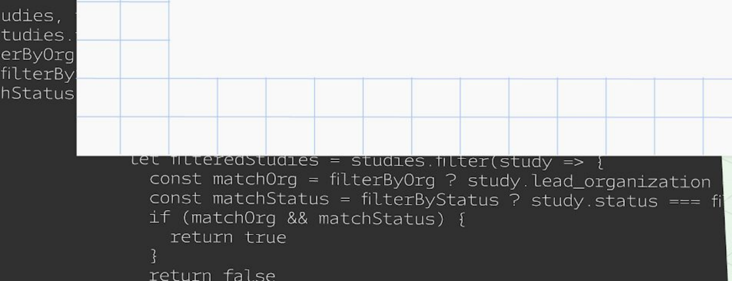
Road to Google Developers Certification

Google Cloud

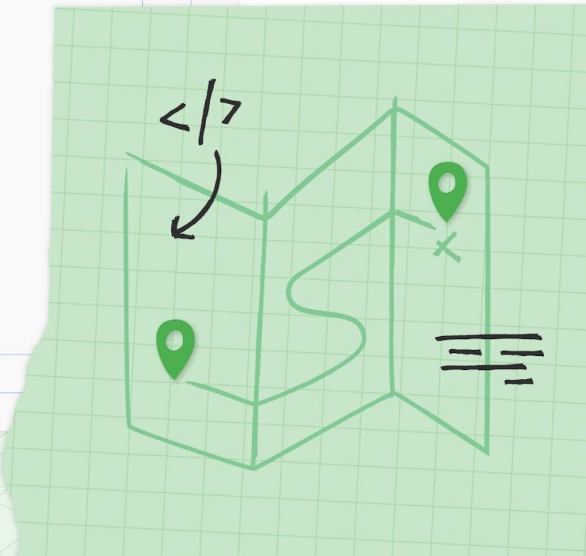
Week 5

Professional Machine Learning Engineer

ChengCheng Tan



```
udies,
tudies.
erByOrg
filterBy
hStatus
    let filteredStudies = studies.filter(study => {
      const matchOrg = filterByOrg ? study.lead_organization
      const matchStatus = filterByStatus ? study.status === fi
      if (matchOrg && matchStatus) {
        return true
      }
      return false
```



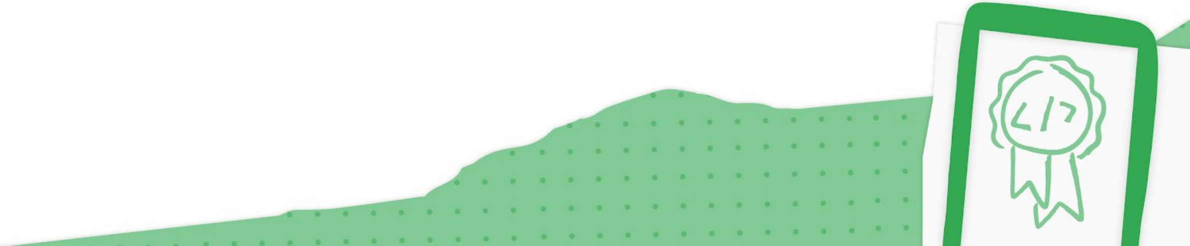
Who am I?

ChengCheng Tan

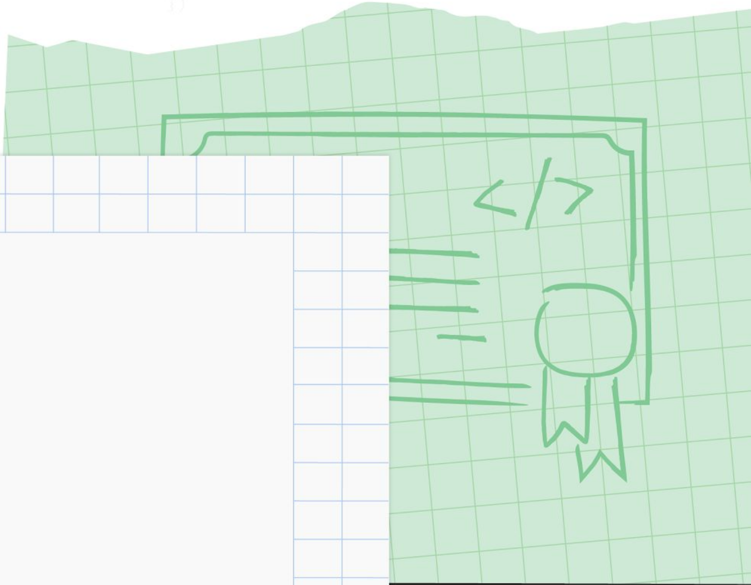
- BA Linguistics & Computer Science, UCLA
- MS Computer Science, Stanford
- LLM + AI Safety

  FAR AI Communications

  AISafety.info + chatbot
- Google Women Techmakers Ambassador



Where are we on our journey



Professional Machine Learning Certification

Learning Journey Organized by Google Developer Groups Surrey co hosting with GDG Seattle

Session 1 Feb 24, 2024 Virtual	Session 2 Mar 2, 2024 Virtual	Session 3 Mar 9, 2024 Virtual	Session 4 Mar 16, 2024 Virtual	Session 5 Mar 23, 2024 Virtual	Session 6 Apr 6, 2024 Virtual	**Review the** Professional ML Engineer Exam Guide **Review the** Professional ML Engineer Sample Questions	**Self study (and potential exam)**
Lightning talk + Kick-off & Machine Learning Basics + Q&A	Lightning talk + GCP- Tensorflow & Feature Engineering + Q&A	Lightning talk + Enterprise Machine Learning + Q&A	Production ML Systems and Computer Vision with Google Cloud + Q&A	Lightning talk + NLP & Recommendation Systems on GCP + Q&A	Lightning talk + MOPs & ML Pipelines on GCP + Q&A		
Complete course: Introduction to AI and Machine Learning on Google Cloud Launching into Machine Learning	**Complete course:** TensorFlow on Google Cloud Feature Engineering	**Complete course:** Machine Learning in the Enterprise	**Hands On Lab Practice:** Production Machine Learning Systems Computer Vision Fundamentals with Google Cloud	**Complete course:** Natural Language Processing on Google Cloud Recommendation Systems on GCP	**Complete course:** ML Ops - Getting Started ML Pipelines on Google Cloud **Check Readiness:** Professional ML Engineer Sample Questions	**Go through:** Google Cloud Platform Big Data and Machine Learning Fundamentals **Hands On Lab Practice:** Perform Foundational Data, ML, and AI Tasks in Google Cloud (Skill Badge) - 7hrs Build and Deploy ML Solutions on Vertex AI (Skill Badge) - 8hrs	

1 NLP Overview

2 Get Started with Gemini

3 Week 5 Content Review

4 Sample Question Review

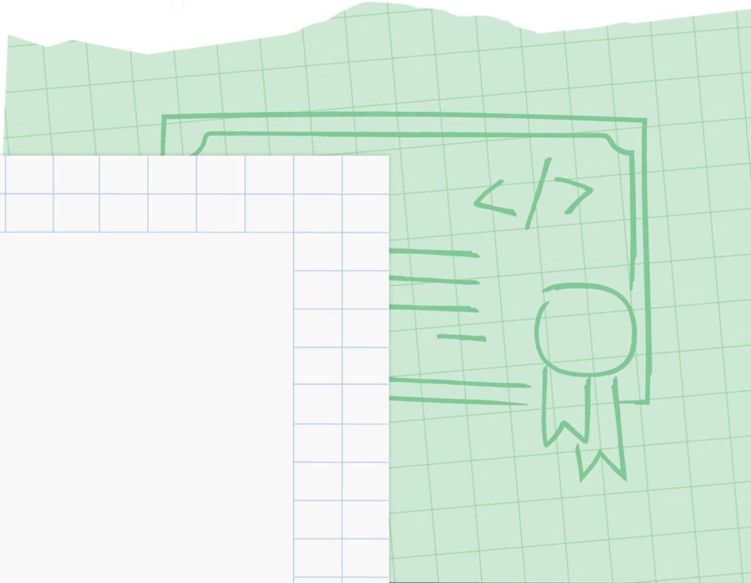
5 Q&A

Google Developer Groups

Road to Google Developers Certification

Google Cloud

NLP Overview



```
return false
})
```

```
nction filterStudies({ studies, filt
  let filteredStudies = studies.filt
    const matchOrg = filterByOrg ?
    const matchStatus = filterBySta
    if (matchOrg && matchStatus) {
      return true
    }
    return false
  })
```

```
Status = false }) {

=== filterByOrg : true
ilterByStatus : true
```

```
lse
          function filterStudies({ studies, filterByOrg = false, filterBy
            let filteredStudies = studies.filter(study => {
```

Natural Language Processing [NLP]: Computers Understand Human Languages

Google Developer Groups

Road to Google Developers Certification

Google Cloud

Pre-1990s:
Rule-Based Expert Systems

1990s-2000s: **Statistics & Probabilities**

bi-grams, tri-grams, n-grams

> You shall know a word by the company it keeps

J.R. Firth, Linguist

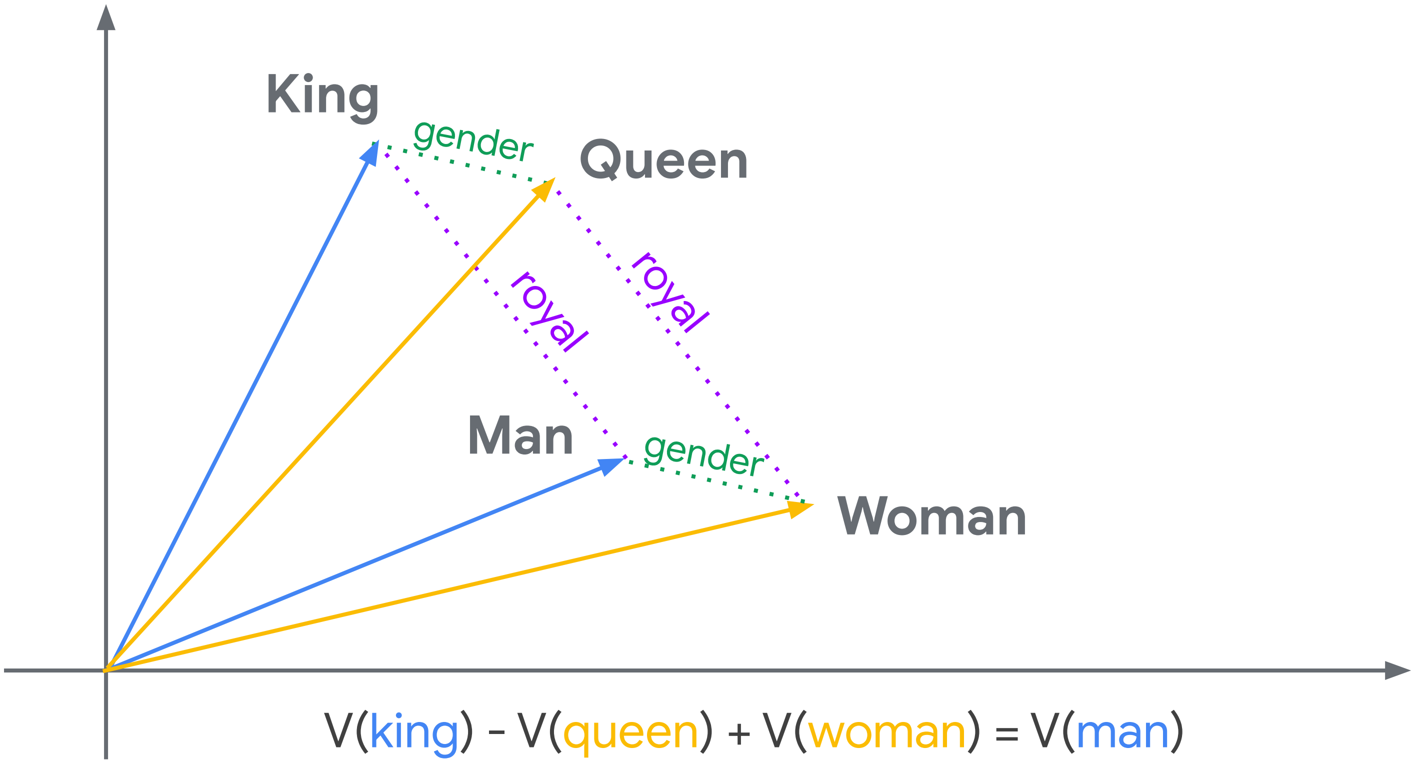
J.R. Firth, Linguist

```
function filterStudies({ studies, filterByOrg = false, filte
                         studies.filter(study => {
                                          organizat
```

2010s:

Rise of Deep Learning and Neural Networks

2013: Word2Vec Embeddings



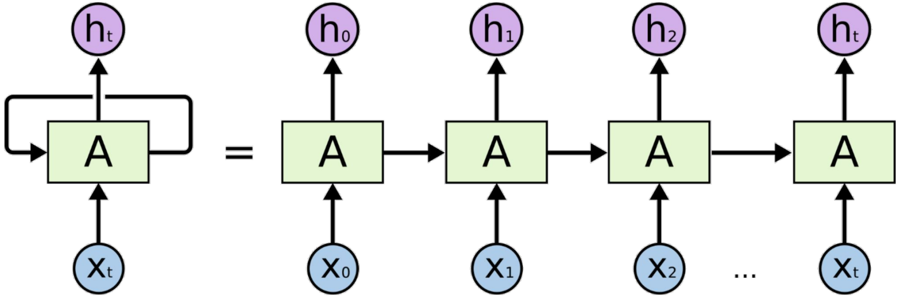
V(king) - V(queen) + V(woman) = V(man)

2013: Word2Vec Embeddings

Analogies	Word Pair 1		Word Pair 2	
Man-Woman	king	queen	man	woman
Capital city	Athens	Greece	Oslo	Norway
City-in-state	Chicago	Illinois	Sacramento	California
Opposite	possibly	impossibly	ethical	unethical
Comparative	great	greater	tough	tougher
Nationality adjective	Switzerland	Swiss	Canada	Canadian
Past tense	walking	walked	swimming	swam
Plural nouns	mouse	mice	dollar	dollars

2010s:
Neural Networks

RNN, GRU, LSTM



Early Neural Networks

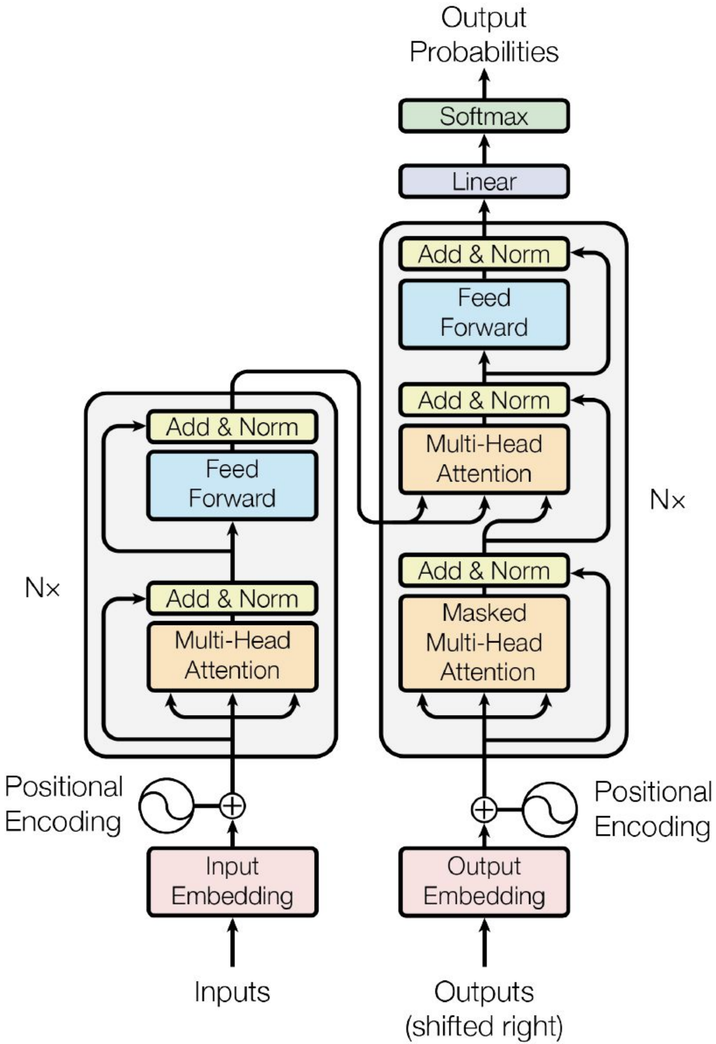
- Slow & Forgetful

2017: Transformers

- Self-Attention
- Algorithm+Data+Compute

Transformer Architecture

Encoder + Decoder



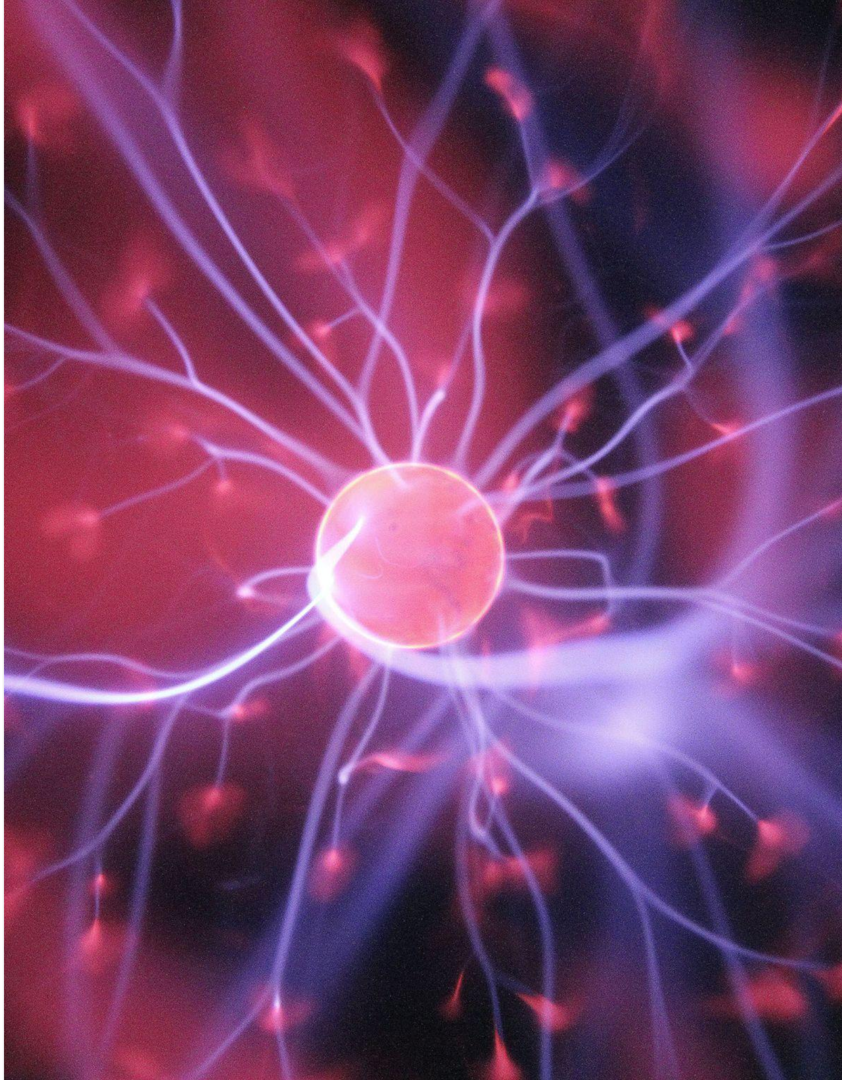
Google Developer Groups

Road to Google Developers Certification

Google Cloud

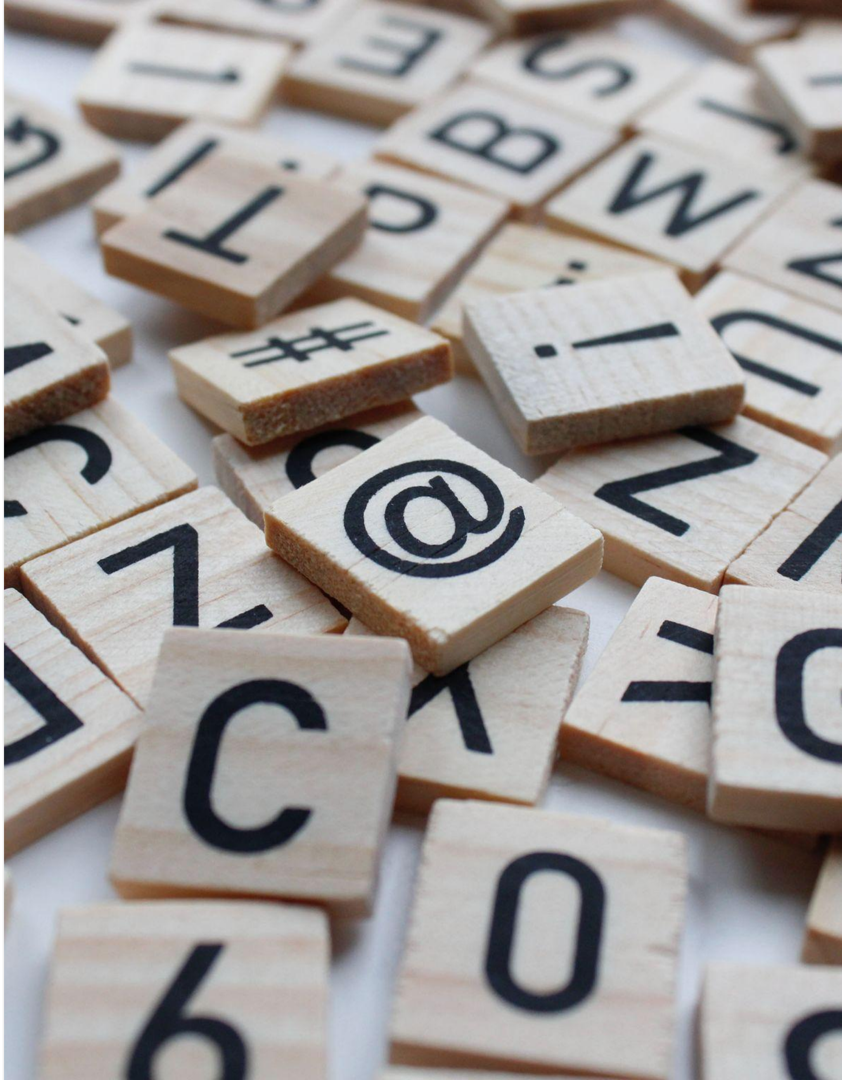
Rise of LLMs

>1 Billion Neurons



Trained for

Next Word Prediction

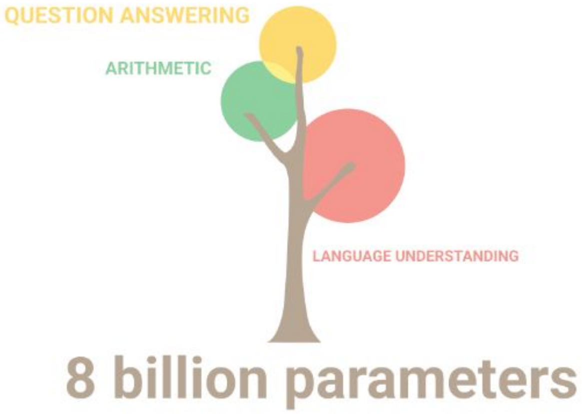


Google Developer Groups

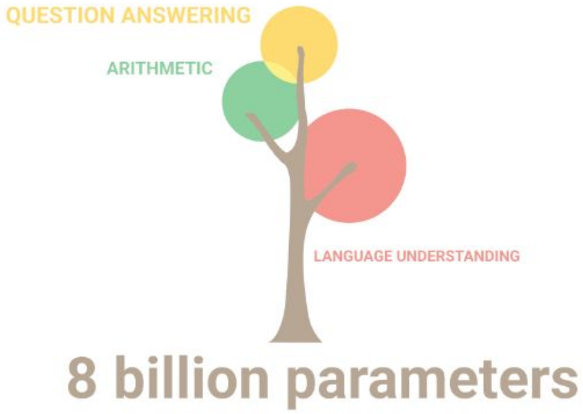
Road to Google Developers Certification

Google Cloud

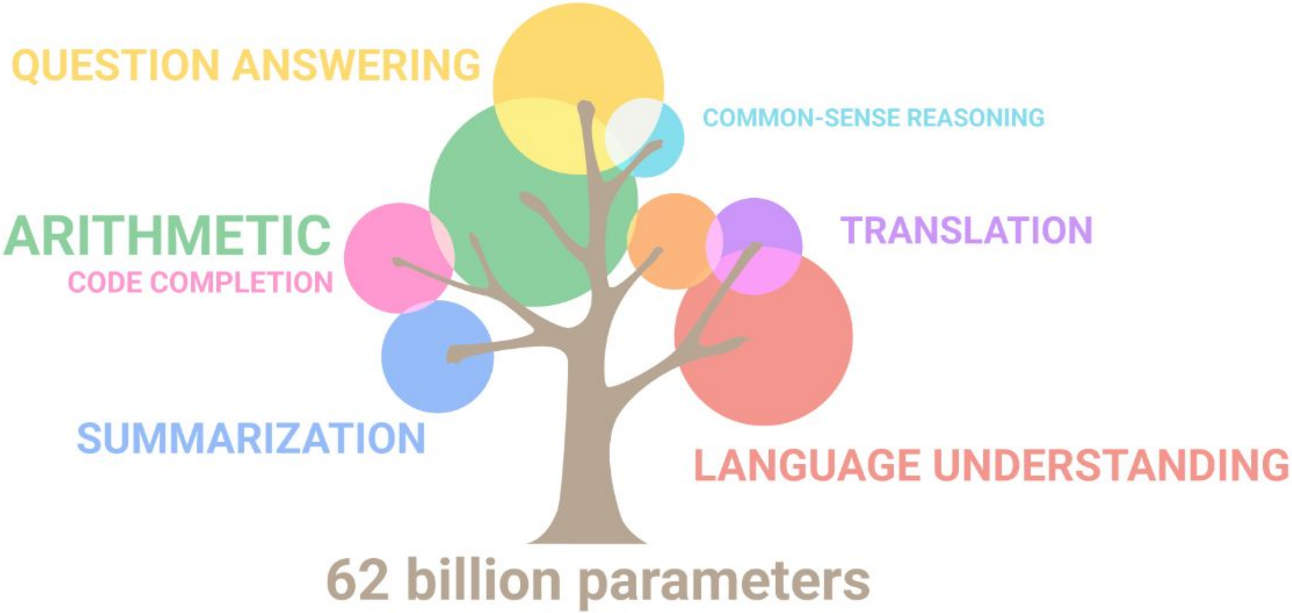
Emergent Abilities



Emergent Abilities



Emergent Abilities



Pre-trained Base

Generalist

vs

Fine-tuned Models

Specialists

RLHF:

Reinforcement Learning
from Human Feedback

Fine-tuned

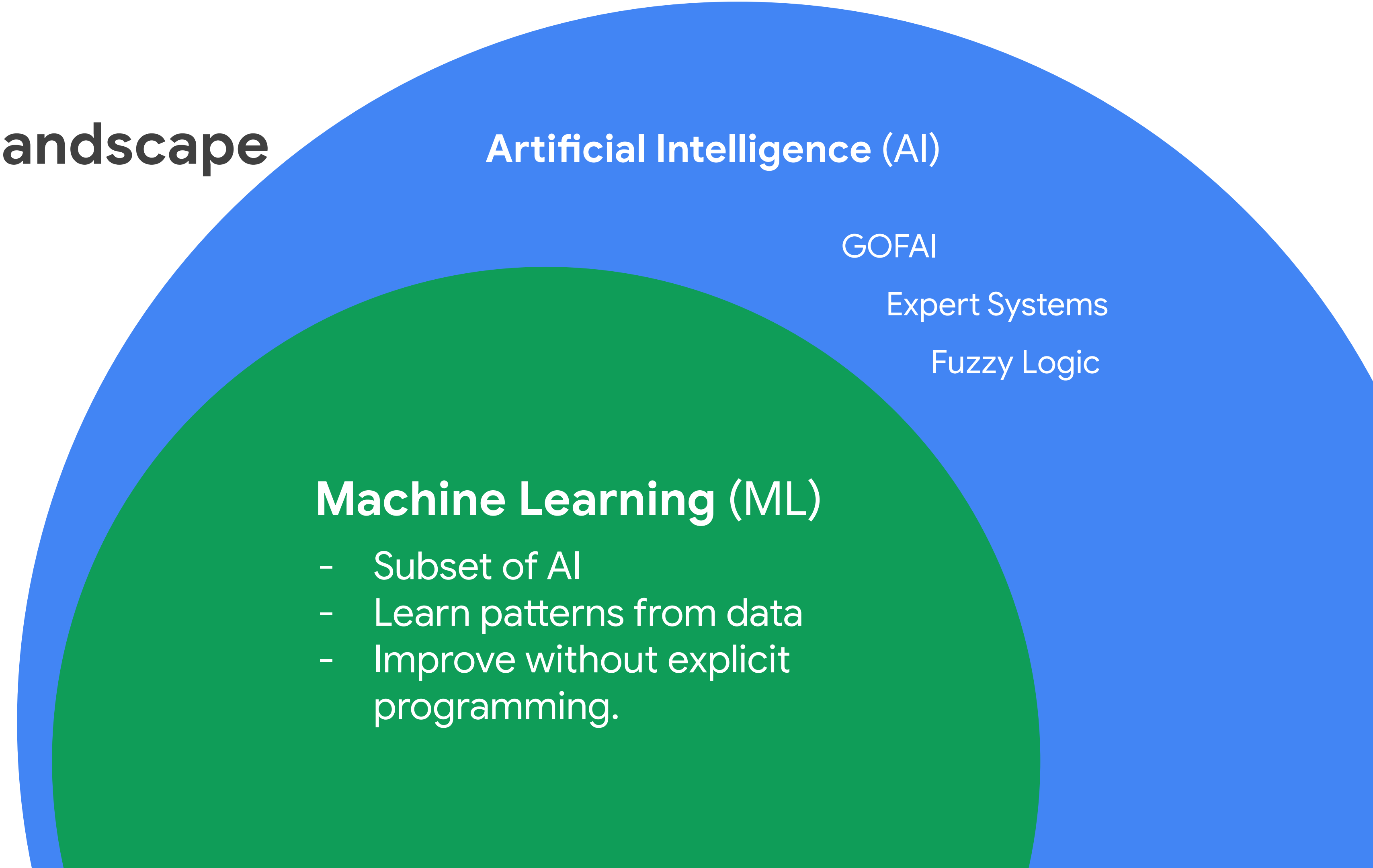
- Follow Instructions
- Conversations

AI Landscape

Artificial Intelligence (AI)

- Field of CS
- Create intelligent machines
- Perform tasks with human-like abilities
- Reasoning, learning & problem-solving

AI Landscape



Artificial Intelligence (AI)

GOFAI

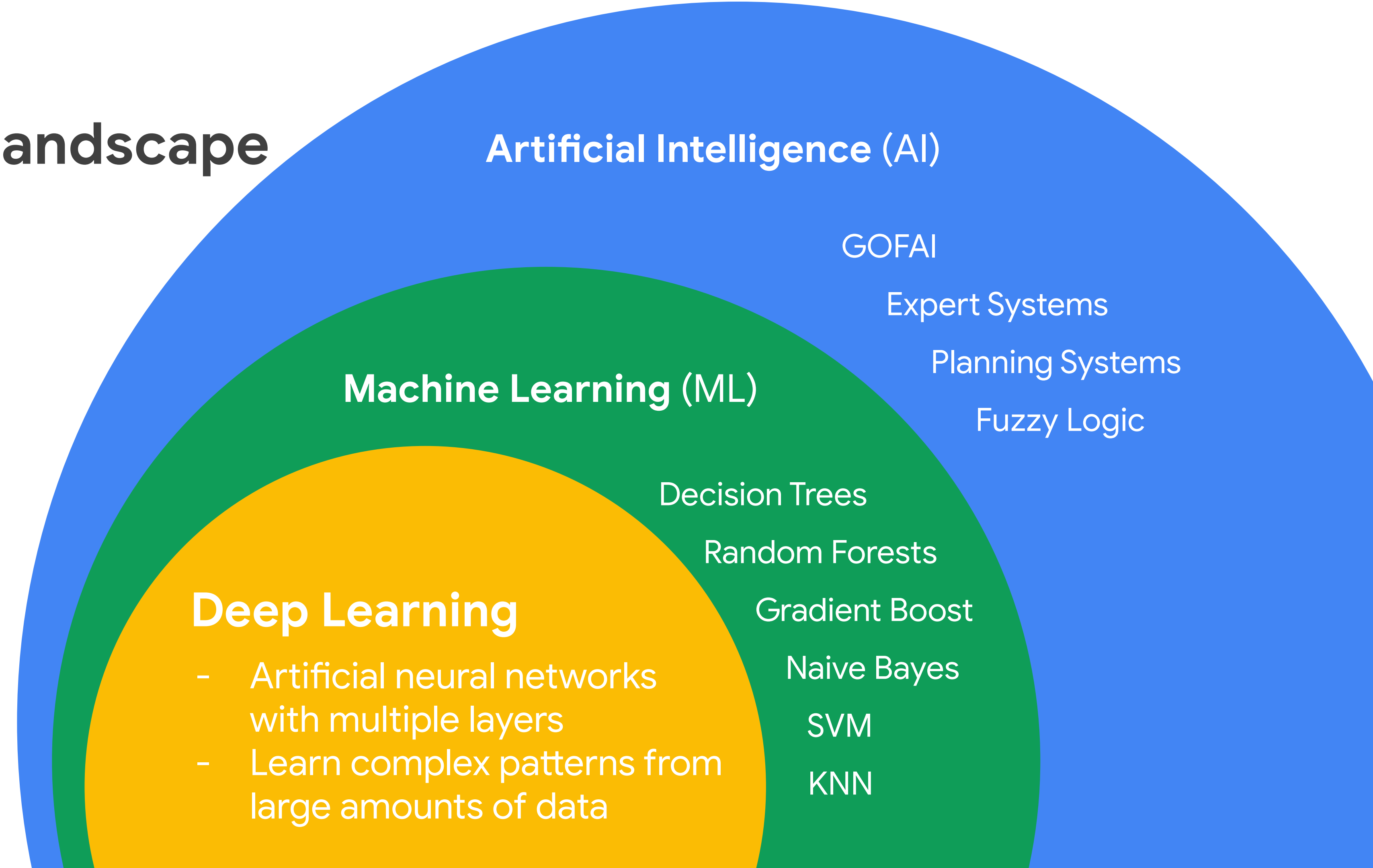
Expert Systems

Fuzzy Logic

Machine Learning (ML)

- Subset of AI
- Learn patterns from data
- Improve without explicit programming.

AI Landscape



Artificial Intelligence (AI)

GOFAI

Expert Systems

Planning Systems

Fuzzy Logic

Machine Learning (ML)

Decision Trees

Random Forests

Gradient Boost

Naive Bayes

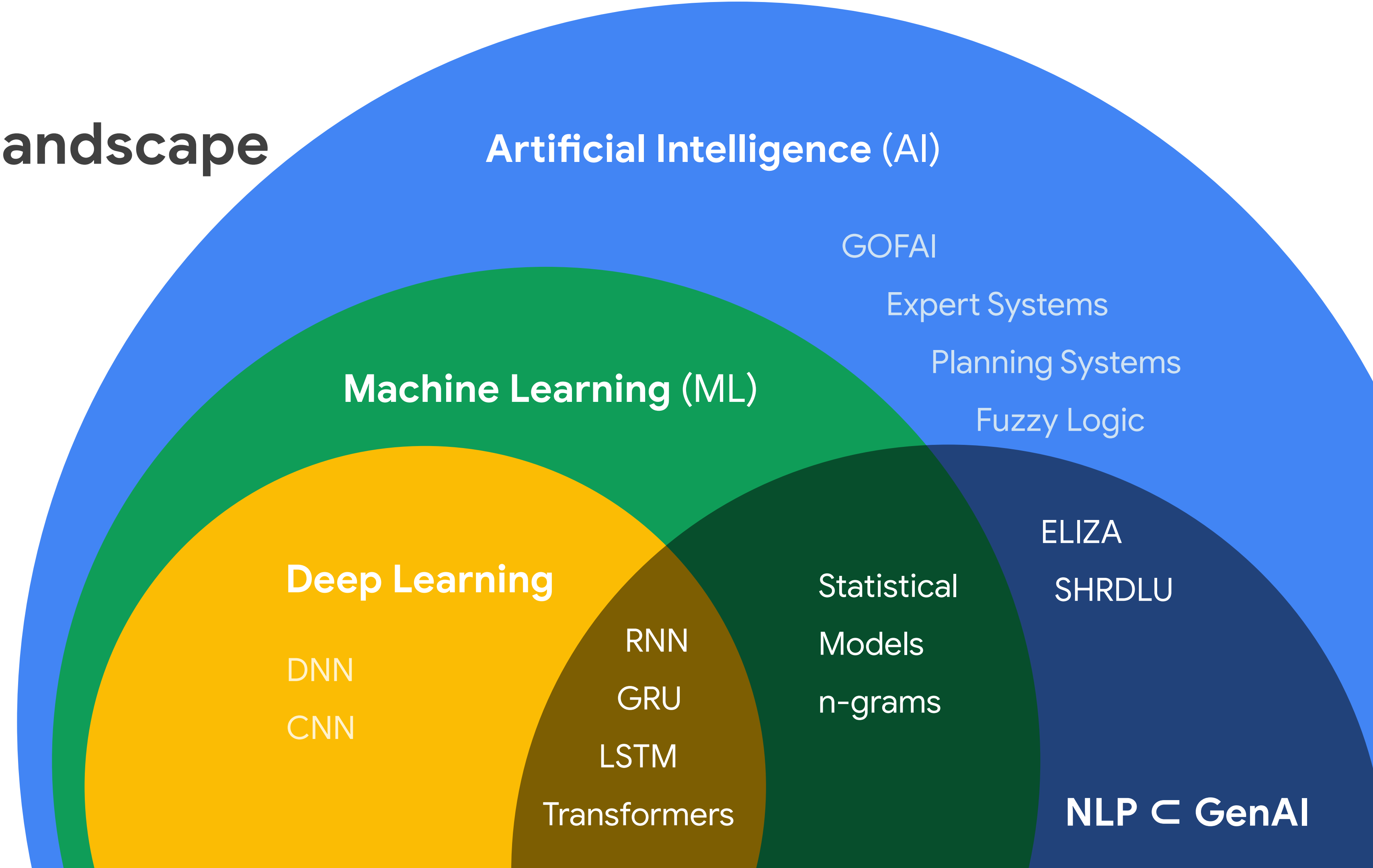
SVM

KNN

Deep Learning

- Artificial neural networks with multiple layers
- Learn complex patterns from large amounts of data

AI Landscape



Artificial Intelligence (AI)

GOFAI

Expert Systems

Planning Systems

Fuzzy Logic

Machine Learning (ML)

Deep Learning

DNN

CNN

RNN

GRU

LSTM

Transformers

Statistical

Models

n-grams

ELIZA

SHRDLU

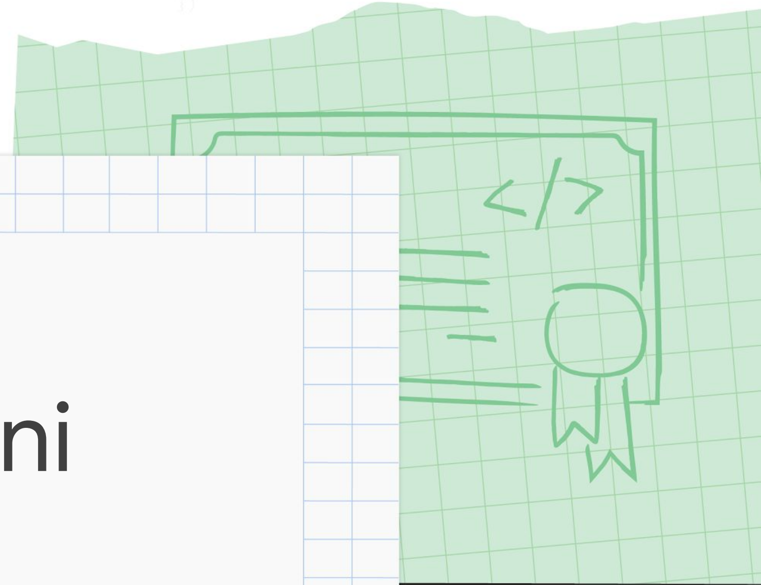
NLP ⊂ GenAI

Google Developer Groups

Road to Google Developers Certification

Google Cloud

Get Started with Gemini



```
nction filterStudies({ studies, filt
  let filteredStudies = studies.filt
    const matchOrg = filterByOrg ?
    const matchStatus = filterBySta
    if (matchOrg && matchStatus) {
      return true
    }
    return false
  })
```

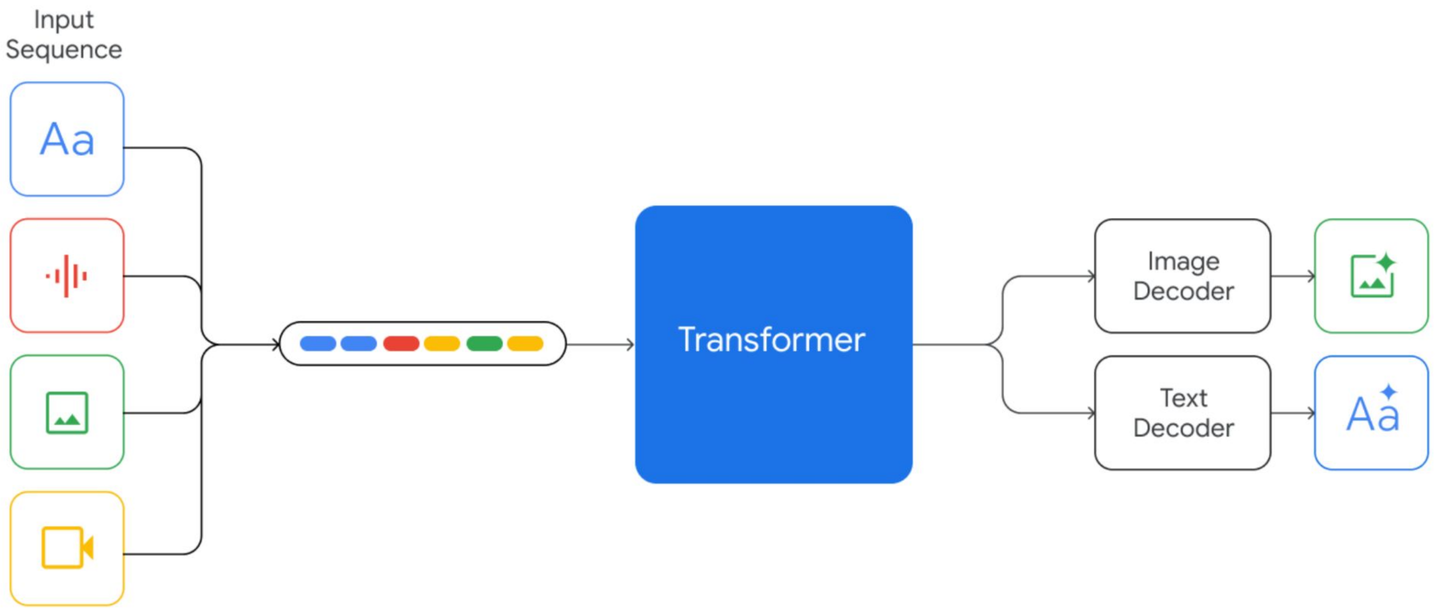
```
yStatus = false }) {

=== filterByOrg : true
ilterByStatus : true
```

```
lse
```

```
function filterStudies({ studies, filterByOrg = false, filterBy
  let filteredStudies = studies.filter(study => {
```

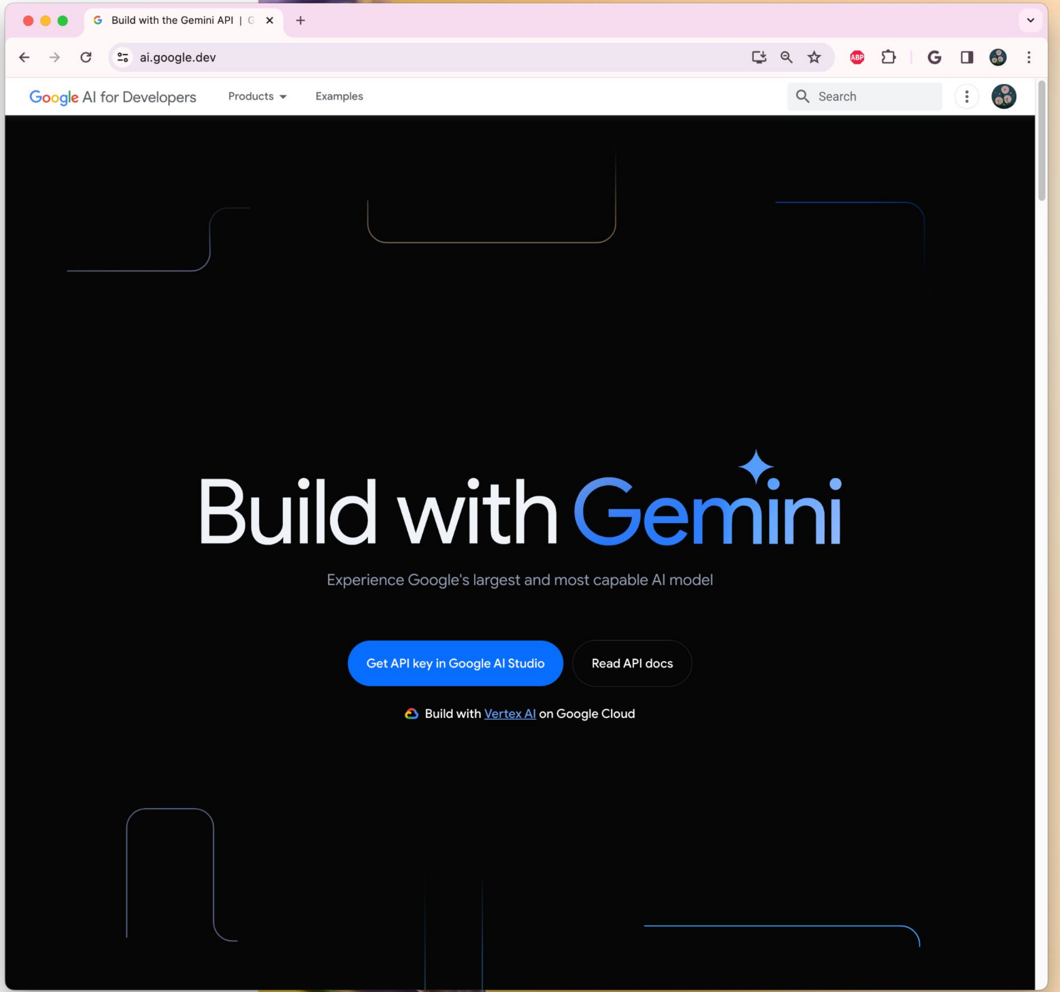
Gemini

Generalized Multimodal Intelligence Network



Prototyping with **Google AI Studio**

ai.google.dev



Build with the Gemini API | G

ai.google.dev

Build with Gemini

Experience Google's largest and most capable AI model

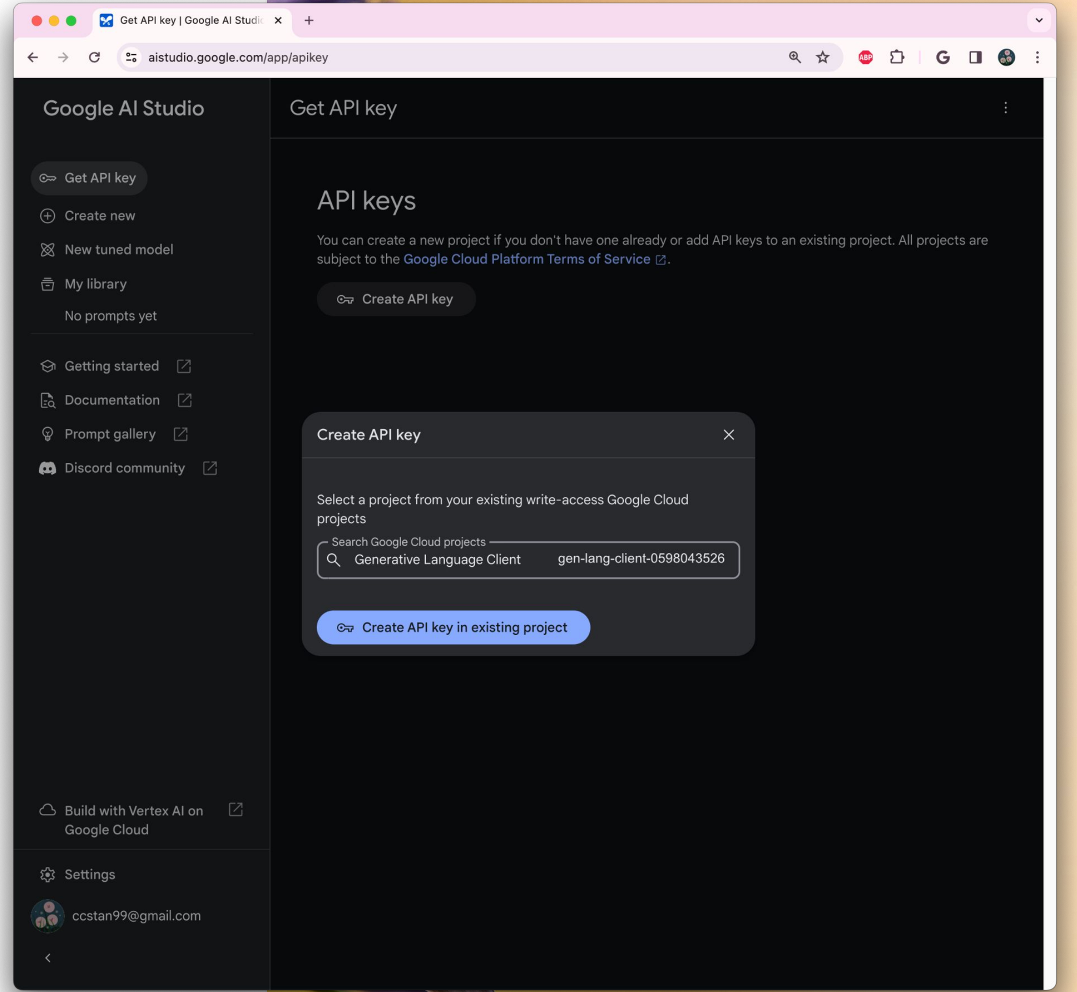
Get API key in Google AI Studio

Read API docs

Build with Vertex AI on Google Cloud

Get API Key

Treat like password



Get API key

API keys

You can create a new project if you don't have one already or add API keys to an existing project. All projects are subject to the Google Cloud Platform Terms of Service.

Create API key

Create API key

Select a project from your existing write-access Google Cloud projects

Search Google Cloud projects: Generative Language Client gen-lang-client-0598043526

Create API key in existing project

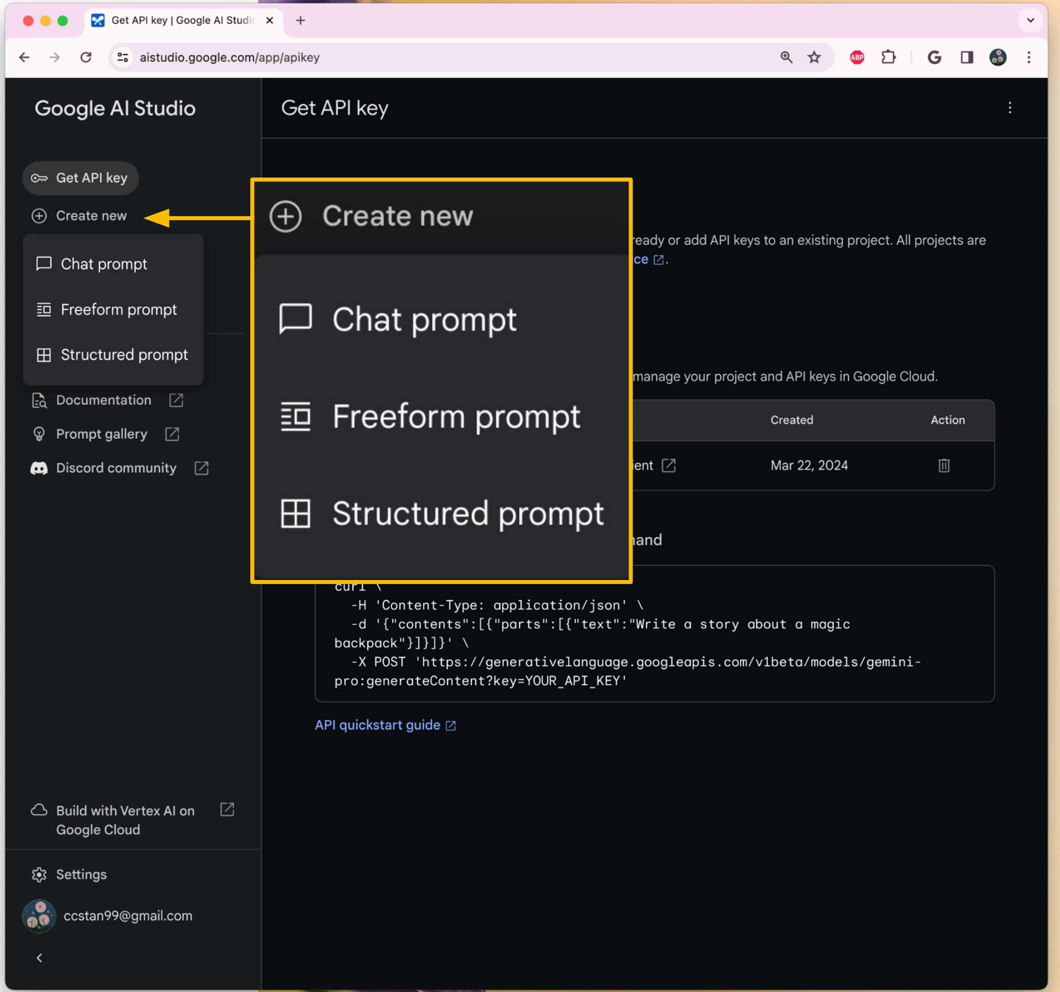
Google Developer Groups

Road to Google Developers Certification

Google Cloud

Create new

- Chat prompt
- Freeform prompt
- Structured prompt



ready or add API keys to an existing project. All projects are
ce.

manage your project and API keys in Google Cloud.

	Created	Action
ient	Mar 22, 2024	

and

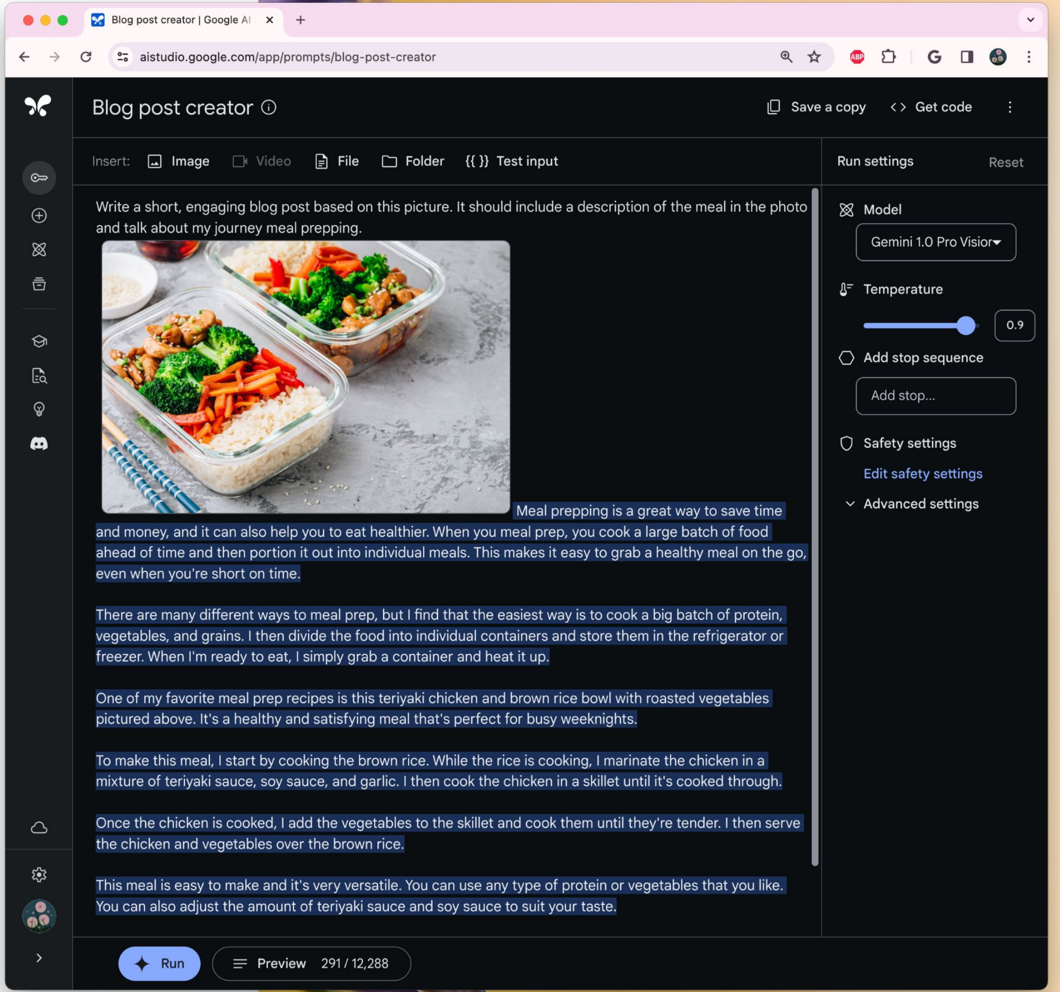
```
curl \
  -H 'Content-Type: application/json' \
  -d '{"contents":[{"parts":[{"text":"Write a story about a magic
backpack"}]}]}' \
  -X POST 'https://generativelanguage.googleapis.com/v1beta/models/gemini-
pro:generateContent?key=YOUR_API_KEY'
```

API quickstart guide

Freeform Prompt

Blog post creator

Write a prompt as text and image for the model to auto continue.



Model: Gemini 1.0 Pro Vision

Temperature: 0.9

291 / 12,288

Write a short, engaging blog post based on this picture. It should include a description of the meal in the photo and talk about my journey meal prepping.

Meal prepping is a great way to save time and money, and it can also help you to eat healthier. When you meal prep, you cook a large batch of food ahead of time and then portion it out into individual meals. This makes it easy to grab a healthy meal on the go, even when you're short on time.

There are many different ways to meal prep, but I find that the easiest way is to cook a big batch of protein, vegetables, and grains. I then divide the food into individual containers and store them in the refrigerator or freezer. When I'm ready to eat, I simply grab a container and heat it up.

One of my favorite meal prep recipes is this teriyaki chicken and brown rice bowl with roasted vegetables pictured above. It's a healthy and satisfying meal that's perfect for busy weeknights.

To make this meal, I start by cooking the brown rice. While the rice is cooking, I marinate the chicken in a mixture of teriyaki sauce, soy sauce, and garlic. I then cook the chicken in a skillet until it's cooked through.

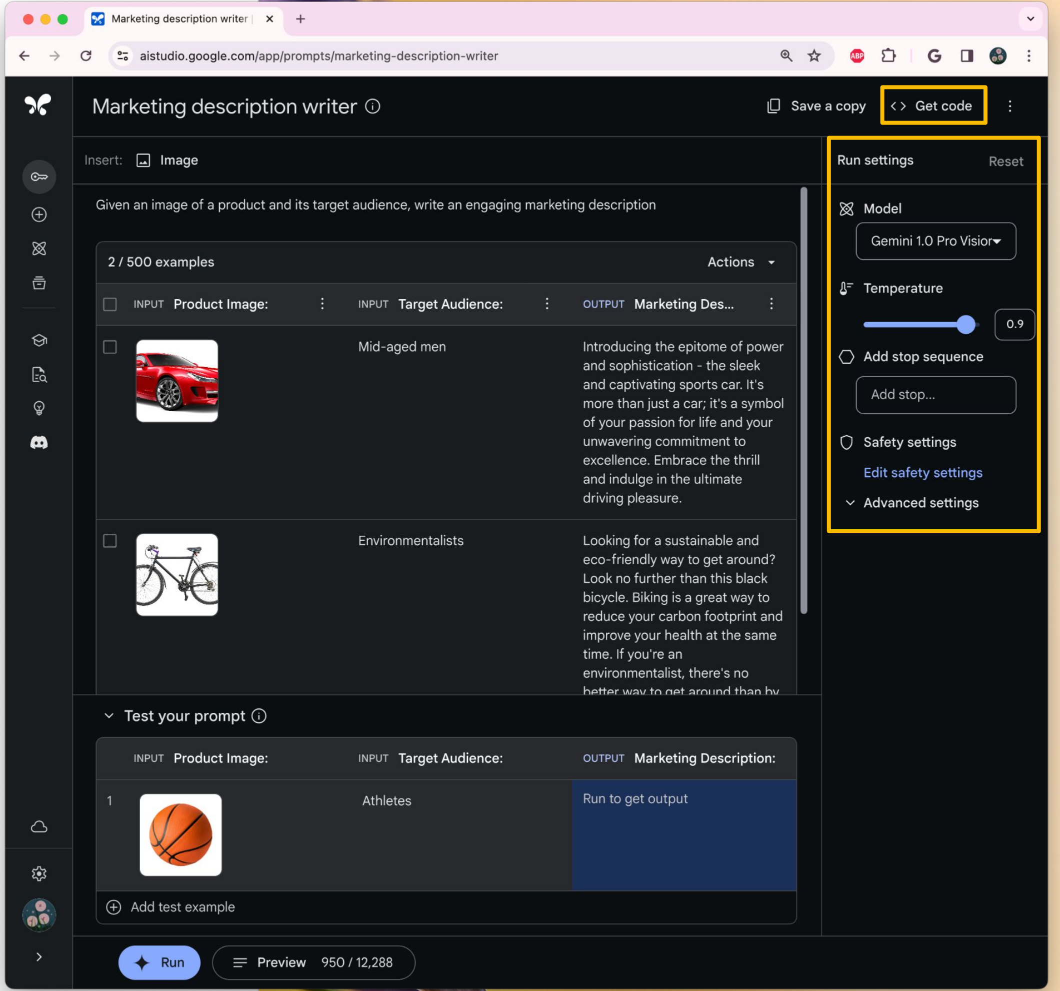
Once the chicken is cooked, I add the vegetables to the skillet and cook them until they're tender. I then serve the chicken and vegetables over the brown rice.

This meal is easy to make and it's very versatile. You can use any type of protein or vegetables that you like. You can also adjust the amount of teriyaki sauce and soy sauce to suit your taste.

Structured Prompt

Marketing description

Table-based interface for more complex model priming and prompting



Marketing description writer |
aistudio.google.com/app/prompts/marketing-description-writer
Save a copy
Get code
Insert: Image

Marketing description writer

Given an image of a product and its target audience, write an engaging marketing description

2 / 500 examples Actions

INPUT Product Image:	INPUT Target Audience:	OUTPUT Marketing Des...
	Mid-aged men	Introducing the epitome of power and sophistication - the sleek and captivating sports car. It's more than just a car; it's a symbol of your passion for life and your unwavering commitment to excellence. Embrace the thrill and indulge in the ultimate driving pleasure.
	Environmentalists	Looking for a sustainable and eco-friendly way to get around? Look no further than this black bicycle. Biking is a great way to reduce your carbon footprint and improve your health at the same time. If you're an environmentalist, there's no better way to get around than by

Test your prompt

	INPUT Product Image:	INPUT Target Audience:	OUTPUT Marketing Description:
1		Athletes	Run to get output

Add test example

Run settings Reset

Model: Gemini 1.0 Pro Vision

Temperature: 0.9

Add stop sequence: Add stop...

Safety settings: Edit safety settings

Advanced settings

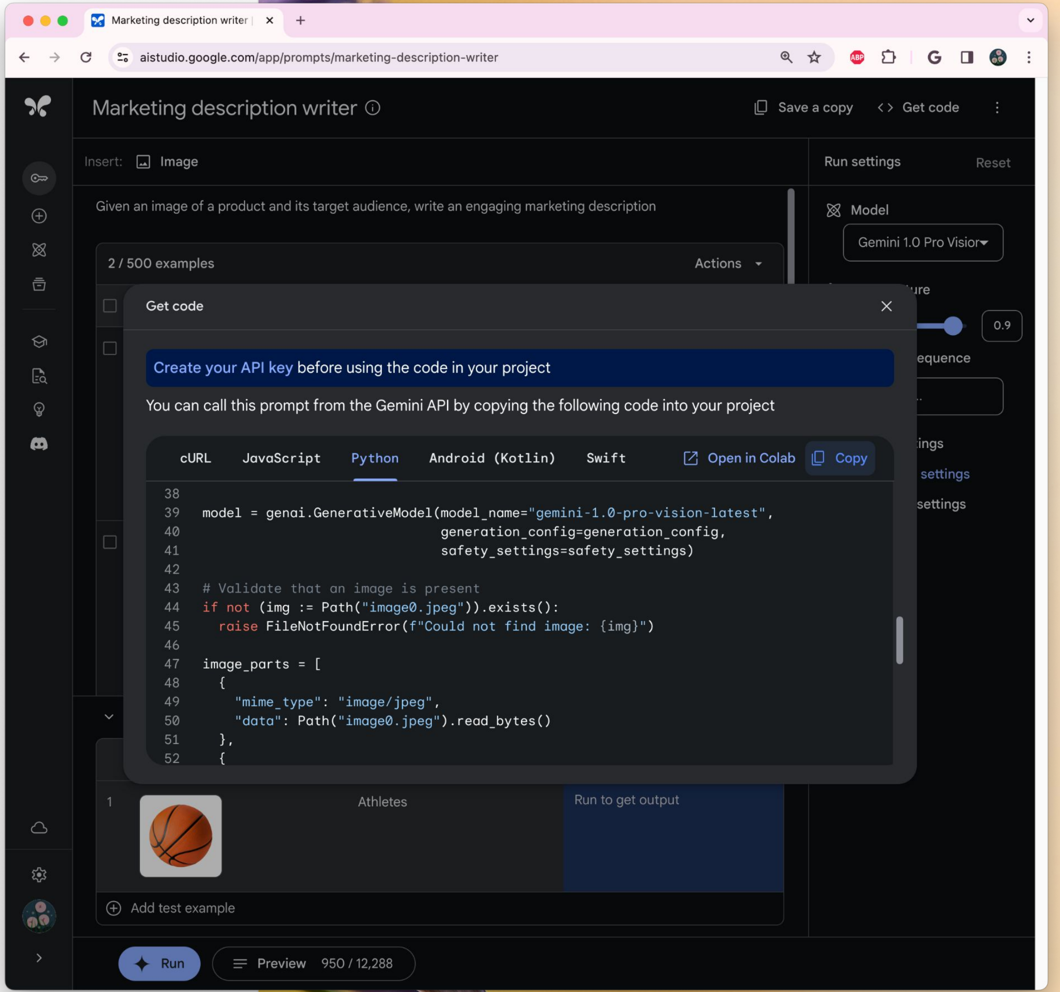
Google Developer Groups

Road to Google Developers Certification

Google Cloud

Get Code

- Choose Language
- Open in Colab
- Copy to Editor



Given an image of a product and its target audience, write an engaging marketing description

Get code

Create your API key before using the code in your project

You can call this prompt from the Gemini API by copying the following code into your project

```python

model = genai.GenerativeModel(model_name="gemini-1.0-pro-vision-latest",
                              generation_config=generation_config,
                              safety_settings=safety_settings)

# Validate that an image is present
if not (img := Path("image0.jpeg")).exists():
  raise FileNotFoundError(f"Could not find image: {img}")

image_parts = [
  {
    "mime_type": "image/jpeg",
    "data": Path("image0.jpeg").read_bytes()
  },
  {
```

Google Developer Groups

Road to Google Developers Certification

Google Cloud

Settings

Tokens

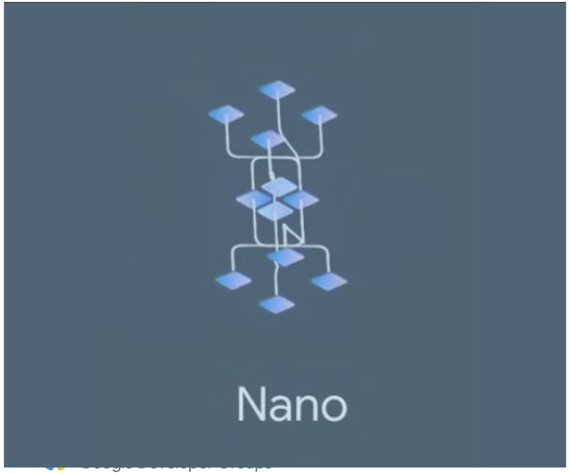
- Words or subwords
- Different LLM tokenizers
- Training data, context window

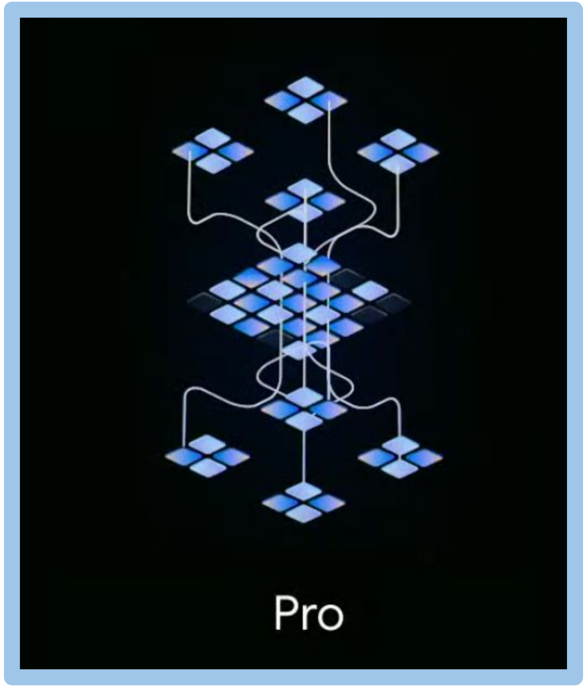
Temperature

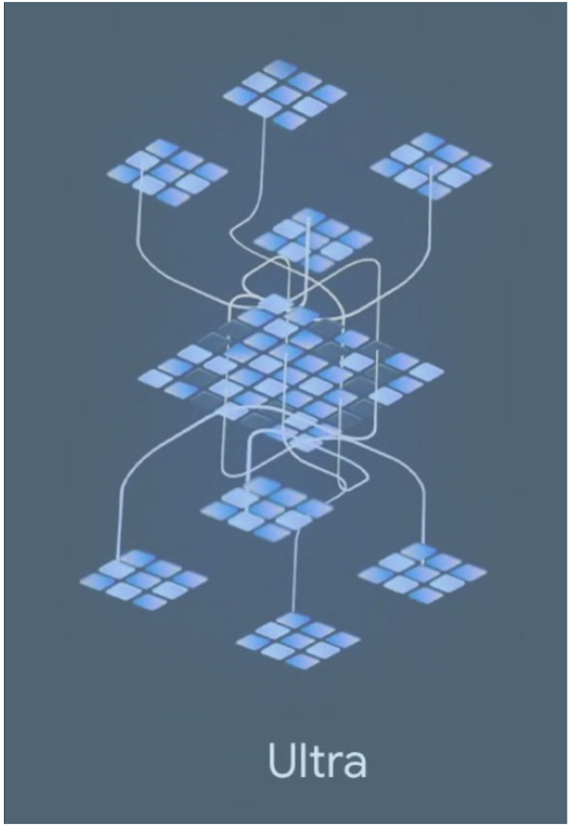
- Selected by probability
- Between 0 to 1.0
- Diversity or “creativity”

Settings

Model Sizes





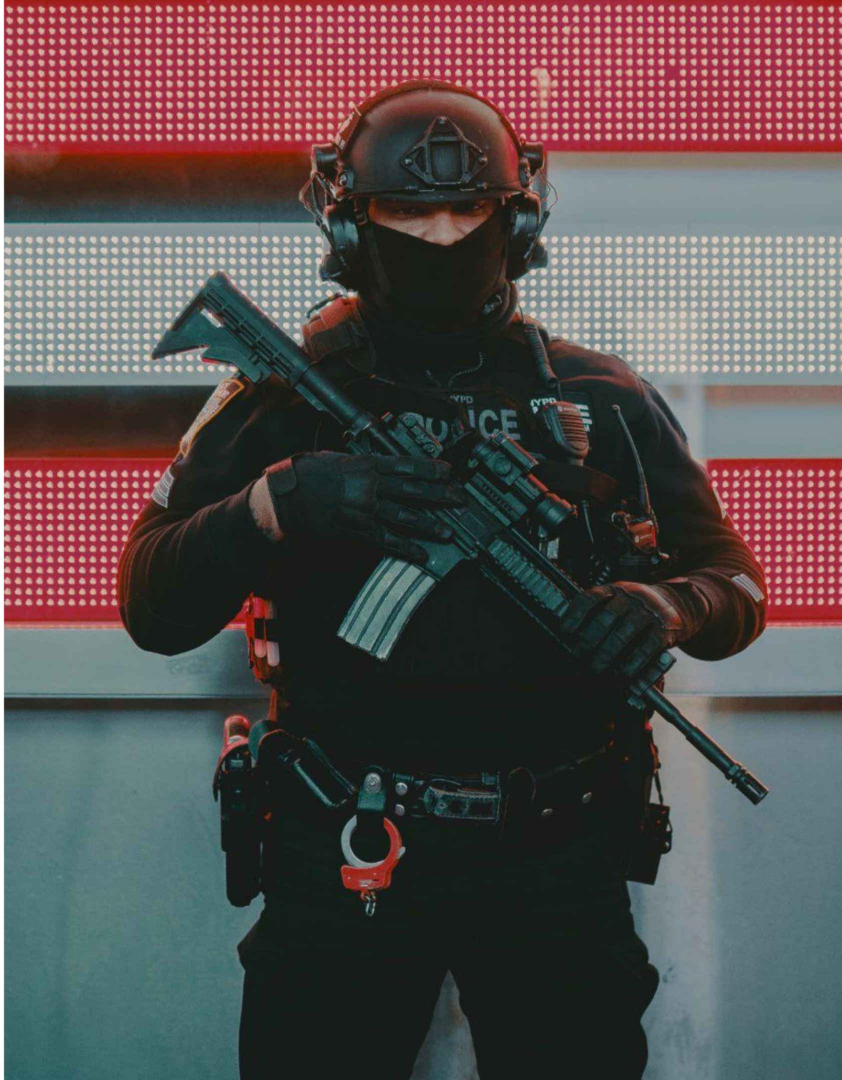


Settings

Safety Ratings

Harm Categories

- Harassment
- Hate Speech
- Sexually Explicit
- Dangerous Content



Settings

Safety Ratings

Harm Categories

- Harassment
- Hate Speech
- Sexually Explicit
- Dangerous Content

Harm Probabilities

- NEGLIGIBLE
- LOW
- MEDIUM
- HIGH

> The hottest new programming language is English.
>
> Andrej Karpathy
> OpenAI

```
function filterStudies({ studies, filterByOrg = false, filte
                studies.filter(study => {
                       and organizat
```

Prompt Engineering

- Clear & Specific Instructions
- Give Examples
- Step by Step

REST APIs

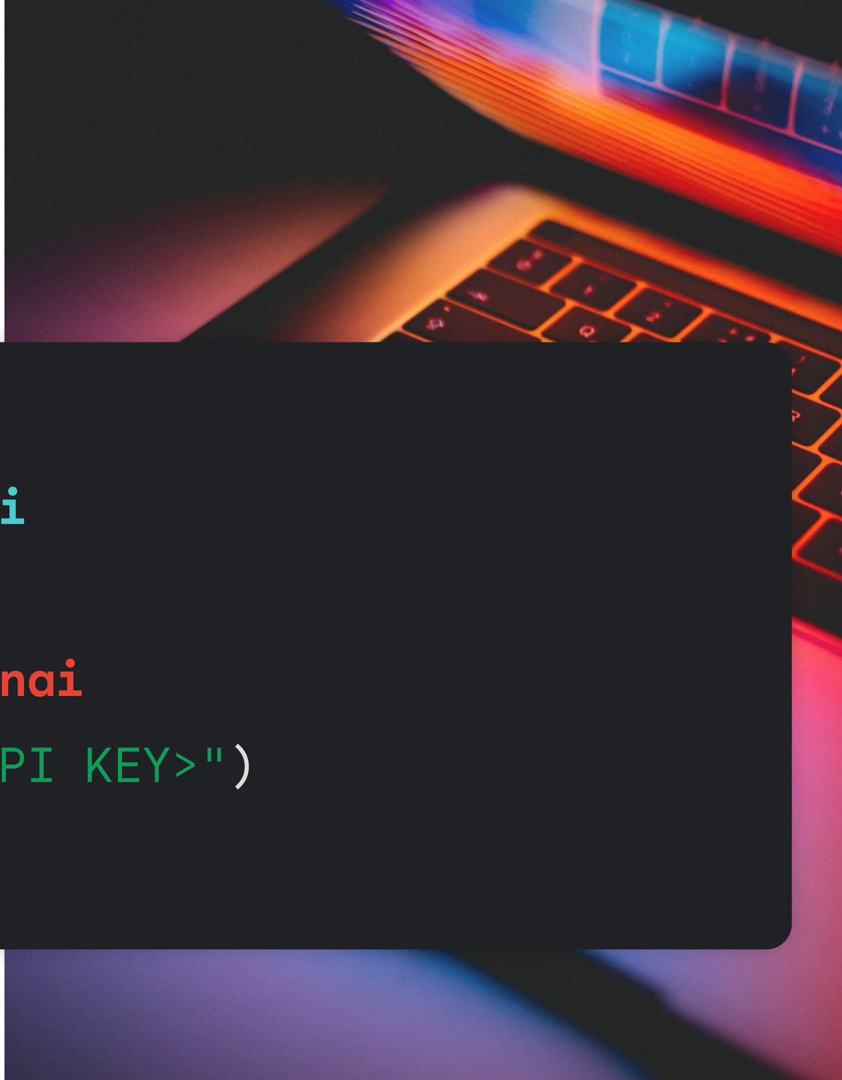
Client libraries for

- Python
- JavaScript
- Android (Kotlin)
- Swift
- cURL

Setup

Install & import libraries

```
$ pip install google-generativeai


import google.generativeai as genai
genai.configure(api_key="<YOUR API KEY>")
```



Generate Text

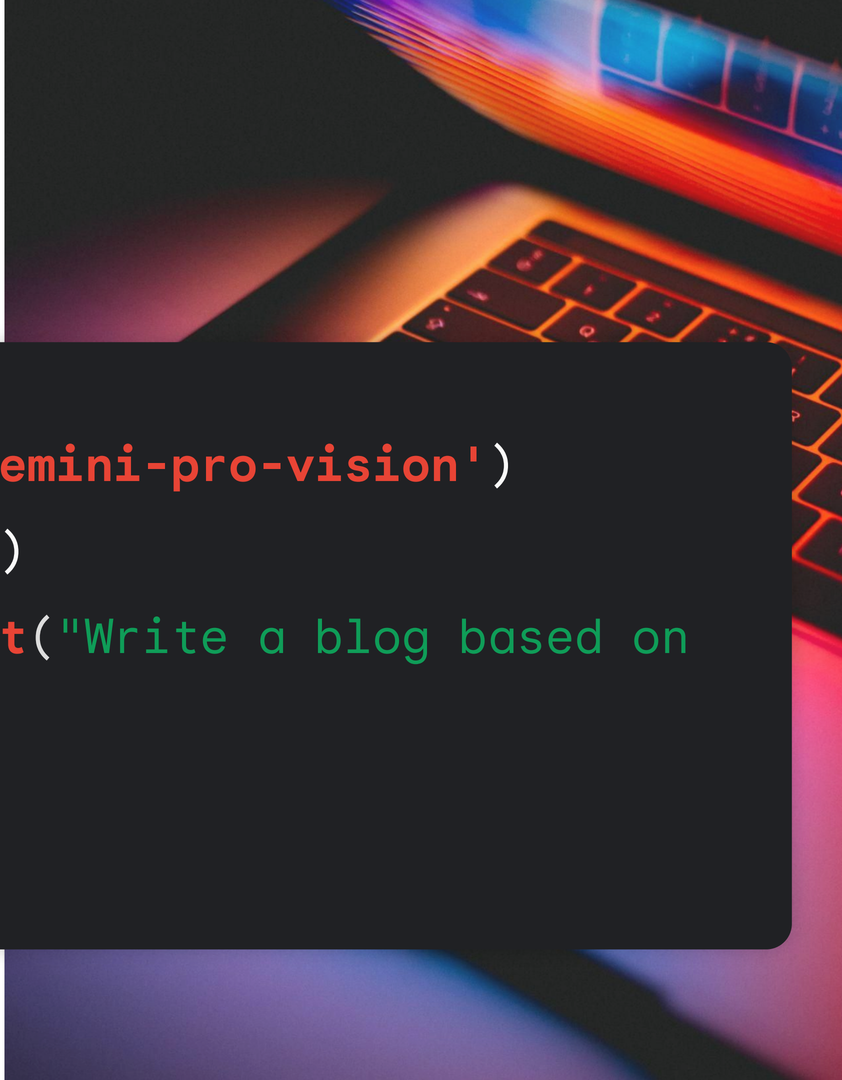
Text only prompt

```python
model = genai.GenerativeModel('gemini-pro')

response = model.generate_content("Write a story about a boy and a backpack.")
print(response.text)
```

Generate Text

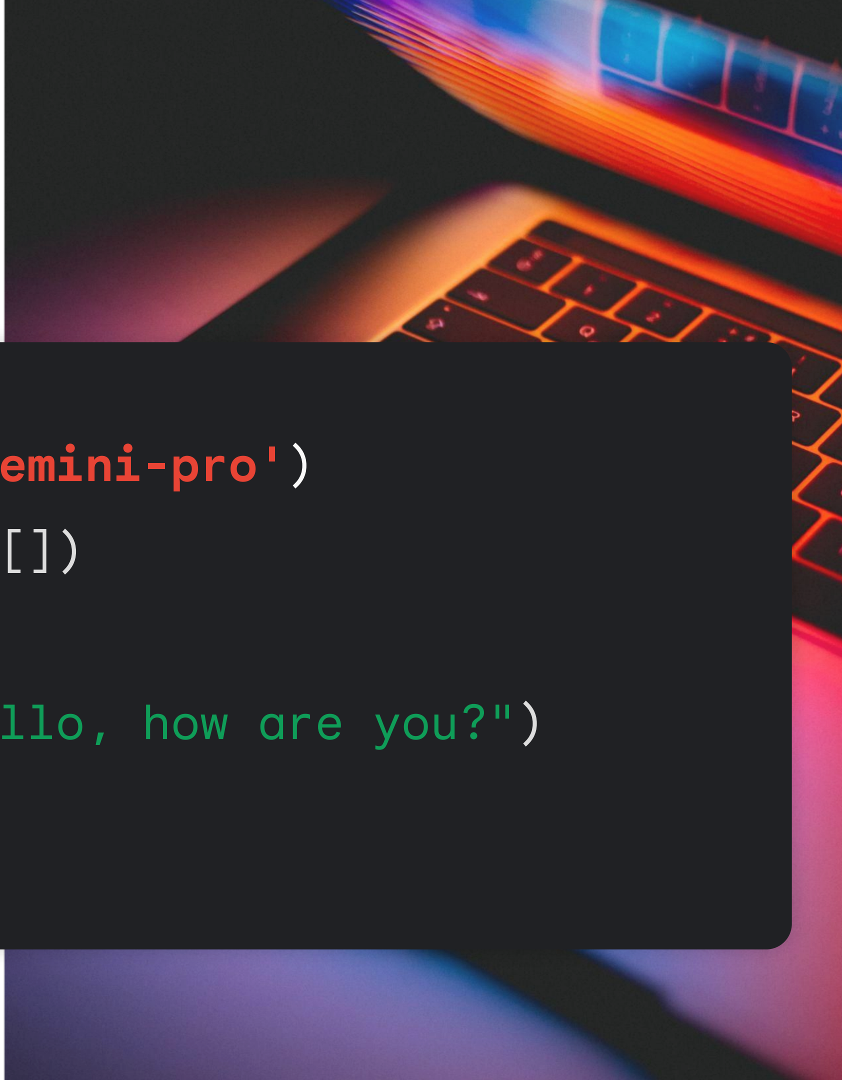
Text and image prompt

```
model = genai.GenerativeModel('gemini-pro-vision')
img = PIL.Image.open('image.jpg')
response = model.generate_content("Write a blog based on
this photo.", img)
print(response.text)
```



Chat Conversations

For interactive applications

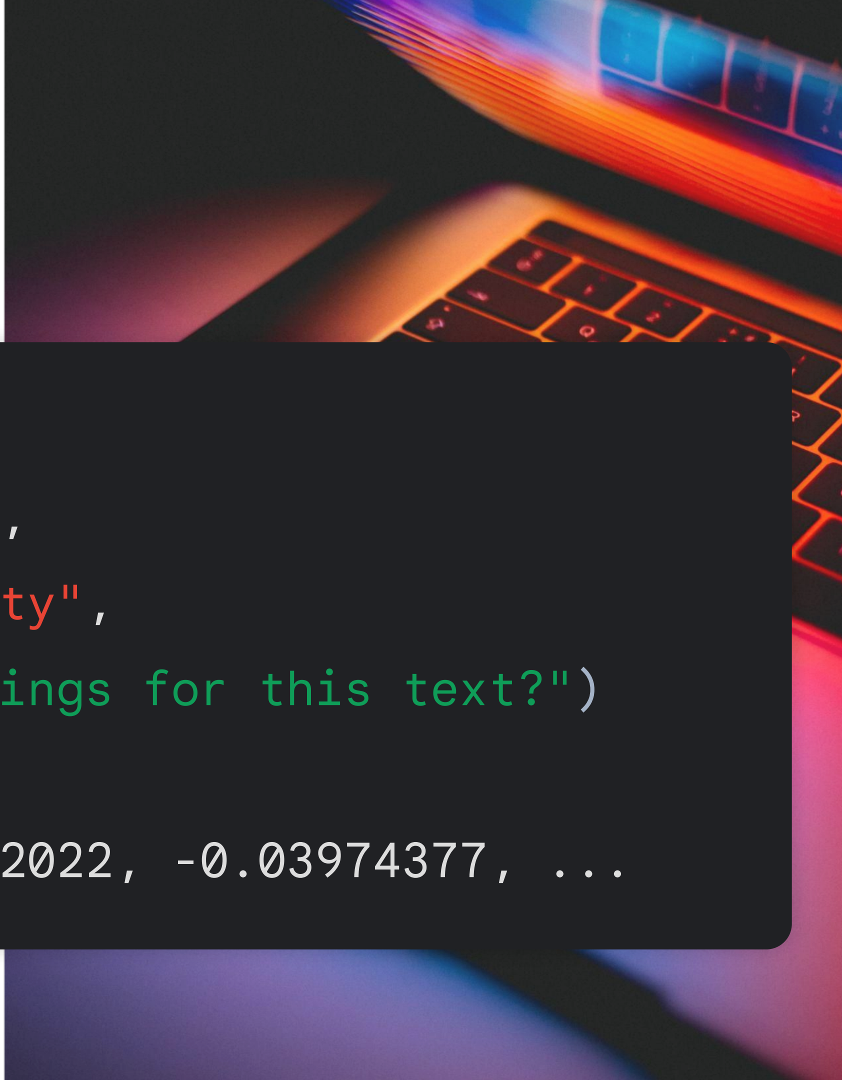


```python
model = genai.GenerativeModel('gemini-pro')
chat = model.start_chat(history=[])

response = chat.send_message("Hello, how are you?")
print(response.text)
```

Embeddings Endpoint

Develop new applications

```
result = genai.embed_content(
  model = "models/embedding-001",
  task_type = "semantic_similarity",
  content = "What are the embeddings for this text?")
print(result['embedding'])
# output: [0.019084517, -0.012082022, -0.03974377, ...
```

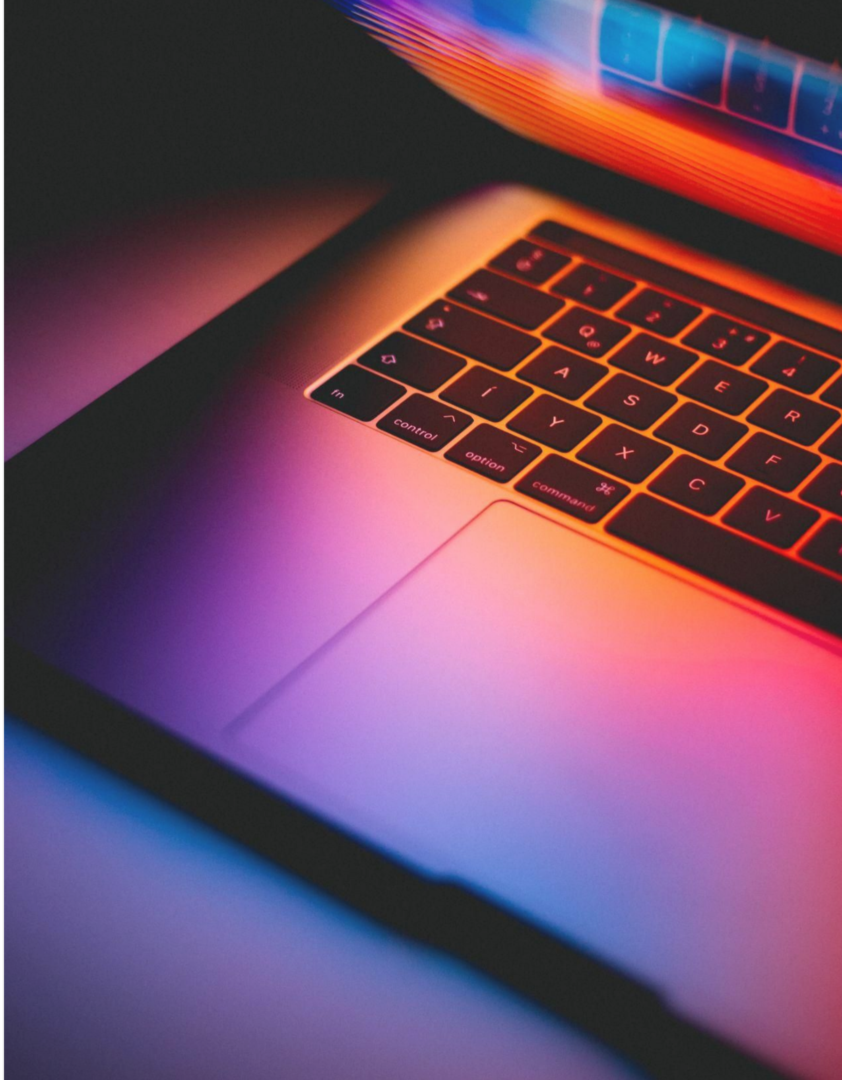


Embeddings Endpoint

Develop new applications

Valid values for task_type

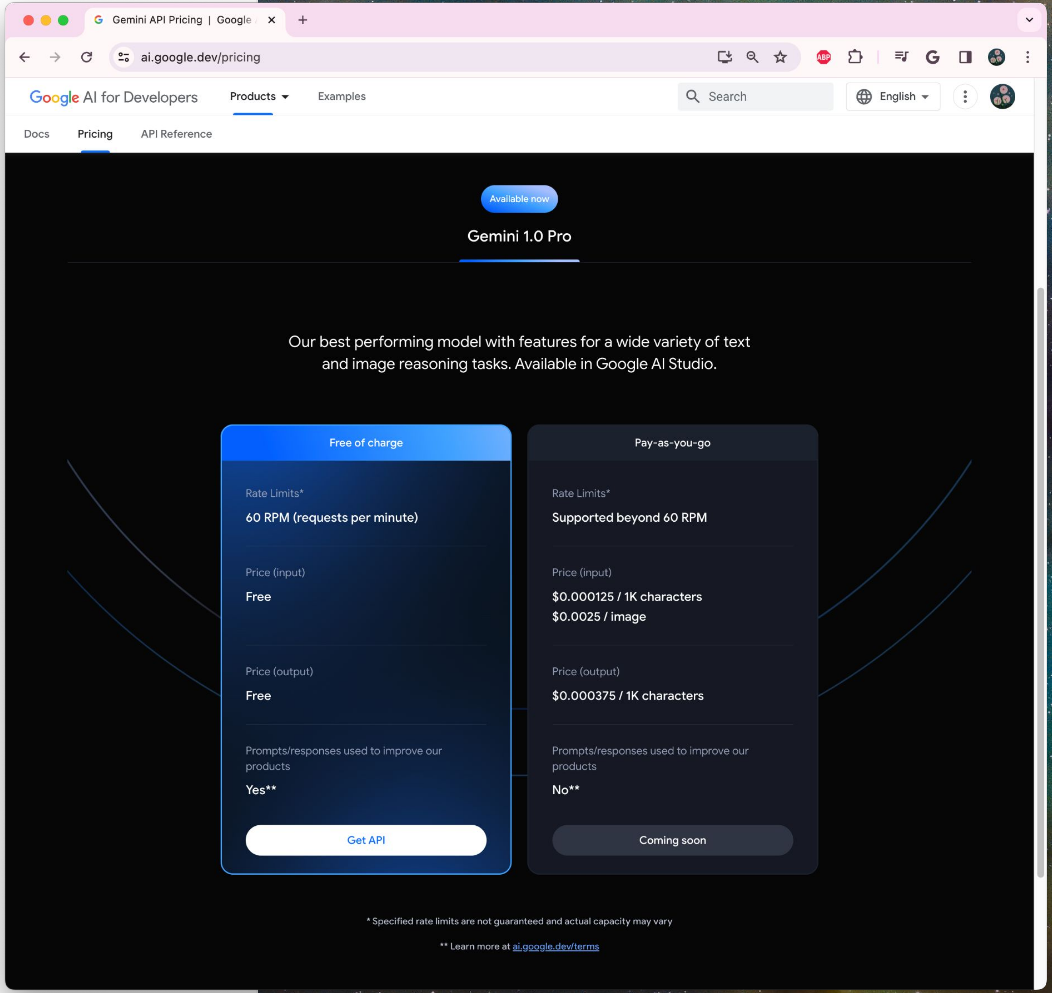
- RETRIEVAL_QUERY
- RETRIEVAL_DOCUMENT
- SEMANTIC_SIMILARITY
- CLASSIFICATION
- CLUSTERING



Gemini

Free for now.
Try it out!

ai.google.dev



Gemini 1.0 Pro

Our best performing model with features for a wide variety of text and image reasoning tasks. Available in Google AI Studio.

	Free of charge	Pay-as-you-go
Rate Limits*	60 RPM (requests per minute)	Supported beyond 60 RPM
Price (input)	Free	$0.000125 / 1K characters $0.0025 / image
Price (output)	Free	$0.000375 / 1K characters
Prompts/responses used to improve our products	Yes**	No**

* Specified rate limits are not guaranteed and actual capacity may vary

** Learn more at ai.google.dev/terms

Google Developer Groups

Road to Google Developers Certification

Google Cloud

Recommendation Systems

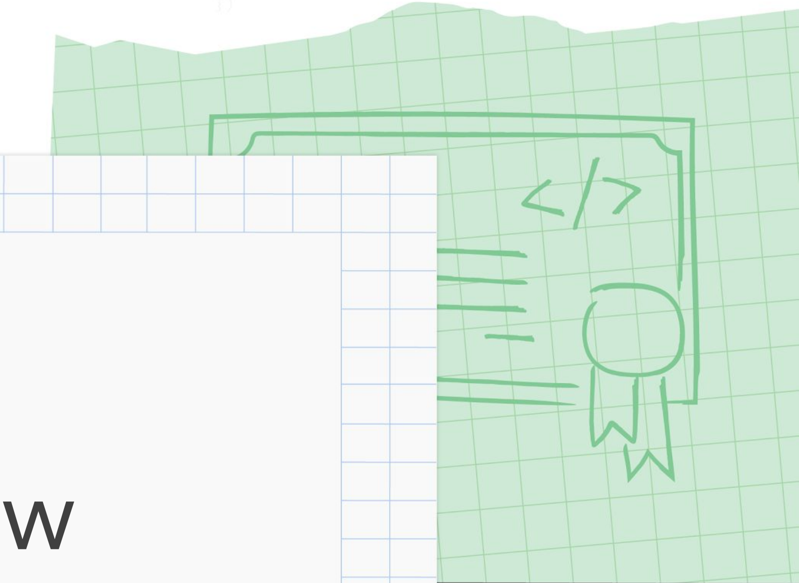
- Content-based
- Collaborative Filtering

Google Developer Groups

Road to Google Developers Certification

Google Cloud

Week 5 Content Review



```
nction filterStudies({ studies, filt
  let filteredStudies = studies.filt
    const matchOrg = filterByOrg ?
    const matchStatus = filterBySta
    if (matchOrg && matchStatus) {
      return true
    }
    return false
  })
```

```
Status = false }) {

=== filterByOrg : true
ilterByStatus : true
```

```
lse
```

```
function filterStudies({ studies, filterByOrg = false, filterBy
  let filteredStudies = studies.filter(study => {
```

Week 5

Study Group

Google Developer Groups

Road to Google Developers Certification

Google Cloud

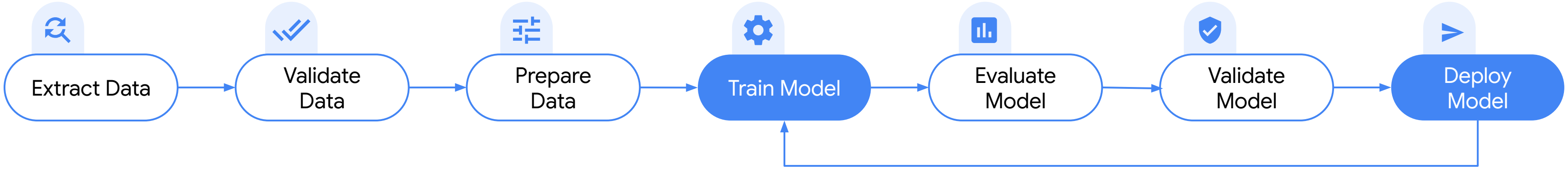
ML Pipeline Automation & Orchestration

- Design pipeline.
- Implement training pipeline.
- Implement serving pipeline.
- Track and audit metadata.
- Use CI/CD to test and deploy models.

ML Solution Monitoring, Optimization, and Maintenance

- Monitor ML solutions.
- Troubleshoot ML solutions.
- Tune performance of ML solutions for training & serving in production.

Pipelines automate the training and deployment of models



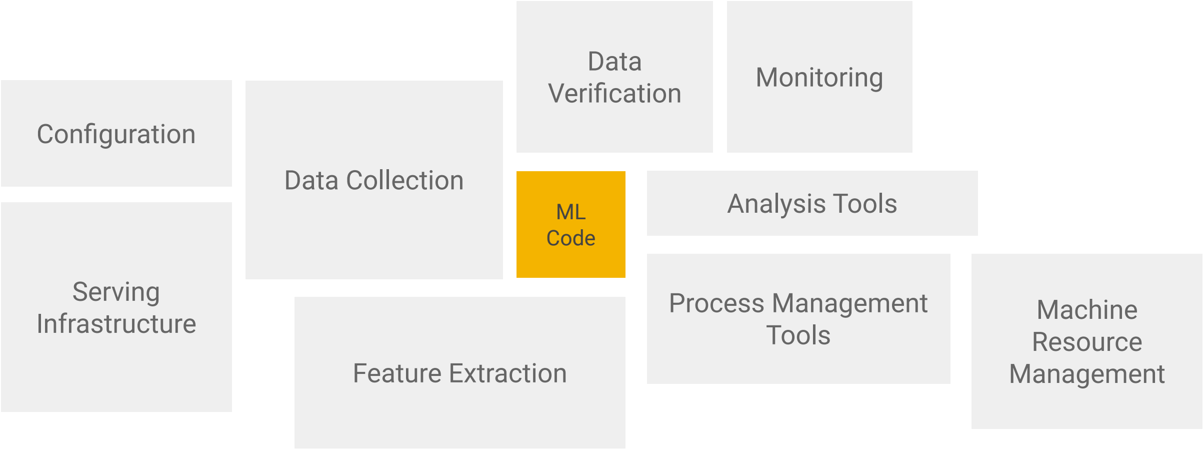
From DevOps to MLOps

- **Continuous Integration (CI)** is no longer only about testing and validating code and components, but also testing and validating data, data schemas, and models.
- **Continuous Deployment (CD)** is no longer about a single software package or a service, but a system (ML training pipeline) that should automatically deploy another service (model prediction service).
- **Continuous Training** is a new property, specific to ML systems, concerning automatically retraining and serving the models.

In addition to the actual ML...

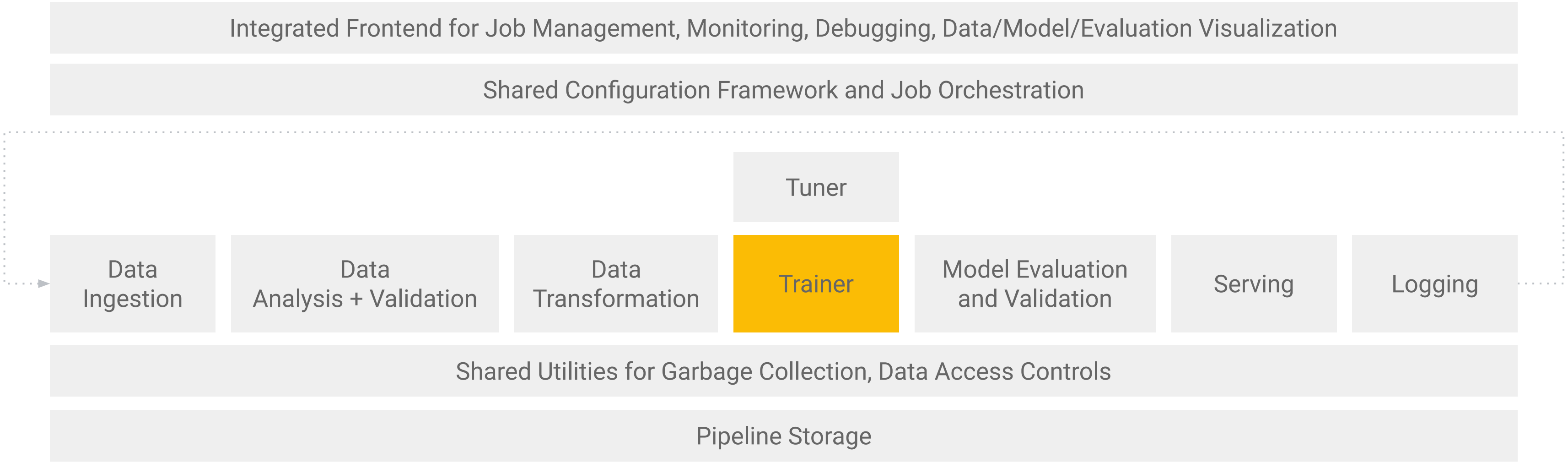
ML
Code

...you have to worry about so much more.



Source: Sculley et al.: Hidden Technical Debt in Machine Learning Systems

TFX is the solution to this problem



Integrated Frontend for Job Management, Monitoring, Debugging, Data/Model/Evaluation Visualization

Shared Configuration Framework and Job Orchestration

Tuner

Data Ingestion

Data Analysis + Validation

Data Transformation

Trainer

Model Evaluation and Validation

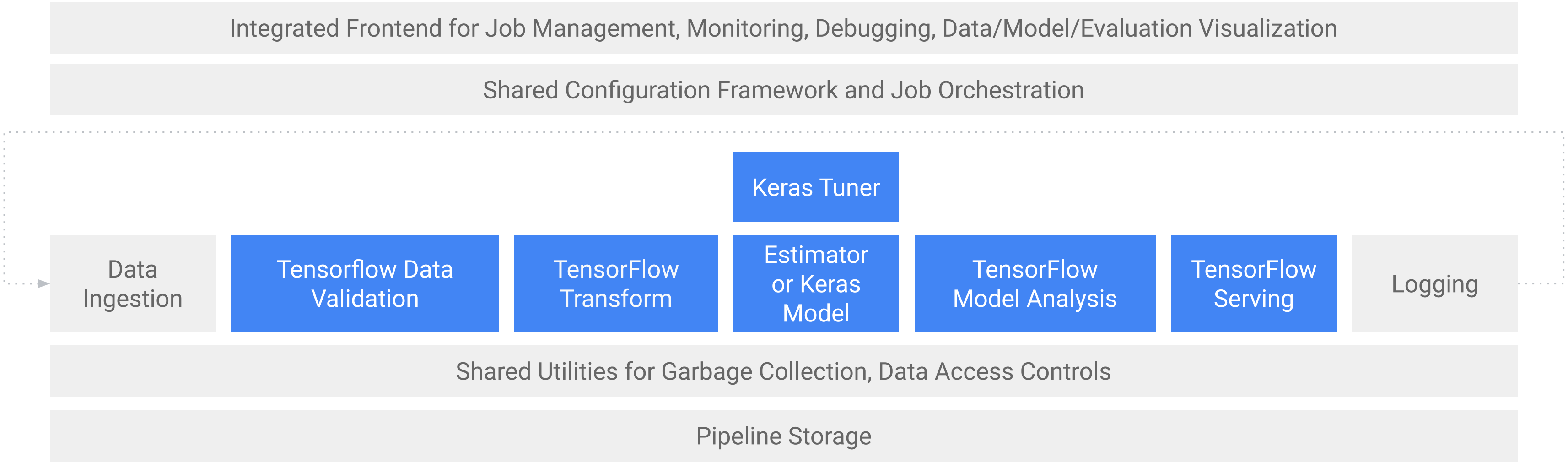
Serving

Logging

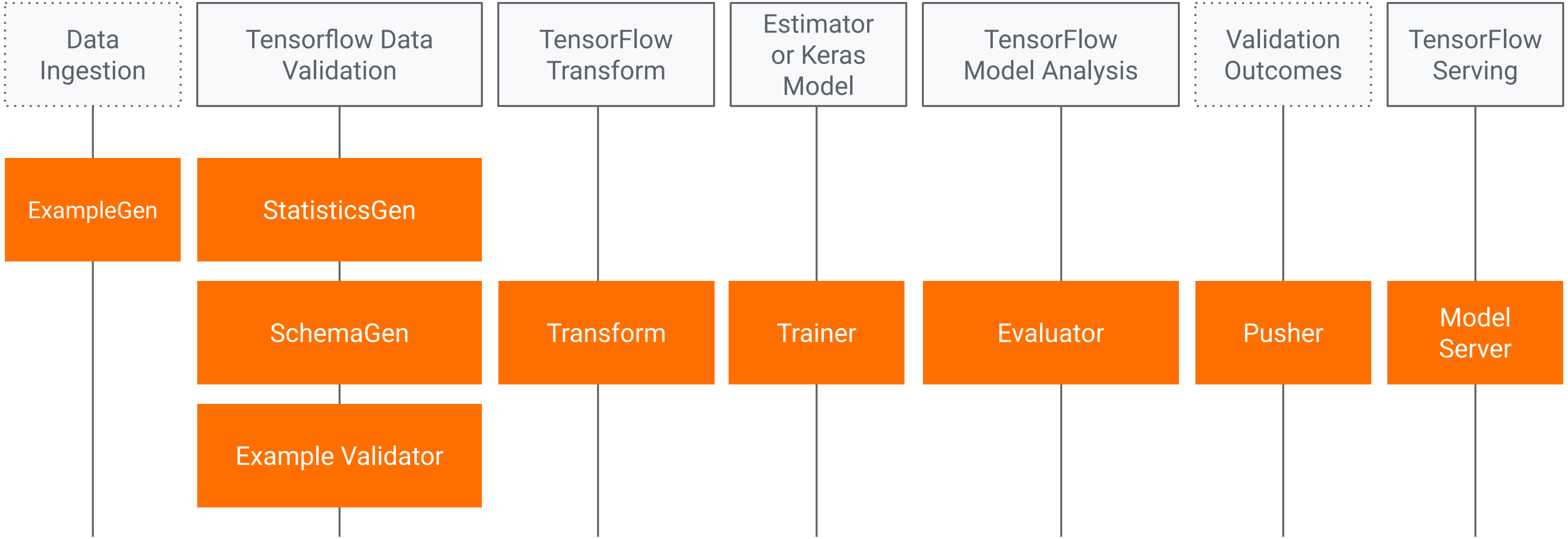
Shared Utilities for Garbage Collection, Data Access Controls

Pipeline Storage

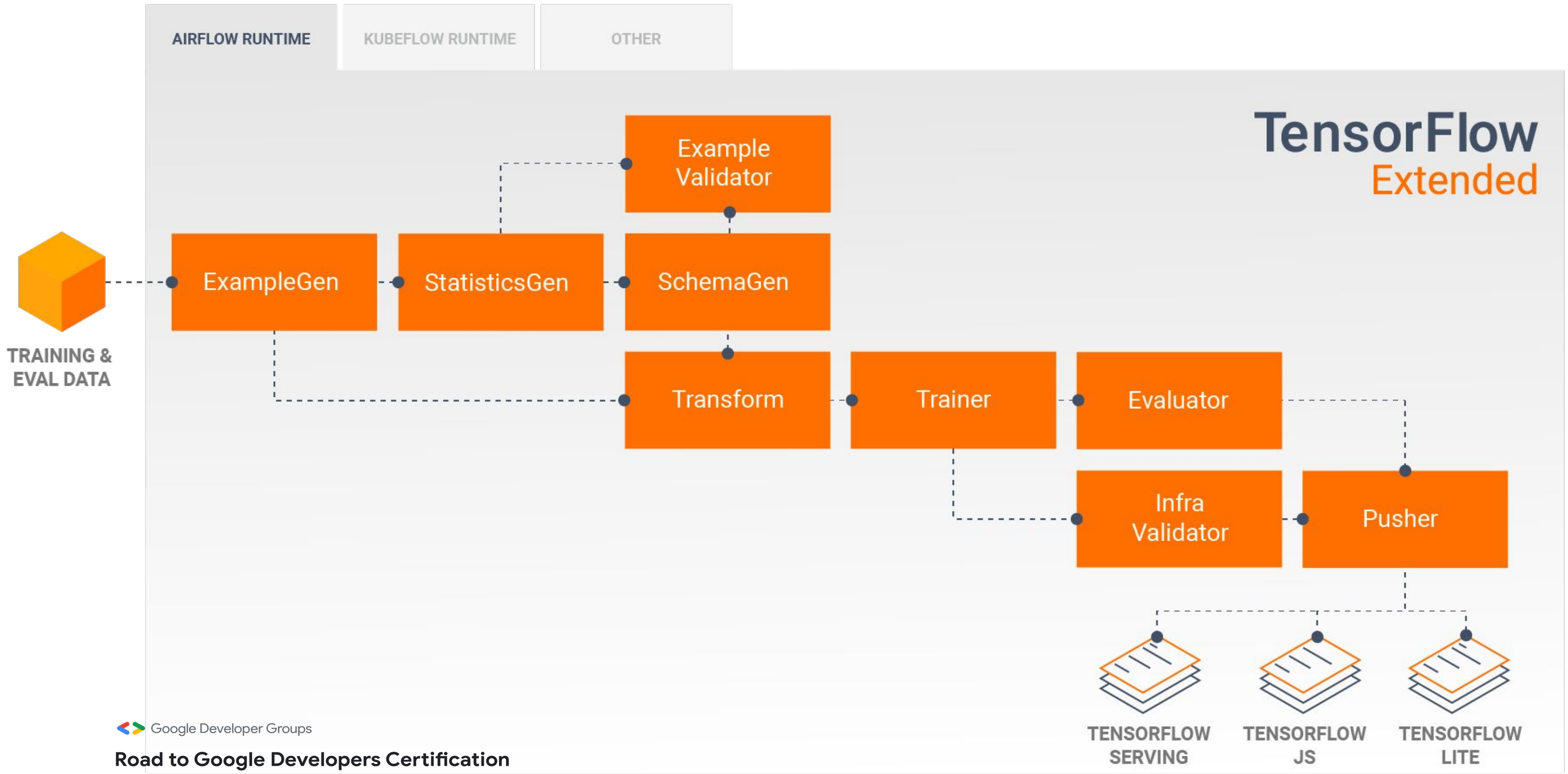
TFX is the solution to this problem

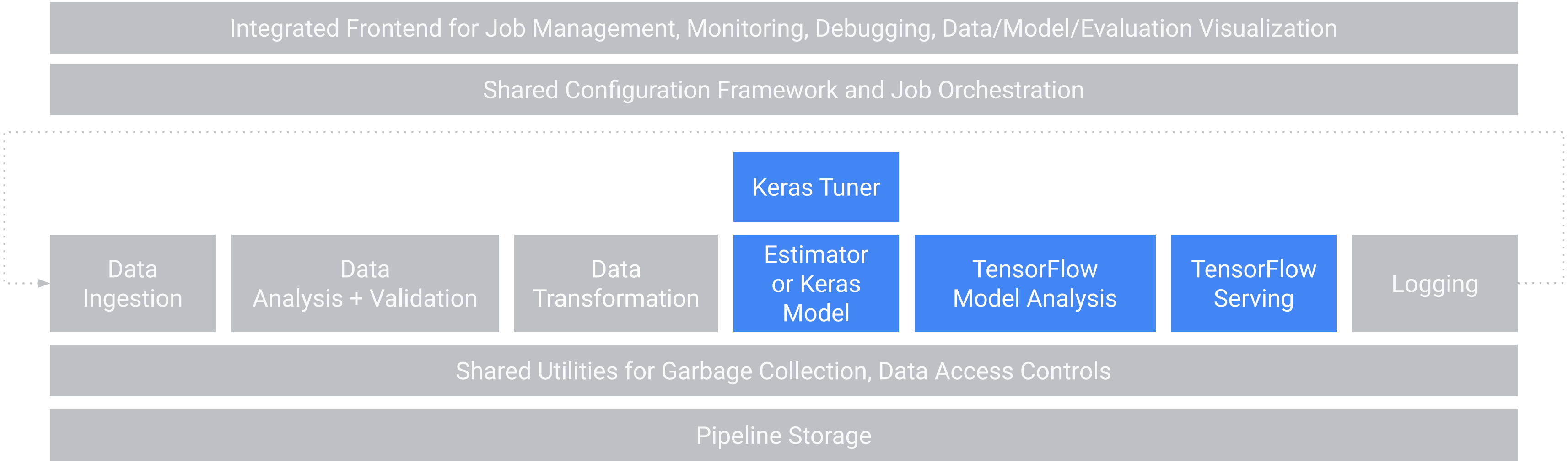

Integrated Frontend for Job Management, Monitoring, Debugging, Data/Model/Evaluation Visualization
Shared Configuration Framework and Job Orchestration
Keras Tuner
Data Ingestion
Tensorflow Data Validation
TensorFlow Transform
Estimator or Keras Model
TensorFlow Model Analysis
TensorFlow Serving
Logging
Shared Utilities for Garbage Collection, Data Access Controls
Pipeline Storage

TFX Components

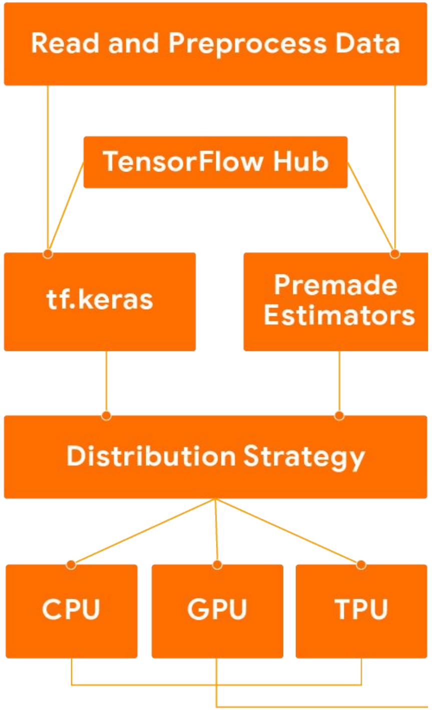


Data Ingestion	Tensorflow Data Validation	TensorFlow Transform	Estimator or Keras Model	TensorFlow Model Analysis	Validation Outcomes	TensorFlow Serving
ExampleGen	StatisticsGen					
	SchemaGen	Transform	Trainer	Evaluator	Pusher	Model Server
	Example Validator					


AIRFLOW RUNTIME
KUBEFLOW RUNTIME
OTHER
TensorFlow
Extended
TRAINING & EVAL DATA
ExampleGen
StatisticsGen
SchemaGen
Example Validator
Transform
Trainer
Evaluator
Infra Validator
Pusher
TENSORFLOW SERVING
TENSORFLOW JS
TENSORFLOW LITE

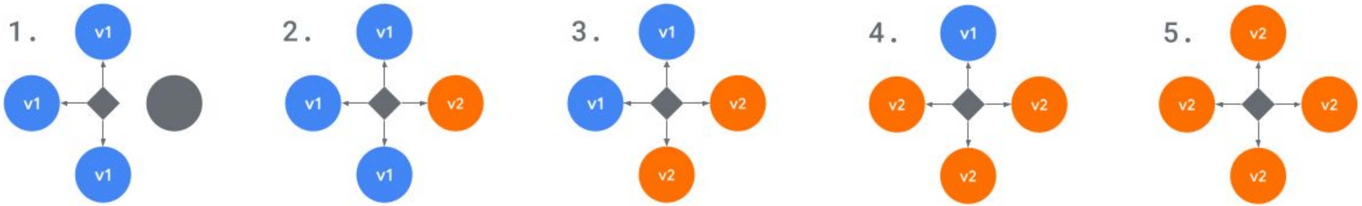

Integrated Frontend for Job Management, Monitoring, Debugging, Data/Model/Evaluation Visualization
Shared Configuration Framework and Job Orchestration
Keras Tuner
Data Ingestion
Data Analysis + Validation
Data Transformation
Estimator or Keras Model
TensorFlow Model Analysis
TensorFlow Serving
Logging
Shared Utilities for Garbage Collection, Data Access Controls
Pipeline Storage

TRAINING

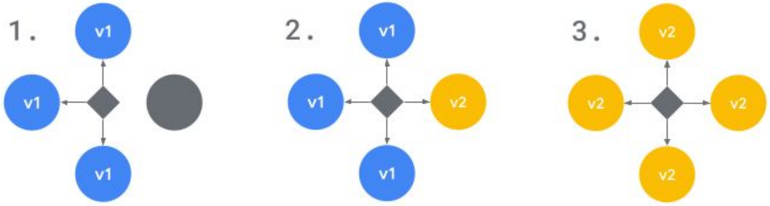

Read and Preprocess Data
TensorFlow Hub
tf.keras
Premade Estimators
Distribution Strategy
CPU
GPU
TPU

Deployment Strategies

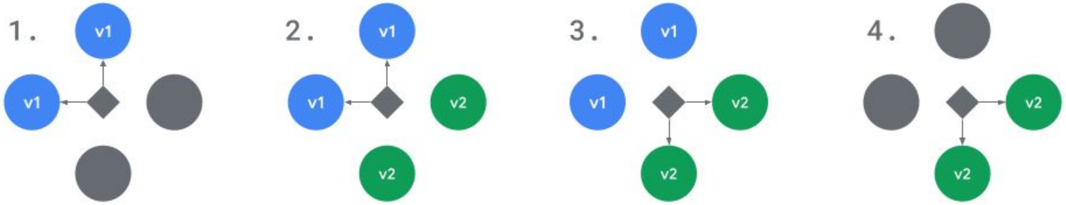
Rolling Update



Canary Deployment

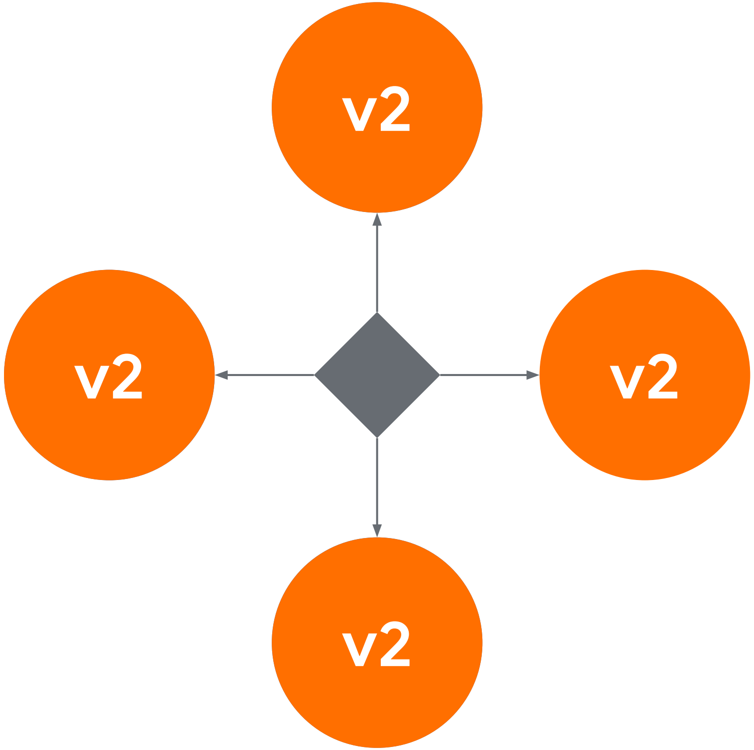


Blue-Green Deployment



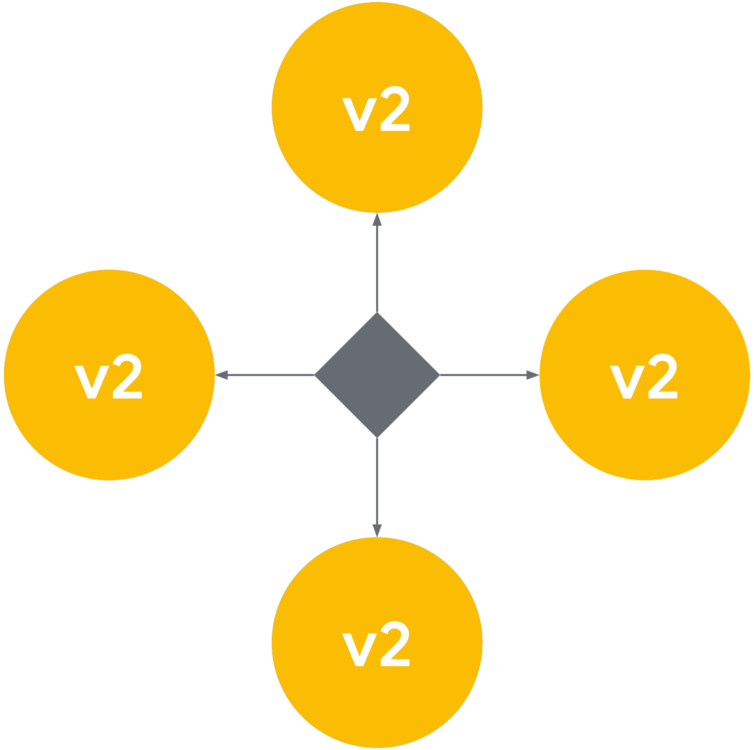
Rolling Update

- Slowly increasing number of NEW
- Slowly decreasing number of OLD



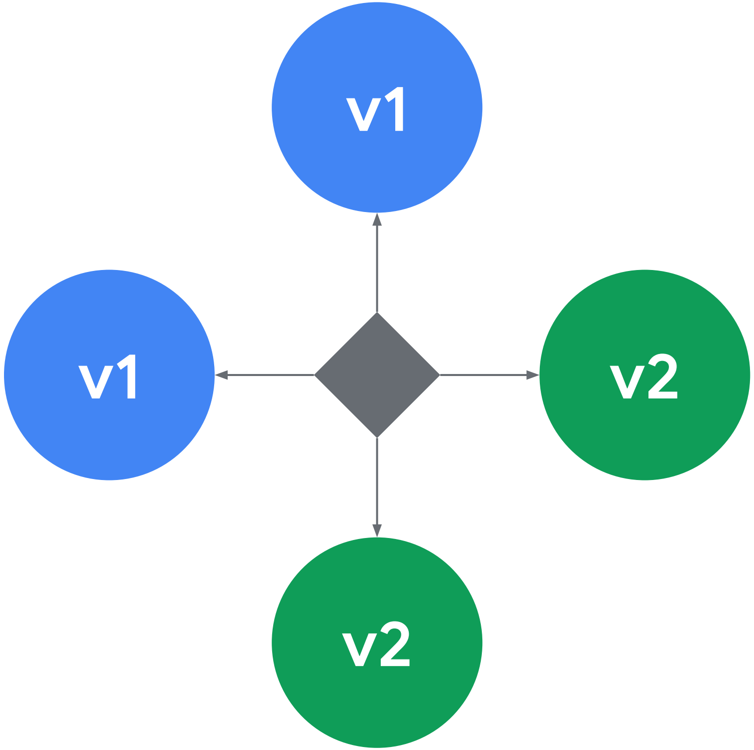
Canary Deployment

- Test on small SUBSET in production
- After testing, traffic shifts to NEW



Blue-Green Deployment

- 2 SEPARATE deployments
- After testing, switch traffic from OLD Blue to NEW Green.



Sample Questions Review

```
nction filterStudies({ studies, filt
  let filteredStudies = studies.filt
    const matchOrg = filterByOrg ?
    const matchStatus = filterBySta
    if (matchOrg && matchStatus) {
      return true
    }
    return false
  })
```

```
yStatus = false }) {

 === filterByOrg : true
ilterByStatus : true
```

```
lse              function filterStudies({ studies, filterByOrg = false, filterBy
                   let filteredStudies = studies.filter(study => {
```

Which of the following are benefit(s) of running TFX on Google Cloud? Select all that apply?

A. Vertex Pipelines is the only supported orchestrator for TFX pipelines

B. Simplify scaling of TFX pipeline data processing as your data grows

C. Automate your ML operational processes for individual and multiple ML pipelines.

D. Increase your pipeline development and experimentation velocity.

Which of the following are benefit(s) of running TFX on Google Cloud? Select all that apply?

A. Vertex Pipelines is the only supported orchestrator for TFX pipelines

B. Simplify scaling of TFX pipeline data processing as your data grows

C. Automate your ML operational processes for individual and multiple ML pipelines.

D. Increase your pipeline development and experimentation velocity.

In addition to CI/CD practiced by DevOps teams, MLOps introduces:

A. Continuous classification
B. Continuous regression
C. Continuous training
D. All of the above.

In addition to CI/CD practiced by DevOps teams, MLOps introduces:

A. Continuous classification

B. Continuous regression

C. Continuous training

D. All of the above.

In what order are the following phases executed in a machine learning project?

I - Selection of ML algorithm
II - Data Exploration
III - Definition of the business use case
IV - Model monitoring
V - Model operationalization
VI - Model Development

A. I, II, III, IV, V, VI

B. III, II, I, VI, V, IV

C. II, III, I, VI, IV, V

D. II, I, III, VI, IV, V

In what order are the following phases executed in a machine learning project?

I - Selection of ML algorithm
II - Data Exploration
III - Definition of the business use case
IV - Model monitoring
V - Model operationalization
VI - Model Development

III - Definition of the business use case
II - Data Exploration
I - Selection of ML algorithm
VI - Model Development
V - Model operationalization
IV - Model monitoring

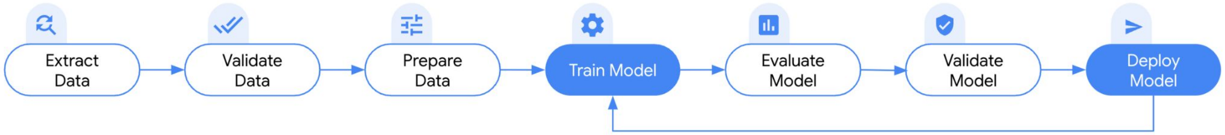
A. I, II, III, IV, V, VI

B. III, II, I, VI, V, IV

C. II, III, I, VI, IV, V

D. II, I, III, VI, IV, V

Pipeline



You want to have two versions of your application in production, but be able to direct a small percentage of traffic to the newer version as a gradual test. This is an example of which deployment strategy?

A. Canary deployment

B. Blue-green deployment

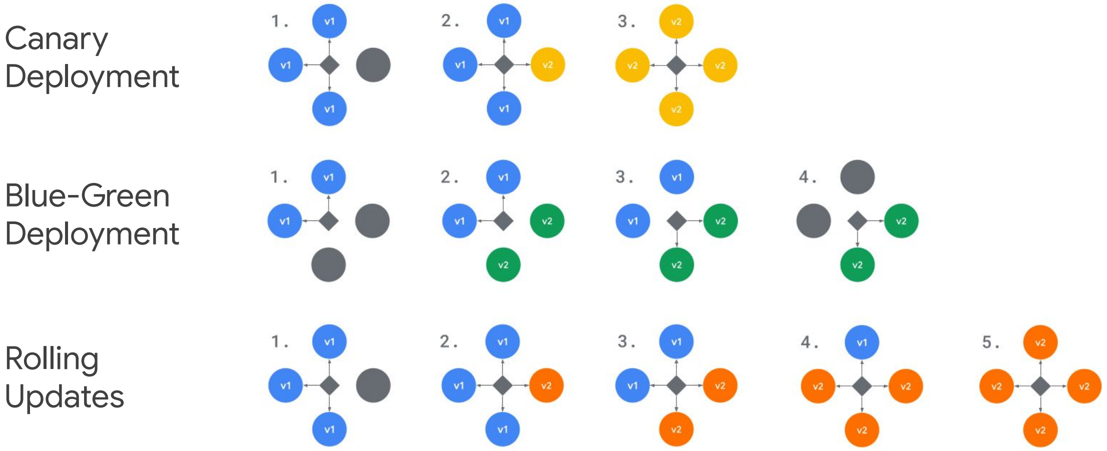
C. Rolling updates

You want to have two versions of your application in production, but be able to direct a small percentage of traffic to the newer version as a gradual test. This is an example of which deployment strategy?

A. Canary deployment

B. Blue-green deployment

C. Rolling updates

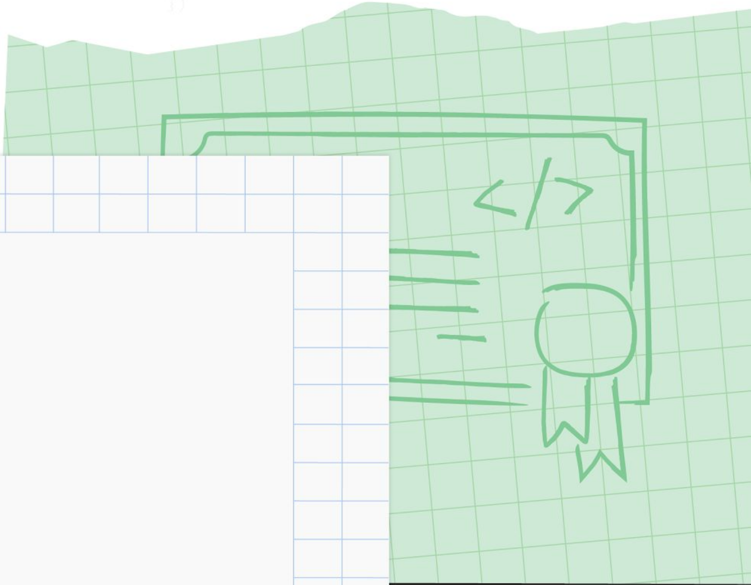


Canary Deployment

Blue-Green Deployment

Rolling Updates

Q&A



```
nction filterStudies({ studies, filt
  let filteredStudies = studies.filt
    const matchOrg = filterByOrg ?
    const matchStatus = filterBySta
    if (matchOrg && matchStatus) {
      return true
    }
    return false
  })
```

```
yStatus = false }) {

=== filterByOrg : true
ilterByStatus : true
```

```
lse          function filterStudies({ studies, filterByOrg = false, filterBy
               let filteredStudies = studies.filter(study => {
```

ChengCheng Tan

ccstan99@gmail.com

in cheng2-tan

@cheng2_tan