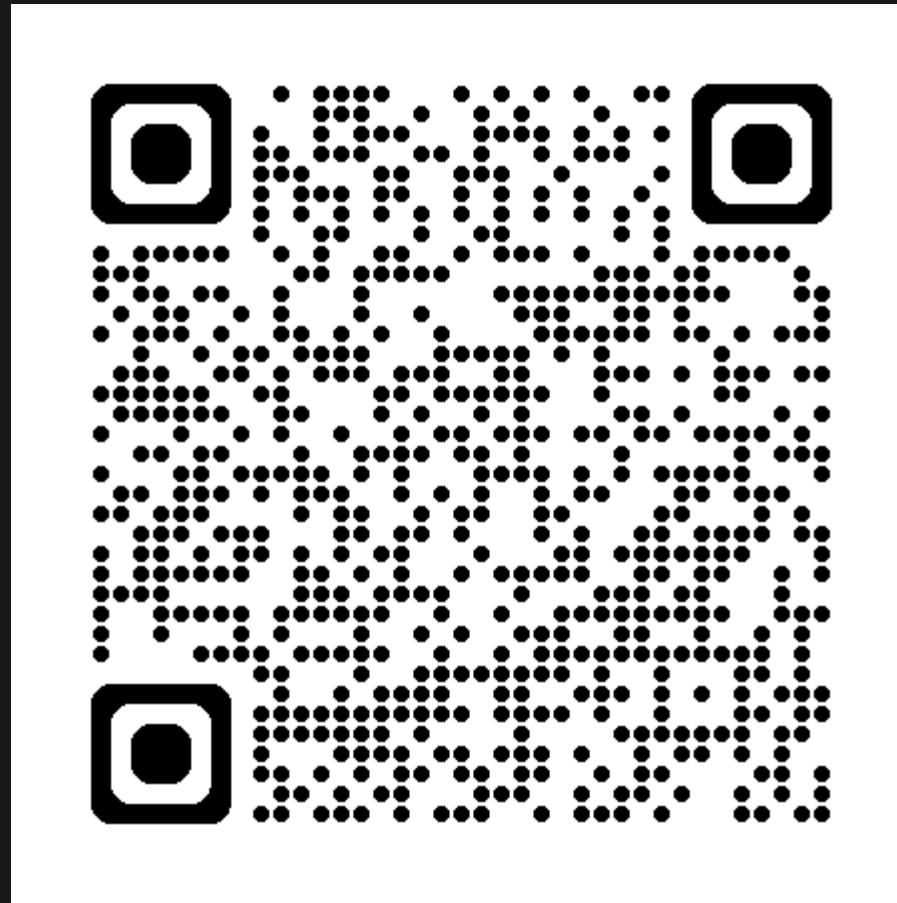


# INTEGRATING GENERATIVE AI IN YOUR CF APPLICATION

# ME

- Raymond Camden
- Senior Developer Evangelist for Adobe
- 🌐 raymondcamden.com
- 📩 raymondcamden@gmail.com
- Former ColdFusion Jedi





[github.com/cfjedimaster/gemini-ai-preso](https://github.com/cfjedimaster/gemini-ai-preso)

# ASSUMPTIONS AND WARNINGS

# WHAT YOU KNOW, WHAT I KNOW

# WHAT YOU KNOW, WHAT I KNOW

- What Generative AI is (**high level!!**)

# WHAT YOU KNOW, WHAT I KNOW

- What Generative AI is (**high level!!**)
- ColdFusion (see previous warning)

# WHAT YOU KNOW, WHAT I KNOW

- What Generative AI is (**high level!!**)
- ColdFusion (see previous warning)
- Noob on Stage (me, not you)

# TLDR: GOOD NEWS

# TLDR: GOOD NEWS

Getting a key, writing the code, it's simple!

# TLDR: BAD NEWS

# TLDR: BAD NEWS

\*Correctly\* using Gen AI is way more complex and "prompt engineering" is no joke.

# GOOGLE GEMINI

# GOOGLE GEMINI

- The artist formally known as Bard (Feb 2023)

# GOOGLE GEMINI

- The artist formally known as Bard (Feb 2023)
- Relaunched as Gemini (Dec 2023)

# GOOGLE GEMINI

- The artist formally known as Bard (Feb 2023)
- Relaunched as Gemini (Dec 2023)
- API access in October (PaLM - text only)

# GOOGLE GEMINI

- The artist formally known as Bard (Feb 2023)
- Relaunched as Gemini (Dec 2023)
- API access in October (PaLM - text only)
- Gemini API (also Dec 2023)

# GOOGLE GEMINI

- The artist formally known as Bard (Feb 2023)
- Relaunched as Gemini (Dec 2023)
- API access in October (PaLM - text only)
- Gemini API (also Dec 2023)
- Don't forget: Gemini App ([gemini.google.com](https://gemini.google.com))

# GOOGLE GEMINI

- The artist formally known as Bard (Feb 2023)
- Relaunched as Gemini (Dec 2023)
- API access in October (PaLM - text only)
- Gemini API (also Dec 2023)
- Don't forget: Gemini App ([gemini.google.com](https://gemini.google.com))
- Also... Vertex

# GETTING STARTED

# GETTING STARTED

- [ai.google.dev](https://ai.google.dev)

# GETTING STARTED

- [ai.google.dev](https://ai.google.dev)
- [aistudio.google.com](https://aistudio.google.com)

# GETTING STARTED

- [ai.google.dev](https://ai.google.dev)
- [aistudio.google.com](https://aistudio.google.com)
- [Docs](#)

# GETTING STARTED

- [ai.google.dev](https://ai.google.dev)
- [aistudio.google.com](https://aistudio.google.com)
- [Docs](#)
- [Support Forum](#)

# MODELS

# MODELS

- Gemini 1.5 Pro

# MODELS

- Gemini 1.5 Pro
- Gemini 1.5 Flash (quicker, not as good, but quicker)

# MODELS

- Gemini 1.5 Pro
- Gemini 1.5 Flash (quicker, not as good, but quicker)
- Changing... like often

# MODELS

- Gemini 1.5 Pro
- Gemini 1.5 Flash (quicker, not as good, but quicker)
- Changing... like often
- [More at docs](#)

# 1.5 PRO

# 1.5 PRO

- Huge Context (2M tokens)

# 1.5 PRO

- Huge Context (2M tokens)
- 2 hours of video

# 1.5 PRO

- Huge Context (2M tokens)
- 2 hours of video
- 19 hours of audio

# 1.5 PRO

- Huge Context (2M tokens)
- 2 hours of video
- 19 hours of audio
- 7200 images

# 1.5 PRO

- Huge Context (2M tokens)
- 2 hours of video
- 19 hours of audio
- 7200 images
- 2000 pages of text (approx)

# 1.5 PRO

- Huge Context (2M tokens)
- 2 hours of video
- 19 hours of audio
- 7200 images
- 2000 pages of text (approx)
- 60K lines of code (approx)

# CODE

# CODE

- REST API

# CODE

- REST API
- Python

# CODE

- REST API
- Python
- Node

# CODE

- REST API
- Python
- Node
- Go

# CODE

- REST API
- Python
- Node
- Go
- Web 🤔

# CODE

- REST API
- Python
- Node
- Go
- Web 🤔
- Dart

# CODE

- REST API
- Python
- Node
- Go
- Web 🤔
- Dart
- Flutter

# CODE

- REST API
- Python
- Node
- Go
- Web 🤔
- Dart
- Flutter
- Android

# ACTUALLY USING IT

# ACTUALLY USING IT

- Step 1: Get a Key

# ACTUALLY USING IT

- Step 1: Get a Key
- Step 2:

# ACTUALLY USING IT

- Step 1: Get a Key
- Step 2:
- Step 3: Profit!

Google AI Studio

- [Get API key](#) 
- [Create new prompt](#)
- [New tuned model](#)
- [My library](#)
- No prompts yet

---

- [Getting started](#)
- [Documentation](#)
- [Prompt gallery](#)
- [Gemini cookbook](#)
- [Discourse forum](#)
- [Build with Vertex AI on Google Cloud](#)

## Untitled prompt

▼ System Instructions

Run a prompt to see model response here

## 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](#), which you agree to when creating a new project, while use of the Gemini API and Google AI Studio is subject to the [Gemini API Terms of Service](#).

Use your API keys securely. Do not share them or embed them in code the public can view.

Starting on May 14, 2024, if you use Gemini API from a project that has billing enabled, your use will be subject to [pay-as-you-go pricing](#).

 Create API key

Your API keys are listed below. You can also view and manage your project and API keys in Google Cloud.

API key	Google Cloud project name	Created	Action
Create an API key to see your projects			

## Create API key

X

Create an API key in a new project if you don't have one already

 Create API key in new project

or

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

 Search Google Cloud projects

 Create API key in existing project

API key generated

X

Use your API keys securely. Do not share them or embed them in code the public can view.

 Copy

```
curl \  
  -H 'Content-Type: application/json' \  
  -d '{"contents": [{"parts": [{"text": "Write a story about a ma  
-X POST 'https://generativelanguage.googleapis.com/v1beta/mo
```

Show demos/curh.sh

```
{  
  "candidates": [  
    {  
      "content": {  
        "parts": [  
          {  
            "text": "In the bustling metropolis of Willow Creek,  
          }  
        ],  
        "role": "model"  
      },  
      "finishReason": "STOP",  
      "index": 0,  
      "safetyRatings": [  
        {  
          "score": 1,  
          "type": "Harmless"  
        }  
      ]  
    }  
  ]  
}
```

## Reference for Response

# AI STUDIO

# AI STUDIO

- Let's you create, and save, prompts

# AI STUDIO

- Let's you create, and save, prompts
- Super, super important for testing and tweaking

# AI STUDIO

- Let's you create, and save, prompts
- Super, super important for testing and tweaking
- Then let's you generate code

# AI STUDIO DEMO AND CF...

Show AI Studio, code output, and then first and second.cfm

# MULTIMODAL

# MULTIMODAL

- Just means multiple types of input

# MULTIMODAL

- Just means multiple types of input
- Images, text files, and video

# MULTIMODAL

- Just means multiple types of input
- Images, text files, and video
- \*Can\* be inlined as Base64, if less than 20 megs

# MULTIMODAL

- Just means multiple types of input
- Images, text files, and video
- \*Can\* be inlined as Base64, if less than 20 megs
- Files API

# MULTIMODAL

- Just means multiple types of input
- Images, text files, and video
- \*Can\* be inlined as Base64, if less than 20 megs
- Files API
- Files API max - 20 gigs per project, 2 per file

# MULTIMODAL

- Just means multiple types of input
- Images, text files, and video
- \*Can\* be inlined as Base64, if less than 20 megs
- Files API
- Files API max - 20 gigs per project, 2 per file
- Files API - auto delete after 2 days (you can manually delete as well)

# FILE TYPE SPECIFICS

# FILE TYPE SPECIFICS

- Image: PNG, JPG, WEBP, HEIC, HEIF

# FILE TYPE SPECIFICS

- Image: PNG, JPG, WEBP, HEIC, HEIF
- Audio: WAV, MP3, AIFF, AAC, OGG, FLAC

# FILE TYPE SPECIFICS

- Image: PNG, JPG, WEBP, HEIC, HEIF
- Audio: WAV, MP3, AIFF, AAC, OGG, FLAC
- Note - Gemini understands 'non-speech' parts (ie, sounds)

# FILE TYPE SPECIFICS

- Image: PNG, JPG, WEBP, HEIC, HEIF
- Audio: WAV, MP3, AIFF, AAC, OGG, FLAC
- Note - Gemini understands 'non-speech' parts (ie, sounds)
- Video: MP4, MPEG, MOV, AVI, X-FLV, MPG, WEBM, WMV, 3GPP

# FILE TYPE SPECIFICS

- Image: PNG, JPG, WEBP, HEIC, HEIF
- Audio: WAV, MP3, AIFF, AAC, OGG, FLAC
- Note - Gemini understands 'non-speech' parts (ie, sounds)
- Video: MP4, MPEG, MOV, AVI, X-FLV, MPG, WEBM, WMV, 3GPP
- Note - for video, image frames are extracted (1 per second)

# FILE TYPE SPECIFICS

- Image: PNG, JPG, WEBP, HEIC, HEIF
- Audio: WAV, MP3, AIFF, AAC, OGG, FLAC
- Note - Gemini understands 'non-speech' parts (ie, sounds)
- Video: MP4, MPEG, MOV, AVI, X-FLV, MPG, WEBM, WMV, 3GPP
- Note - for video, image frames are extracted (1 per second)
- Docs: PDF, TXT, Markdown, CSV

# AI STUDIO DEMO - MM

In AI Studio, upload the cat!

# CODE DEMO - MM

Show image.cfm, and possibles images2.cfm and the fun one.

# CHAT

# CHAT

- Why? Context

# CHAT

- Why? Context
- First and Foremost: Nothing Special

# CHAT

- Why? Context
- First and Foremost: Nothing Special
- Chat is just prompts that... keep... growing

# CHAT

- Why? Context
- First and Foremost: Nothing Special
- Chat is just prompts that... keep... growing
- You manage the history...

# CHAT

- Why? Context
- First and Foremost: Nothing Special
- Chat is just prompts that... keep... growing
- You manage the history...
- Use the SDK! ✨

# CHAT

- Why? Context
- First and Foremost: Nothing Special
- Chat is just prompts that... keep... growing
- You manage the history...
- Use the SDK! ✨
- Except we can't 😭

# DEMO

Show chat.cfm

# SYSTEM INSTRUCTIONS

# SYSTEM INSTRUCTIONS

- Provide context about how the model should respond

# SYSTEM INSTRUCTIONS

- Provide context about how the model should respond
- Think persona, style, as well as output

# SYSTEM INSTRUCTIONS

- Provide context about how the model should respond
- Think persona, style, as well as output
- Can be used in the prompt as well

```
"systemInstruction": {  
    "role": "user",  
    "parts": [  
        {  
            "text": "You are a cat. Your name is Neko."  
        }  
    ]  
}
```

# DEMO

Show si and si2.cfm

# SHAPING RESPONSES

# SHAPING RESPONSES

- Your prompt can ask for JSON

# SHAPING RESPONSES

- Your prompt can ask for JSON
- So can system instructions

# SHAPING RESPONSES

- Your prompt can ask for JSON
- So can system instructions
- Be specific

# SHAPING RESPONSES

- Your prompt can ask for JSON
- So can system instructions
- Be specific
- There's also a response type setting...

# SHAPING RESPONSES - 2

# SHAPING RESPONSES - 2

- By default, Gemini returns text in Markdown

# SHAPING RESPONSES - 2

- By default, Gemini returns text in Markdown
- Even if you ask for JSON, it's JSON... in Markdown

# SHAPING RESPONSES - 2

- By default, Gemini returns text in Markdown
- Even if you ask for JSON, it's JSON... in Markdown
- Setting the response type is essential!

```
"generationConfig": {  
    "temperature": 1,  
    "topK": 64,  
    "topP": 0.95,  
    "maxOutputTokens": 8192,  
    "responseMimeType": "application/json"  
}
```

# DEMO

Show json\_response\*

# DEMOS

# MOVIE RECOMMENDATIONS

- Get user's last movie watched
- Pass it to Gemini and ask - what next?
- Email suggestions to user
- [Blog post](#)

# D&D RECOMMENDATION

- Let users roll for attributes
- Pass it to Gemini and ask - what class makes the most sense?
- [Blog post](#)

# BLACKJACK HELP

- Given a web-based Blackjack game...
- And the status of the game...
- Can AI help you play?
- [Blog post](#)

# THINGS NOT COVERED

# THINGS NOT COVERED

- Function Calling ([Docs](#))

# THINGS NOT COVERED

- Function Calling ([Docs](#))
- Model Tuning ([Docs](#))

# THINGS NOT COVERED

- Function Calling ([Docs](#))
- Model Tuning ([Docs](#))
- Prompt Writing ([Docs](#))

# THINGS NOT COVERED

- Function Calling ([Docs](#))
- Model Tuning ([Docs](#))
- Prompt Writing ([Docs](#))
- [Gemini API Cookbook \(Notebooks\)](#)

# VERTEX - WHAT THE DIFF?

# VERTEX - WHAT THE DIFF?

- Vertex has no free tier

# VERTEX - WHAT THE DIFF?

- Vertex has no free tier
- AI Studio's free tier may use your data for training

# VERTEX - WHAT THE DIFF?

- Vertex has no free tier
- AI Studio's free tier may use your data for training
- AI Studio isn't available (region wise) as much as Vertex

# VERTEX - WHAT THE DIFF?

- Vertex has no free tier
- AI Studio's free tier may use your data for training
- AI Studio isn't available (region wise) as much as Vertex
- Vertex requires Google Cloud auth models

# VERTEX - WHAT THE DIFF?

- Vertex has no free tier
- AI Studio's free tier may use your data for training
- AI Studio isn't available (region wise) as much as Vertex
- Vertex requires Google Cloud auth models
- More info at <https://cloud.google.com/vertex-ai>

# VERTEX - WHAT THE DIFF?

- Vertex has no free tier
- AI Studio's free tier may use your data for training
- AI Studio isn't available (region wise) as much as Vertex
- Vertex requires Google Cloud auth models
- More info at <https://cloud.google.com/vertex-ai>
- Thank you [Allen Firstenberg](#)

# QUESTIONS?

