

AI Guessing Image Game By Using p5.js and Ollama

By **Charunthon Limseelo** (@boatchrnthn)
Microsoft Learn Student Ambassador

And **Natthawat Hongthong** (@natthawathongthong)
as Web Designer



Charunthon Limseelo (Boat)

- Microsoft Learn Student Ambassadors (Thai Student Tech Lead) and Data Engineer Intern at Seven Peaks Software – Klong Toei, Bangkok
- Third Year Student at Computer Engineering (International Program) King Mongkut's University of Technology Thonburi



Charunthon Limseelo



@boatchrnthn



Charunthon Limseelo



Charunthon Limseelo



Contents

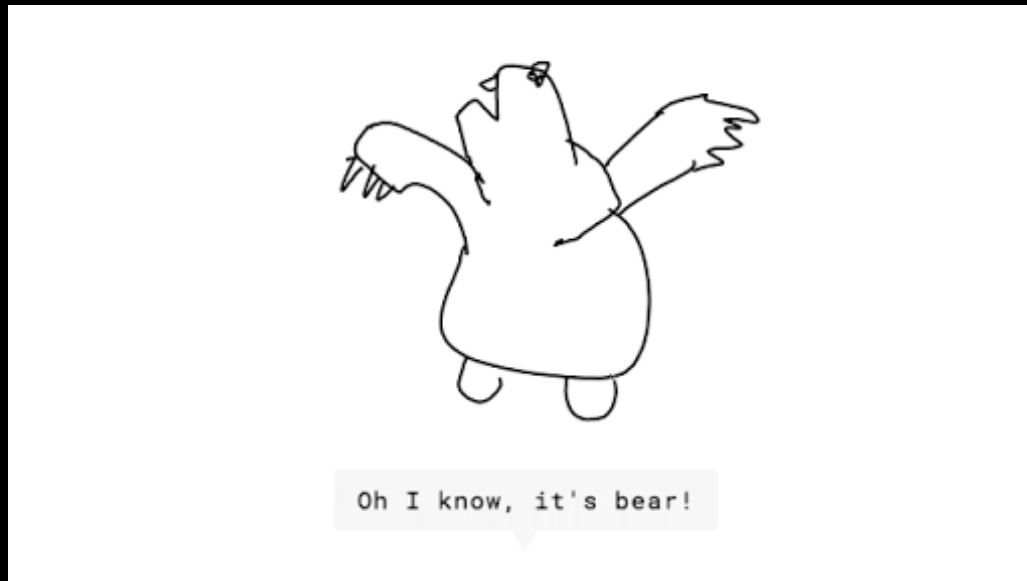
- Quick, Draw! – My Work Inspiration
- Introducing LLaVA, Large Language Vision Assistant
- Introducing Ollama - Like Docker, But for LLMs
- Getting into the project repository and live drawing demonstration
- Key Takeaways and **FOSSASIA**
(how to dev local model based on sustainability)



Quick, Draw! – My Work Inspiration

By Google Creative Lab

A game where a neural net tries to guess what you're drawing.



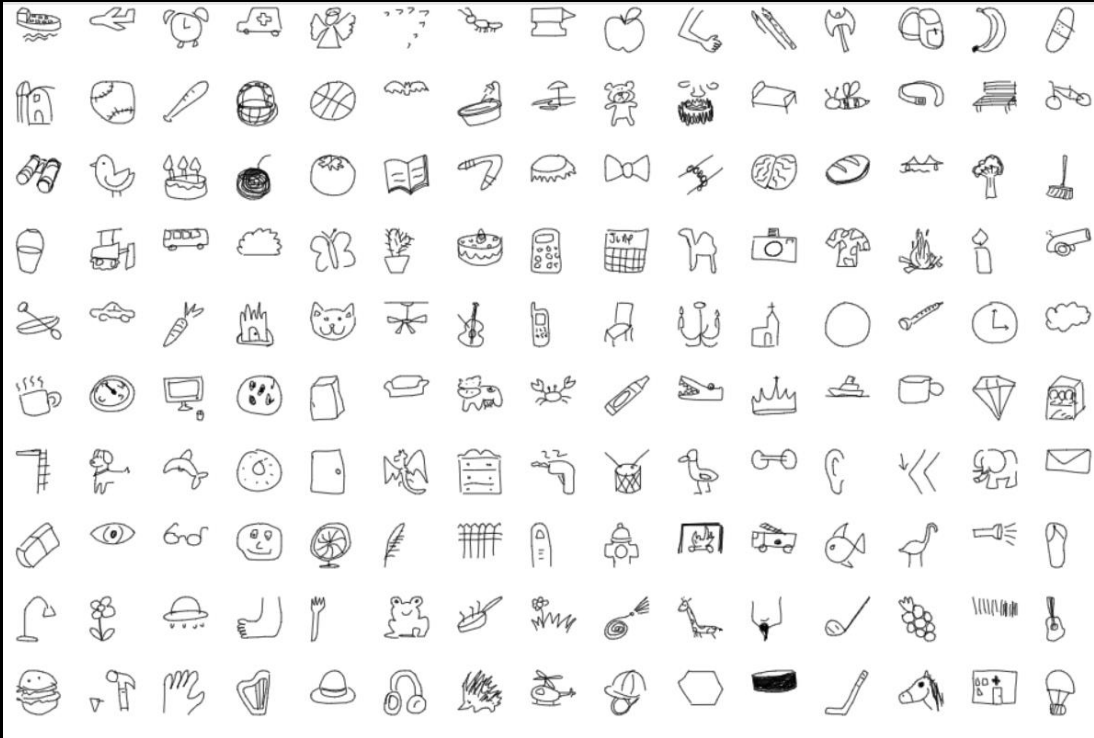
BUILT WITH:

Handwriting Recognition

- This is a game built with machine learning. You draw, and a neural network tries to guess what you're drawing. Of course, it doesn't always work. But the more you play with it, the more it will learn. It's just one example of how you can use machine learning in fun ways.
- Built by Jonas Jongejan, Henry Rowley, Takashi Kawashima, Jongmin Kim, Nick Fox-Gieg, with friends at Google Creative Lab and Data Arts Team.

Quick, Draw! Dataset

By Google Creative Lab



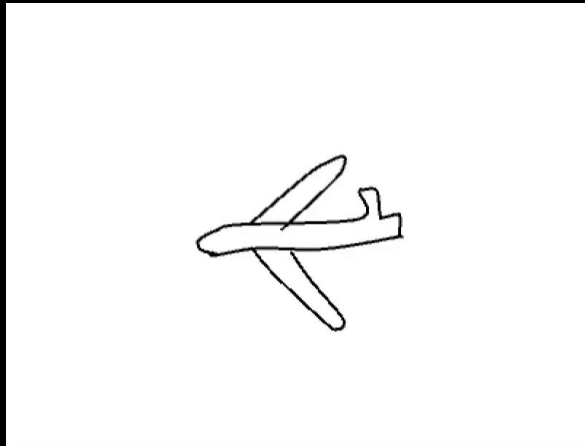
- Over 15 million players have contributed millions of drawings playing Quick, Draw! These doodles are a unique data set that can help developers train new neural networks, help researchers see patterns in how people around the world draw, and help artists create things we haven't begun to think of. That's why we're open-sourcing them, for anyone to play with.

GitHub Repository

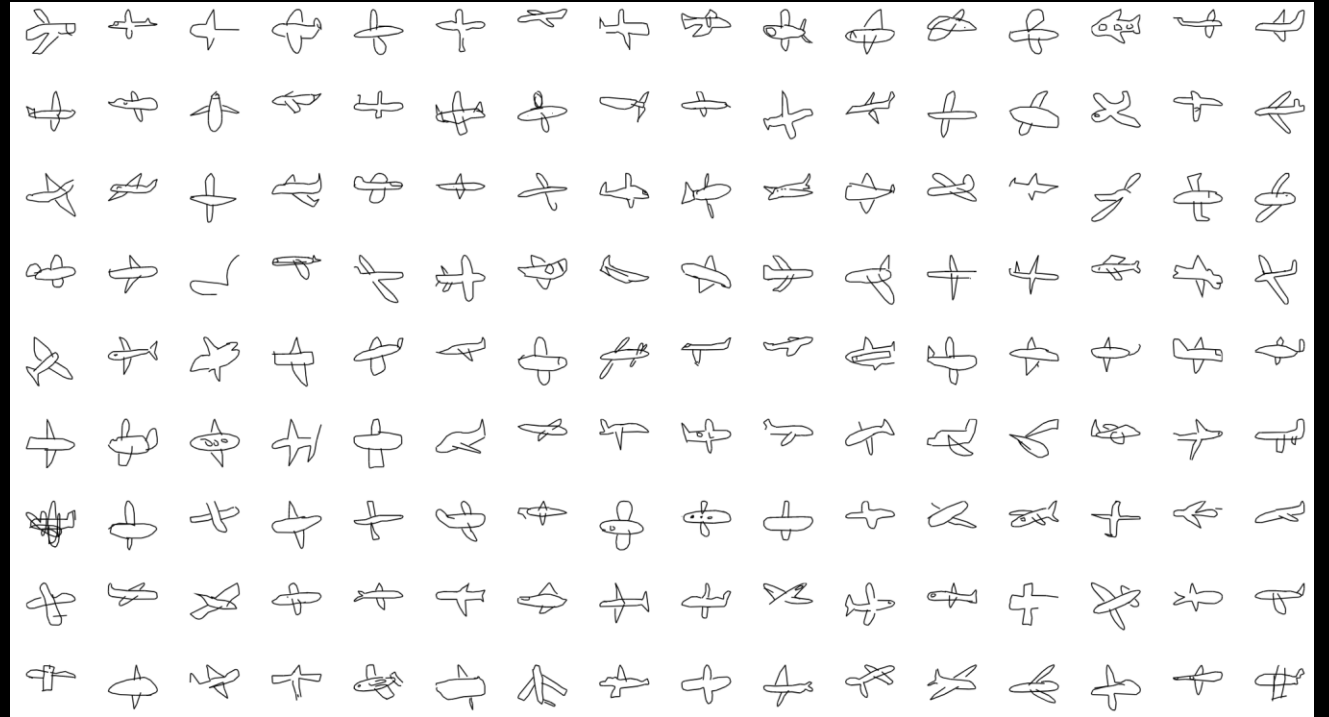
<https://github.com/googlecreativelab/quickdraw-dataset>

It's not easy in the past...

When you let the machine be able to classify the image that you draw (It's all about **Neural Network**)

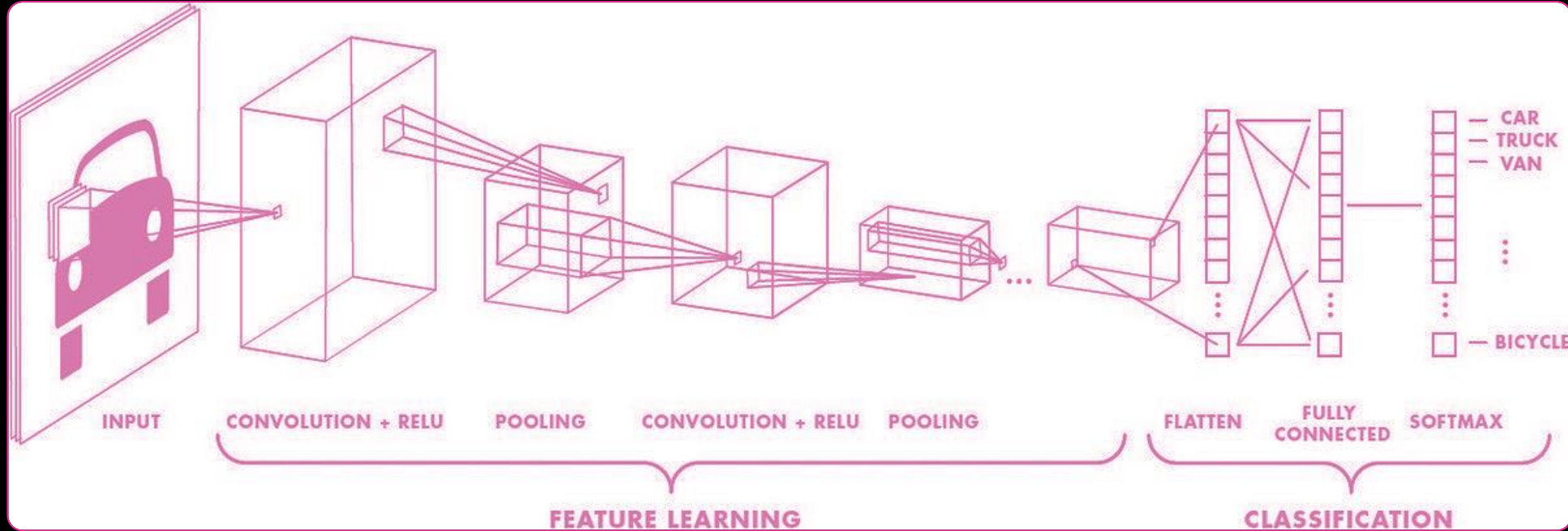


Human's Airplane Drawing



Finding similarities and make a classification and finally can say that
"This is an airplane"

Convolution Neural Network (CNN)



- A convolutional neural network is a regularized type of feedforward neural network that learns features by itself via filter optimization. This type of deep learning network has been applied to process and make predictions from many different types of data including text, images and audio.

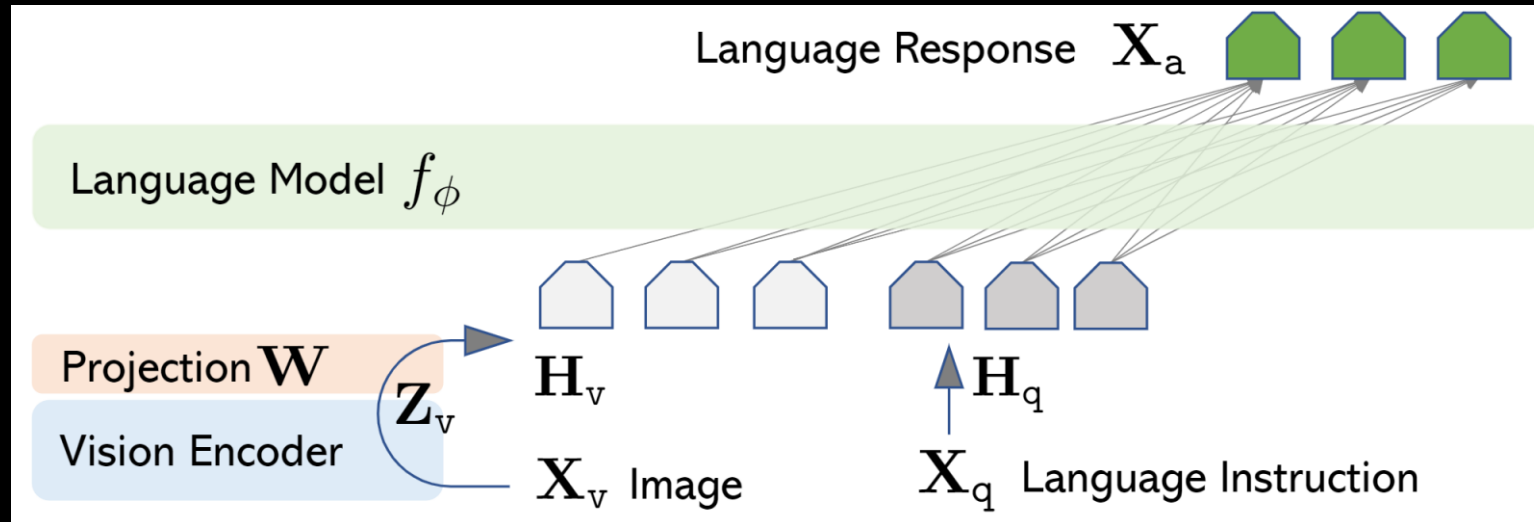
Things are different

As the upcoming of Generative AI, and Multimodal Language Models in nowadays

We don't have to train our own base models for each work, and you can use the open-source models from Hugging Face to work on...

LLaVA – Large Language Vision Assistant

First Multimodal Model Research, created by Microsoft Research



Features:

- Visual Reasoning Chat: Towards building multimodal GPT-4 level chatbot
- Optical character recognition (OCR)

- A novel end-to-end trained large multimodal model that combines a vision encoder and Vicuna for general-purpose visual and language understanding, achieving impressive chat capabilities mimicking spirits of the multimodal GPT-4 and setting a new state-of-the-art accuracy on Science QA.

Ollama – Like Docker, But for LLMs

Created by Jeffrey Morgan



Ollama



- Ollama is a lightweight, extensible framework for building and running language models on the local machine. It provides a simple API for creating, running, and managing models, as well as a library of pre-built models that can be easily used in a variety of applications.

```
ollama
+
Update Warp
base ~/Code/ecosystem
ollama run zephyr
eval count:      31/ token(s)
eval duration:   6.37296s
eval rate:      49.74 tokens/s
>>> what animals do not?
As I mentioned earlier, many animal species do not have clear and rigid social hierarchies among adult individuals in their wild populations. Some examples include:

1. Polar bears: They are solitary animals that do not live in groups with a social hierarchy.

2. Bald eagles: While they form pairs during mating season, there is no dominant leader or hierarchy within the pair or between pairs.

3. Sea otters: Although they form social groups, there is no hierarchy among them. The social structure is more cooperative and based on sharing resources.

4. Tiger sharks: They are solitary predators that do not live in groups with a social hierarchy.

5. Orangutans: While they may form social groups led by a dominant female, there is generally less rigid and hierarchical structure compared to some other primates such as chimpanzees, gorillas, and bonobos.

These are just a few examples of animals that do not have clear social hierarchies among adult individuals in their wild populations. However, it's important to note that the presence or absence of social hierarchy can vary depending on factors such as species, environment, and context.

total duration: 8.592375916s
load duration:  1.528084666s
prompt eval count: 1169 token(s)
prompt eval duration: 2.050488s
prompt eval rate: 570.11 tokens/s
eval count:      255 token(s)
eval duration:   4.95492s
eval rate:      51.46 tokens/s
>>> end a message (/? for help)
```

Inside's Architecture

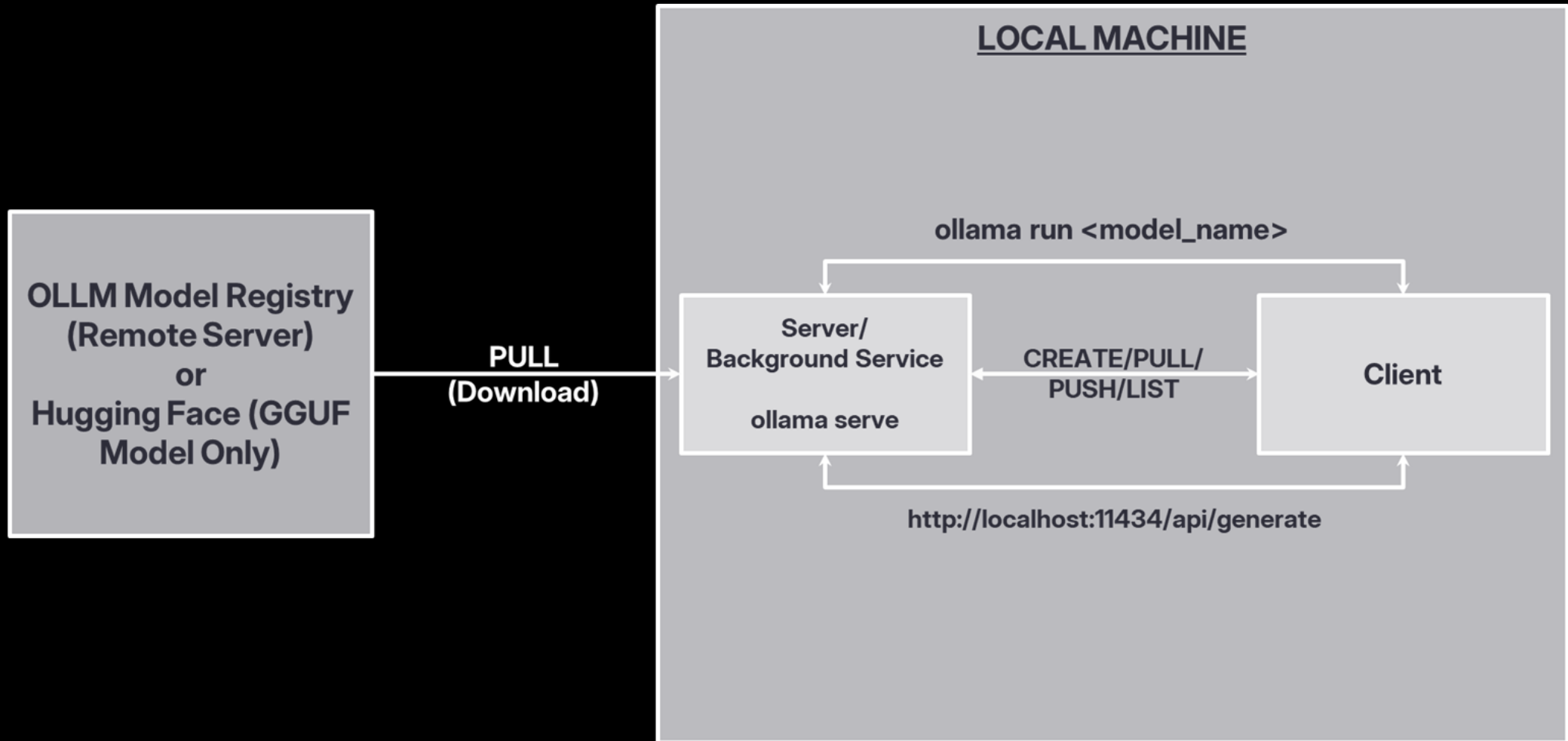
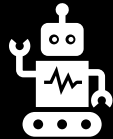


Image-to-text Generation Process



User: Do you know who drew this painting?



LLaVA processing



LLaVA: (Generate)

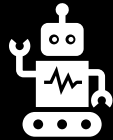
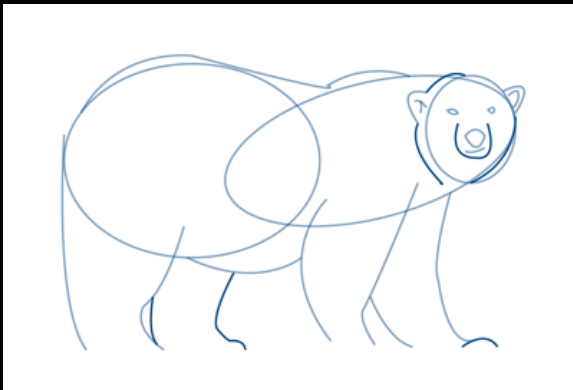
The painting depicts a woman, commonly believed to be Mona Lisa, the famous artwork by Leonardo da Vinci. It is a portrait painting that showcases the woman's enigmatic smile and has become one of the most famous and iconic art pieces in the world. The original work is displayed in the Louvre Museum in Paris, and it is known for its intricate details, use of oil paint, and the artist's innovative techniques that contributed to its enduring appeal and mystery.

Today's Planned Architecture



User:

Finish Drawing > Click Submit



LlaVA processing



LlaVA: (Generate)

This is a bear!

Today's Planned Architecture



User:

Finish drawing > [unclear]

"Oh, I Want Something Just Like This"



Chris Martin from Coldplay

LlaVA: (Generate)

This is a bear!


Demo Section


Getting into the project repository and live drawing demonstration


Visit the repository and try it


From My GitHub Page


chrnthnkmutt/
Ollama_Drawing_p5js




 1
Contributor

 0
Issues

 0
Stars

 0
Forks





Key Takeaways and FOSSASIA

Why do I need to talk about it?

Multimodal Models + Creativity

Brings us to the creative solution that could see many possibilities on surrounded environment

Perception + Creativity

Brings to the humans' beyond thinking for the new things

Creativity + AI + Sustainability = SLMs

Brings us everything that could bring any possibilities for our world
with endless thinking and environment caring, and more...

Thank you for your participation

Hope we could see each other again
next time! 🙌

Follow me at @boatchrnthn on Social
Media and @chrnthnkmutt on GitHub



FOSSASIA SUMMIT 2025

True Digital Park, Bangkok
summit.fossasia.org

March 13-15



การประชุมเทคโนโลยีโอเพนซอร์สชั้นนำของเอเชีย

Charunthon Limseelo

Microsoft Learn Student Ambassador
King Mongkut University of
Technology Thonburi (KMUTT)

Small Language Models: The New
Frontier of Developing Open-source
Generative AI to Sustainability Goals