

# Generative

Bocheng Lin

2024.3.13

PART I

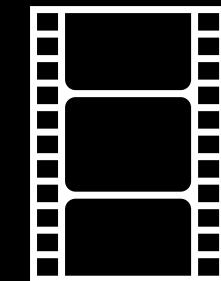
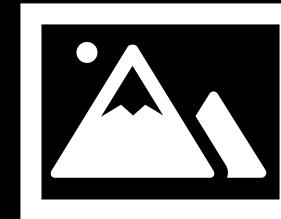
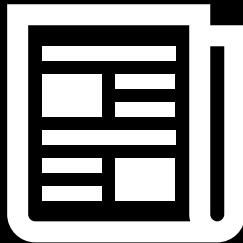
# Generative Artificial Intelligence

The goal of Artificial Intelligence  
is to enable machines to exhibit  
Intelligence.

**Generative AI**  
is to enable machines to generate  
complex and structured things.

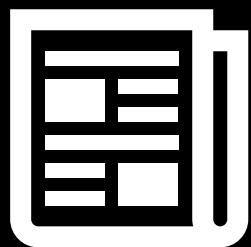
# Generative AI

is to enable machines to **generate**  
complex and structured things.



Generative AI  
is to enable machines to generate  
**complex** and structured things.

# Generative AI is to enable machines to generate **complex** and structured things.



To generate a **100-word** article, assuming there are **1000** common words, the number of possibilities for this article would be

**$10^{300}$**

The estimated number of atoms in the universe is around  **$10^{80}$**

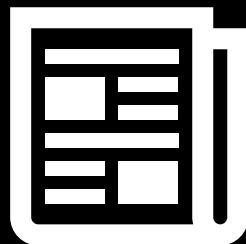
# Generative AI is to enable machines to generate **complex** and structured things.



To find out appropriate  
combinations from 

Generative AI  
is to enable machines to generate  
complex and structured things.

To generate a **100-word** article



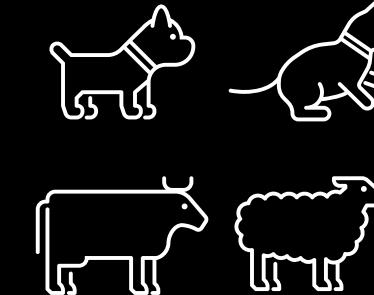
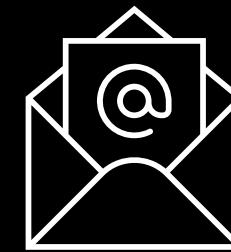
**Classification?**

## Generative AI

is to enable machines to generate complex and structured things.

## Classification?

Make choices from a limited set of options



Generative AI  
is to enable machines to generate  
**complex** and structured things.



almost impossible to exhaustively enumerate

PART II

# Machine Learning

Machine Learning  
is to automatically find a function  
from data.

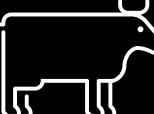
# Machine Learning is to automatically find a **function** from data.

$$f(x, y) = [a \quad b] \begin{bmatrix} x \\ y \end{bmatrix} + c$$

We know a set of input and output, and we want to find out the parameters **by machine, automatically.**

# Machine Learning is to automatically find a **function** from data.

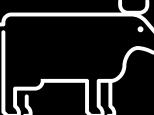
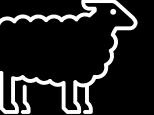
$$f(\text{dog}) = \begin{bmatrix} \dots \\ \vdots \\ \ddots \\ \dots \end{bmatrix}$$

Input                

Output **Dog, Cat, Cow, Sheep**

# Machine Learning is to automatically find a **function** from data.

$$f(\text{dog}) = \begin{bmatrix} \vdots & \ddots & \vdots \\ \dots & \ddots & \vdots \\ \vdots & \ddots & \vdots \end{bmatrix}$$

<b>Input</b>				
<b>Output</b>	<b>Dog, Cat, Cow, Sheep</b>			
	<b>Training data</b>			

**Machine Learning can  
find out the parameters**

# Machine Learning is to automatically find a **function** from data.

$$f(\text{Dog}) = \begin{bmatrix} \vdots & \ddots & \vdots \\ \dots & \ddots & \dots \end{bmatrix} \text{Model}$$

Input



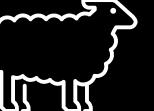
Output **Dog, Cat, Cow, Sheep**  
Training data

Machine Learning can  
find out the parameters

# Machine Learning is to automatically find a **function** from data.

**Inference (Testing)**  $f(\text{dog}) =$  

**Model**

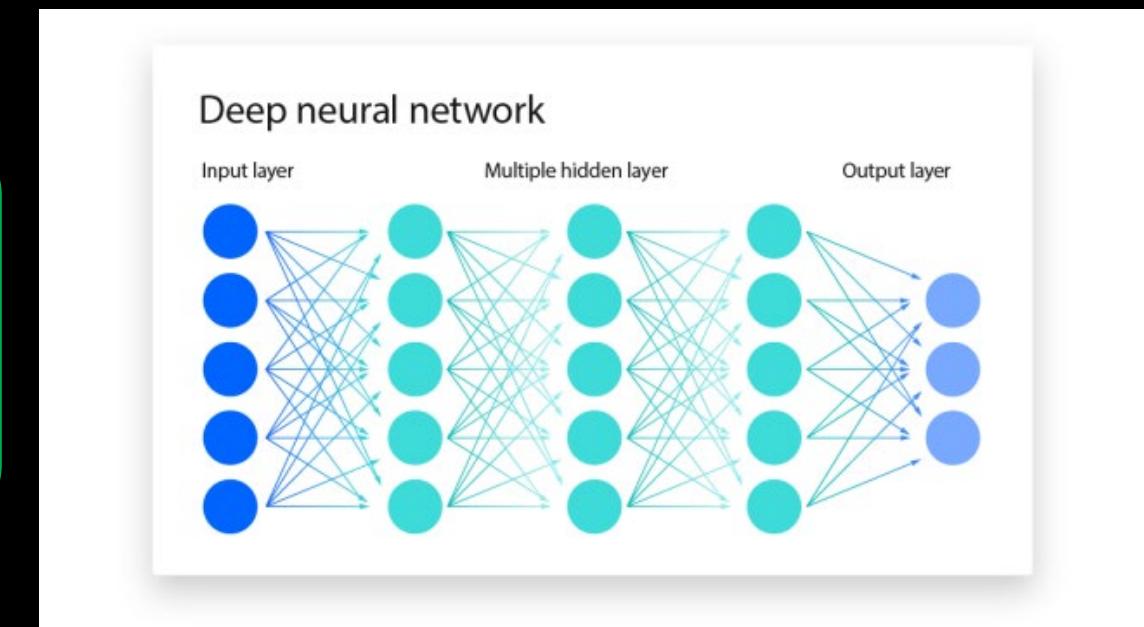
**Input**    

**Output** **Dog, Cat, Cow, Sheep**  
**Training data**

**Learning (Training)**

# Machine Learning is to automatically find a **function** from data.

$$f(\text{dog}) = \begin{bmatrix} \vdots & \cdots & \vdots \end{bmatrix} \text{Model}$$



## Deep Learning

Machine Learning  
is to automatically find a **function** from data.

How?

Machine Learning  
is to automatically find a **function** from data.

**Set scope**

**Establish criteria**

**Achieve the goal**

Machine Learning  
is to automatically find a **function** from data.

**Set scope** Define the set of candidate functions

Establish criteria

Achieve the goal

Machine Learning  
is to automatically find a **function** from data.

**Set scope**

Establish criteria

Achieve the goal

Define the set of candidate functions  
**Model**

In Deep Learning, the structures of neural networks, such as *CNN*, *RNN*, and *Transformer*, etc., refer to the sets of different candidate functions.

Machine Learning  
is to automatically find a **function** from data.

Set scope

Define the set of candidate functions

Establish criteria

Define the criteria for evaluating  
the quality of functions **Loss**

Achieve the goal

Machine Learning  
is to automatically find a **function** from data.

**Set scope**

Define the set of candidate functions

**Establish criteria**

Define the criteria for evaluating  
the quality of functions **Loss**

**Achieve the goal** Find the best function → Optimization

# Machine Learning is to automatically find a **function** from data.

**Set scope**

Define the set of candidate functions

**Deep Learning (CNN, Transformer...),  
Decision Tree, etc.**

**Establish criteria**

Define the criteria for evaluating the quality of functions

**Supervised Learning, Semi-supervised  
Learning, Reinforcement Learning, etc.**

**Achieve the goal**

Find the best function → Optimization

**Gradient Descent, Genetic Algorithm, etc.**

**Artificial Intelligence**

**Generative AI**

**Machine Learning**

**Deep Learning**

**The Ultimate Target: Artificial Intelligence**

**One of the Ultimate Target: Generative AI**

**Method: Machine Learning**

**A powerful Method: Deep Learning**

**The Ultimate Target: Artificial Intelligence**

**One of the Ultimate Target: Generative AI**

**Method: Machine Learning**

**A powerful Method: Deep Learning**

**Current Generative AI is mostly  
achieved through Deep Learning.**

PART III

# Generate Text

## Large Language Model

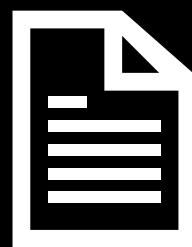
# ChatGPT

# Generative ChatGPT

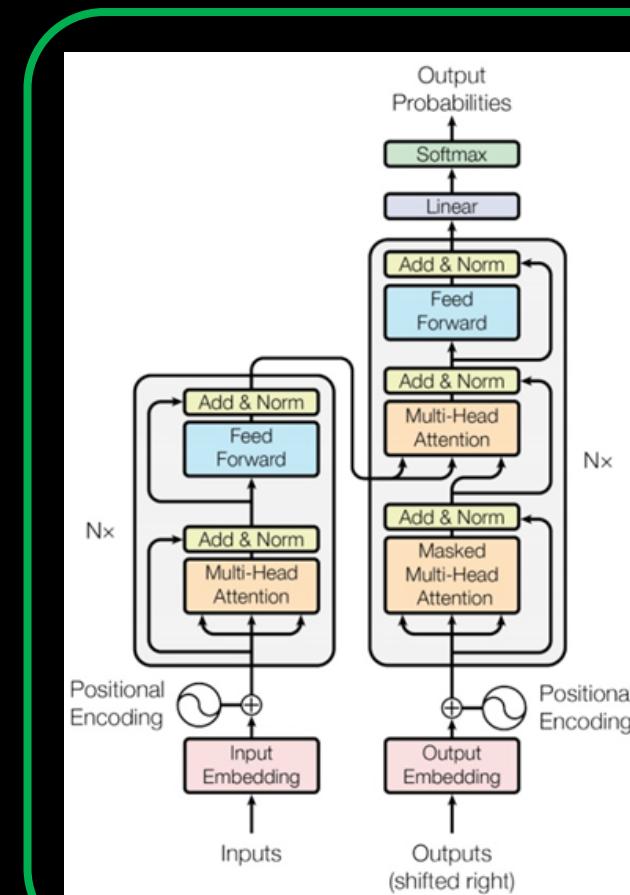
Pre-trained  
**ChatGPT**

# Transformer ChatGPT

# ChatGPT is a function



$$= f( \text{ } \cdots \text{ } ) =$$



Billions of  
Parameters  
Transformer

# What is the relationship between ChatGPT and Large Language Models?

What is the relationship between  
ChatGPT and Large Language Models?

Socialization

# How to create a ChatGPT?

# How to create a **ChatGPT**?

**Set scope**

Define the set of candidate functions

**Deep Learning (CNN, Transformer...),  
Decision Tree, etc.**

**Establish criteria**

Define the criteria for evaluating the quality of functions

**Supervised Learning, Semi-supervised  
Learning, Reinforcement Learning, etc.**

**Achieve the goal**

Find the best function → Optimization

**Gradient Descent, Genetic Algorithm, etc.**

# Supervised Learning?

We need to prepare **paired** data.

## Input

What is the capital of New Zealand?

Please help me polish the following  
text: .....

How to crack the Wi-Fi password of my  
neighbor's house?

## Output

Wellington.

Yes, here's the polished text: .....

I'm sorry, but I cannot assist with hacking or  
unauthorized access to networks or devices, as it is  
illegal and unethical.

# Supervised Learning?

## Input

What is the capital of New Zealand?

Please help me polish the following text: .....

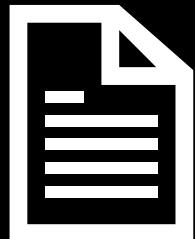
How to crack the Wi-Fi password of my neighbor's house?

## Output

Wellington.

Yes, here's the polished text: .....

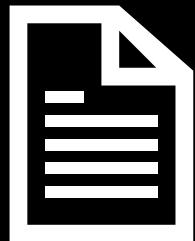
I'm sorry, but I cannot assist with hacking or unauthorized access to networks or devices, as it is illegal and unethical.



$= f( \text{ } \dots \text{ } ) =$

**Billions of Parameters**

# Supervised Learning?



$= f( \text{ } \dots \text{ } ) =$

**Billions of Parameters**

**GPT 1**

**2018**

**GPT 2**

**2019**

**GPT 3**

**2020**

**GPT 4**

**2023**

**Model Size  
(Parameters)**

**117M**

**1.5B**

**175B**

**8\*220B?**

**Training Data**

**1 GB**

**40 GB**

**45TB**  
**570GB(after Filtering)**

**?**

# Supervised Learning?



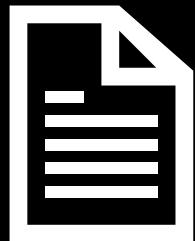
$= f( \text{ } \dots \text{ } ) =$

**Billions of Parameters**

	GPT 1 2018	GPT 2 2019	GPT 3 2020	GPT 4 2023
Model Size (Parameters)	117M	1.5B	175B	8*220B?
Training Data	1 GB	40 GB	45TB 570GB(after Filtering)	?

**The paired data that humans  
can provide is limited.**

# Supervised Learning?



$= f( \text{ } \dots \text{ } ) =$

**Billions of Parameters**

	GPT 1 2018	GPT 2 2019	GPT 3 2020	GPT 4 2023
Model Size (Parameters)	117M	1.5B	175B	8*220B?
Training Data	1 GB	40 GB	45TB 570GB(after Filtering)	?

## How to create paired data?

# Pre-trained ChatGPT

**The capital of New Zealand is Wellington.**

**The wind in Wellington is so strong that it makes  
me want to go home.**

**The capital of New Zealand is Wellington.**

**The wind in Wellington is so strong that it makes  
me want to go home.**

**The capital of New Zealand is Wellington.**

**The wind in Wellington is so strong that it makes  
me want to go home.**



**Any piece of text can be used to teach a machine to continue a text chain.**



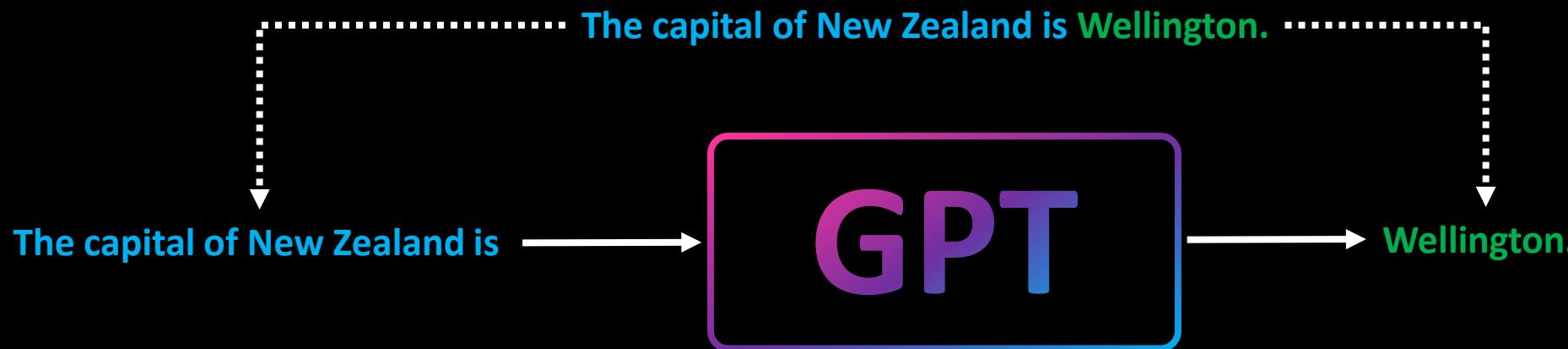
Any piece of text can be used to teach a machine to continue a text chain.

Pre-trained (Self-supervised Learning)

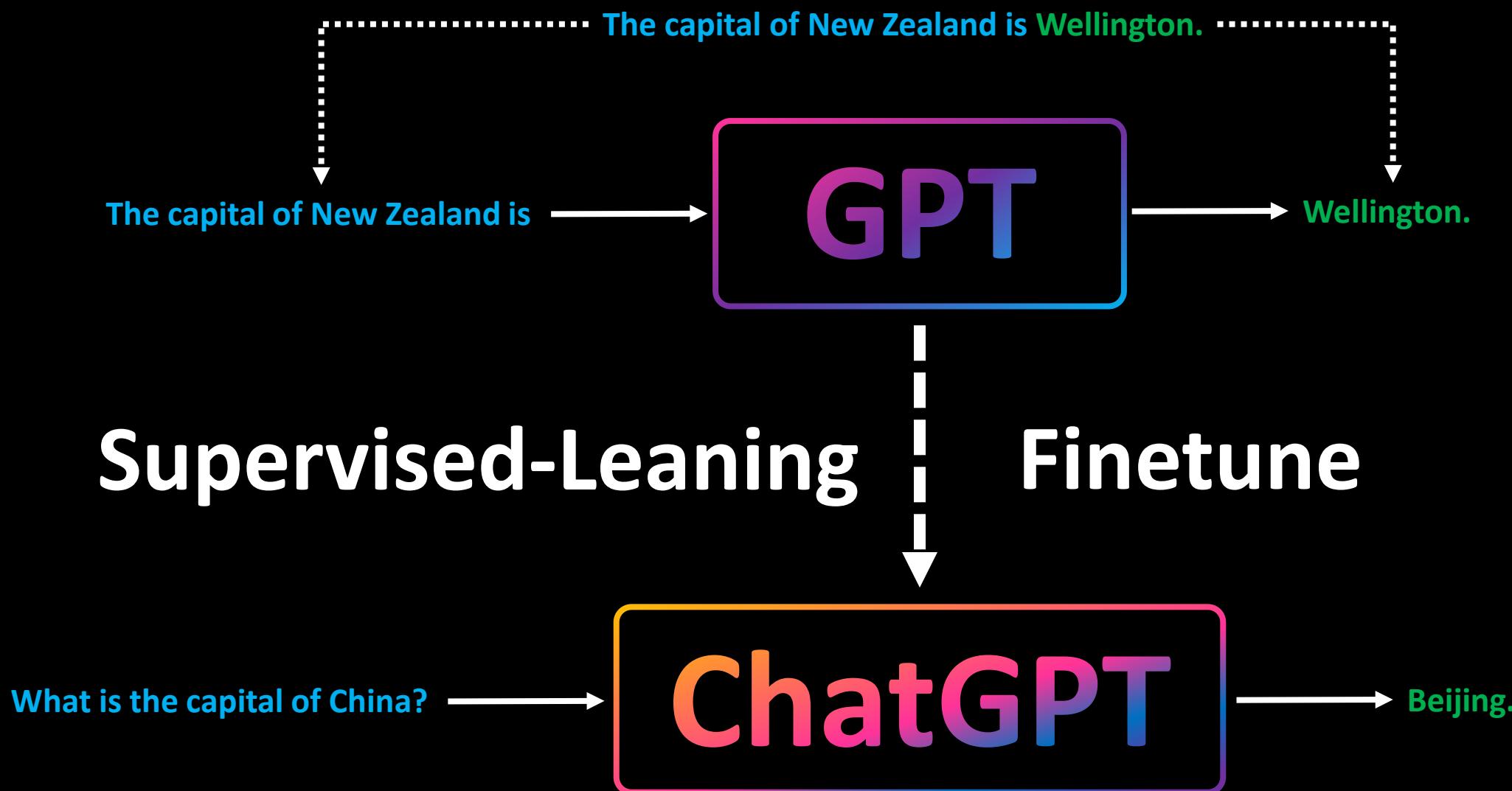


What is the relationship between  
ChatGPT and Large Language Models?

Socialization



**At this point, the model has become a Large Language Model, but currently, it only possesses world knowledge and does not know how to answer questions like human.**



The capital of New Zealand is Wellington.

The capital of New Zealand is



Wellington.

Supervised-Leaning

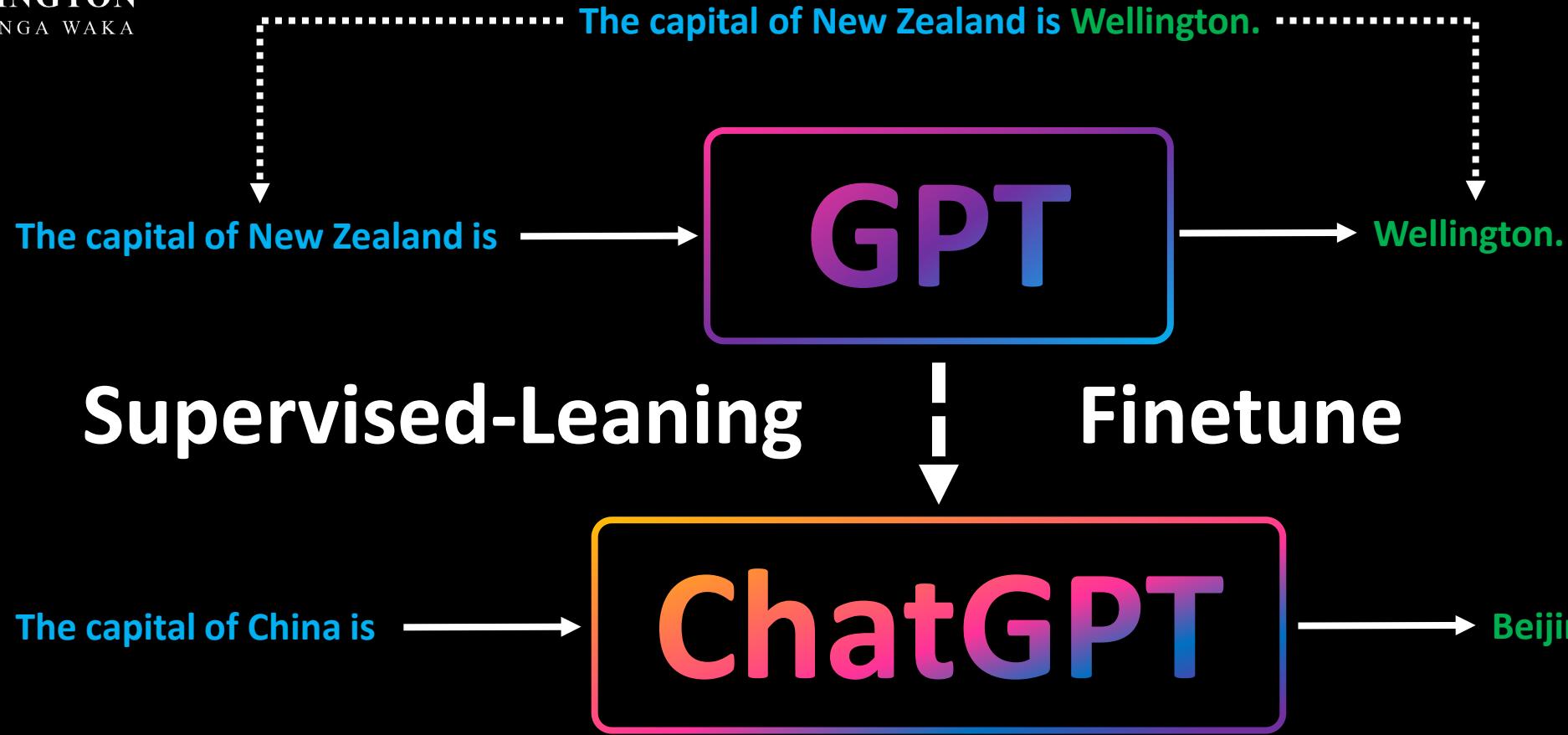
Finetune

The capital of China is

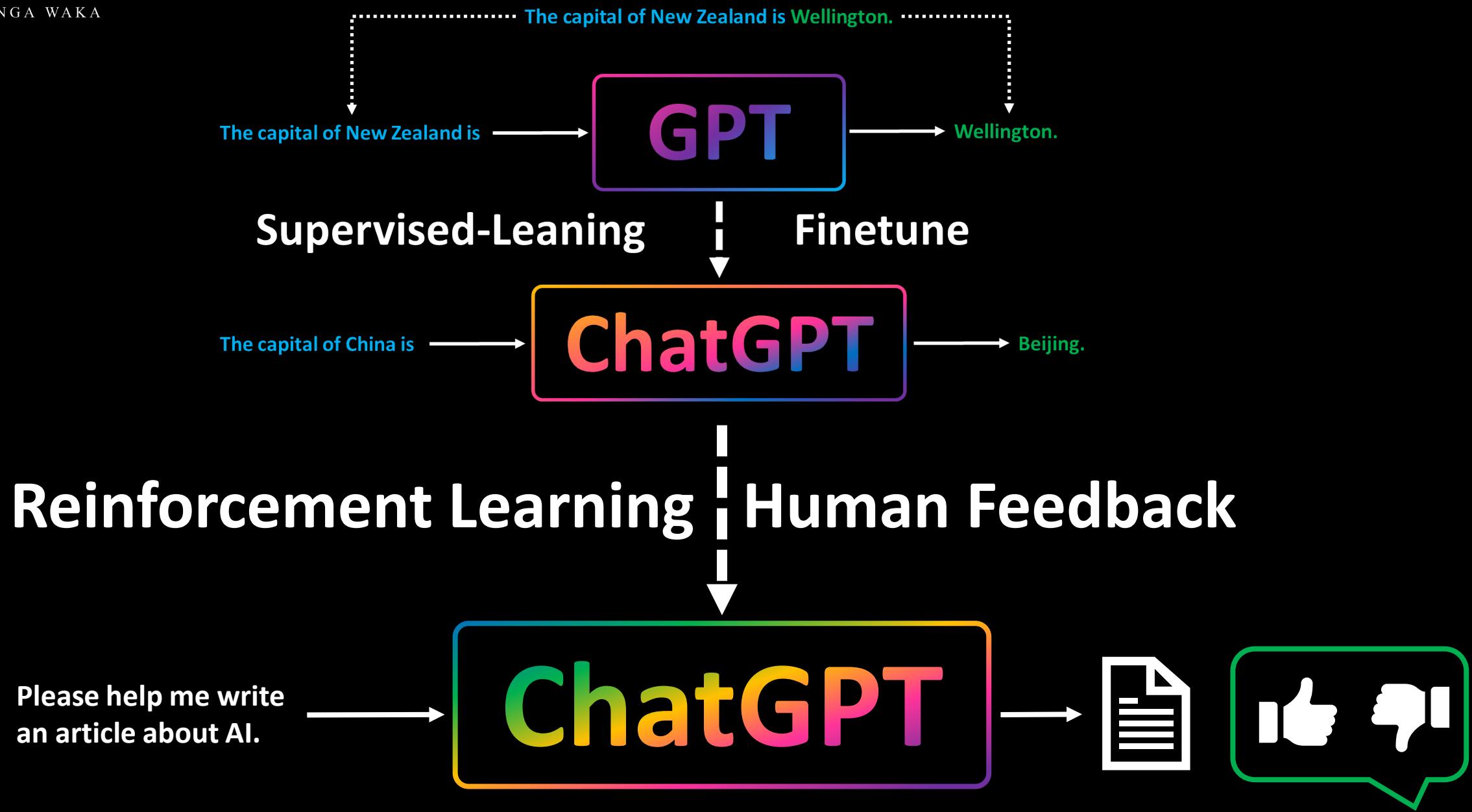


Beijing.

At this point, the Large Language Model has learned to answer questions like a human, completing its **socialization** process.



# How to make a good answer?



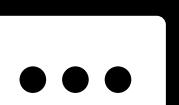
Now that Artificial Intelligence has  
evolved from a "**tool**" to a "**helper**,"  
what actions can we take?

**Fuck the world if you are rich,  
otherwise fuck yourself.**

**穷则独善其身，达则兼济天下**

If we can't change the model,  
we change ourselves.

Better Prompt, More Information

=  $f($     $)$

↑

Model (e.g., ChatGPT)

# Prompt Engineering



**Better Prompt, More Information**

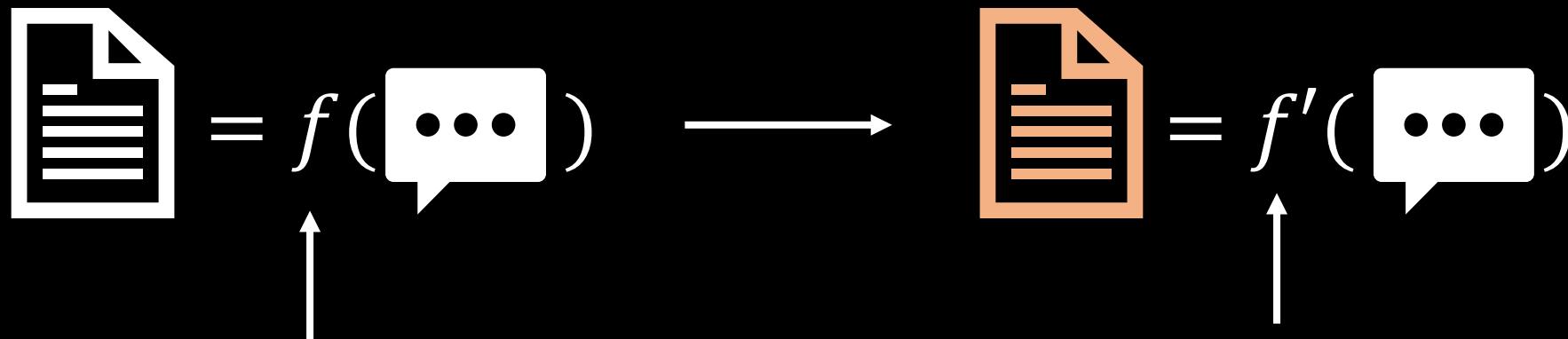
$$= f( \text{ } \text{ } \text{ } \text{ } \text{ } )$$

↑

A Model (e.g., ChatGPT)

The Art of Human Communication with  
Generative AI

# Train our own Model



An Opensource Model  
(e.g., LLaMA)

Adjust the parameters.

If we can't change the model, we change ourselves.



Better Prompt, More Information

$$= f( \text{ } \text{ } \text{ } \text{ } )$$

↑

Low Cost

A Model (e.g., ChatGPT)

Train our own Model



$$= f( \text{ } \text{ } \text{ } \text{ } )$$

↑



$$= f'( \text{ } \text{ } \text{ } \text{ } )$$

↑

High Cost

An Opensource Model  
(e.g., LLaMA)

Adjust the parameters.

# Two different expectations of Large Language Models

## Become an expert in a specific field, dedicated to solving a particular task.

Fuck the world if you are rich,  
otherwise fuck yourself. → **LLM** → 穷则独善其身，达则兼济天下

solve XXX problem → **LLM** → < / >

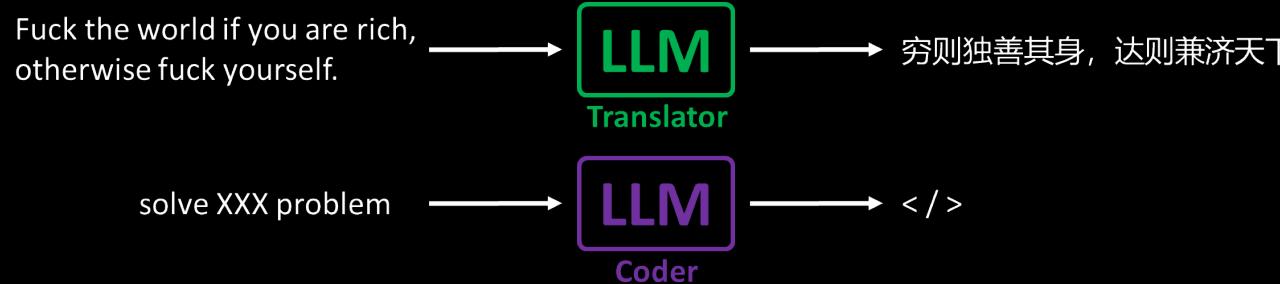
## Become a generalist

Please Translate:

Fuck the world if you are rich,  
otherwise fuck yourself. → **LLM** → 如果你很富有，那么去享受世界；  
如果不是，那么只能自求多福。

Please write a code to  
solve XXX problem → **LLM** → < / >

**Become an expert in a specific field, dedicated to solving a particular task.**



**GitHub Copilot**

**Become a generalist**



**ChatGPT**



$= f( \text{ } \dots \text{ } ) =$

**Billions of Parameters**

**GPT 1**

**2018**

**GPT 2**

**2019**

**GPT 3**

**2020**

**GPT 4**

**2023**

**Model Size  
(Parameters)**

**117M**

**1.5B**

**175B**

**8\*220B?**

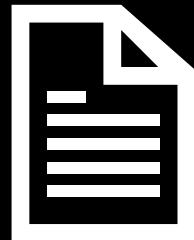
**Training Data**

**1 GB**

**40 GB**

**45TB**  
570GB(after Filtering)

**?**



$= f( \text{ } \dots \text{ } ) =$

**Billions of Parameters**

GPT 1

GPT 2

GPT 3

GPT 4

2018

2019

2020

2023

## Mixture of Experts model

Model Size  
(Parameters)

**(MoE)**

1 GB

40 GB

175B

8\*220B?

Training Data

1.17M

1.5B

45TB

570GB(after Filtering)

?

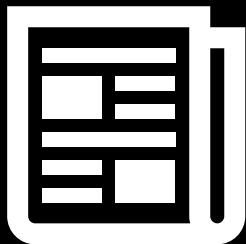
I strongly recommend that everyone  
learns to use Large Language Models  
efficiently as your assistant.

PART IV

# Generate Pictures

## Diffusion Models

# Generative AI is to enable machines to **generate** complex and structured things.



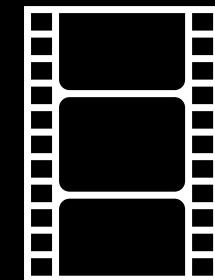
Tokens



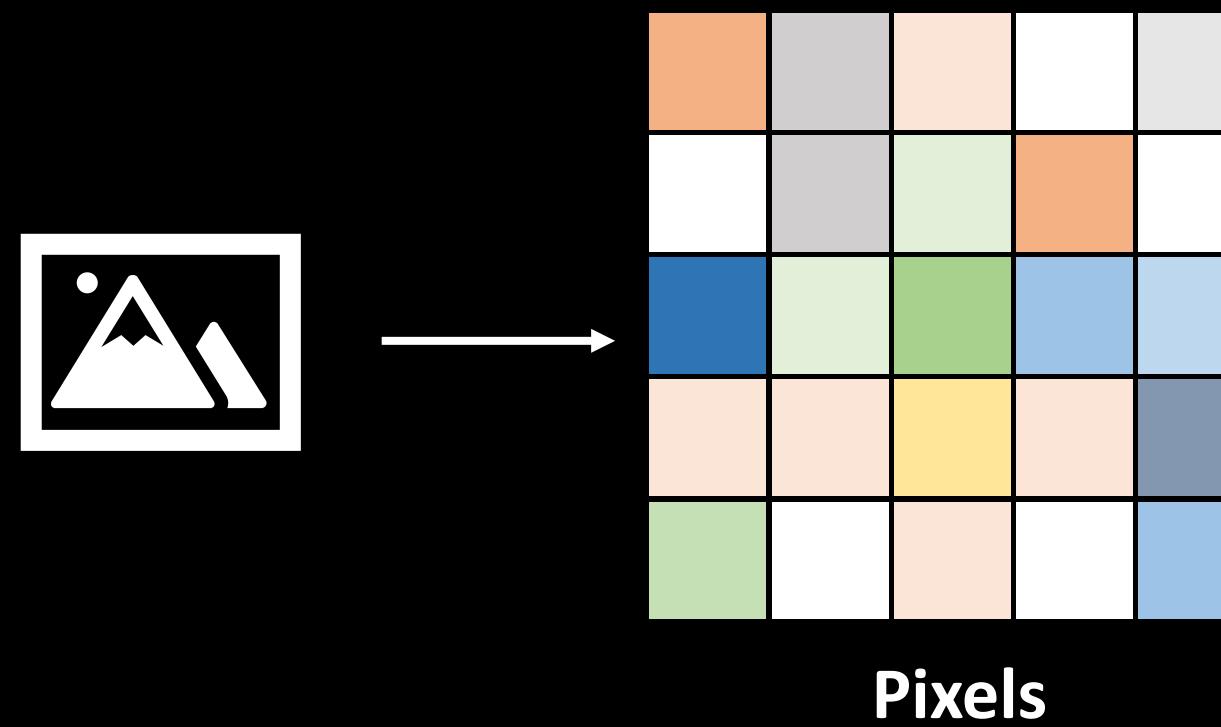
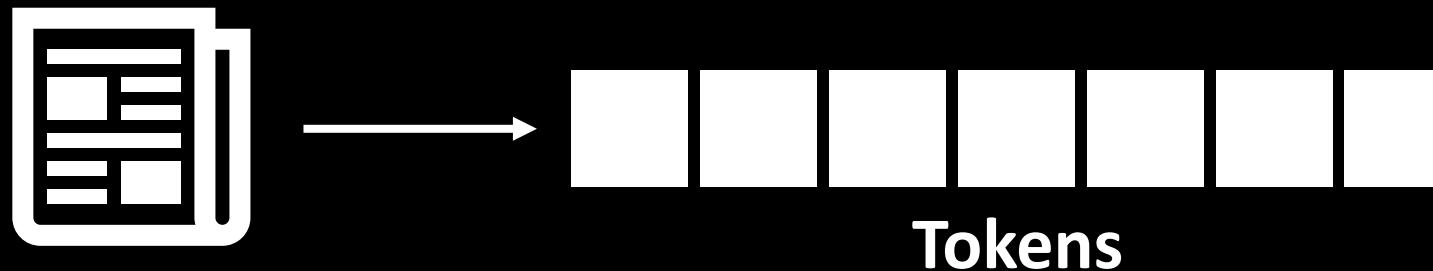
Pixels



Samples



Frames



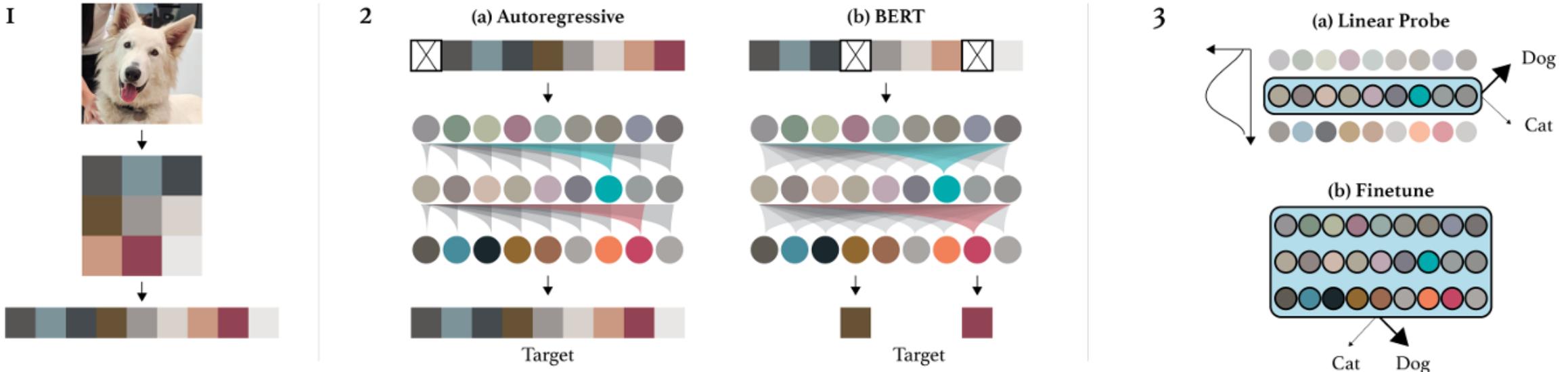


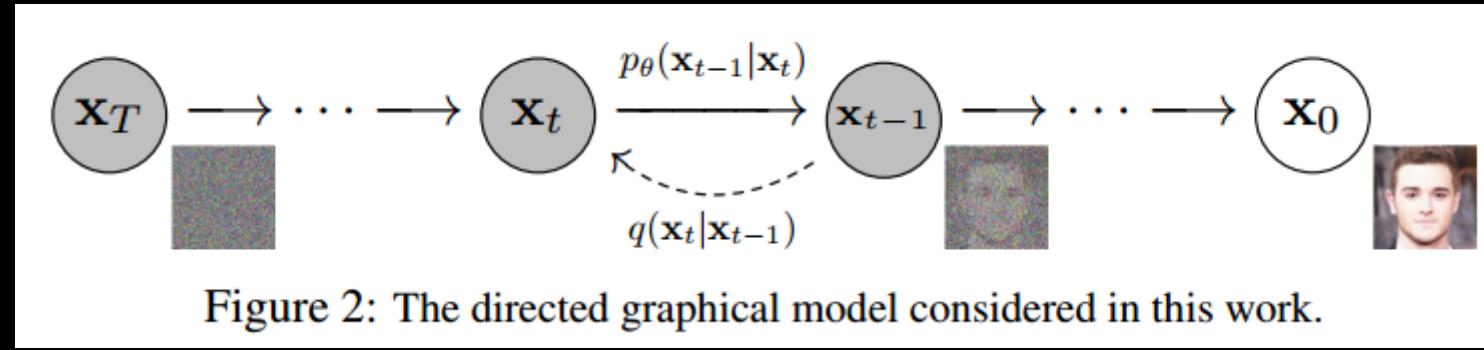
Figure 1. An overview of our approach. First, we pre-process raw images by resizing to a low resolution and reshaping into a 1D sequence. We then chose one of two pre-training objectives, auto-regressive next pixel prediction or masked pixel prediction. Finally, we evaluate the representations learned by these objectives with linear probes or fine-tuning.

# Autoregressive Generation

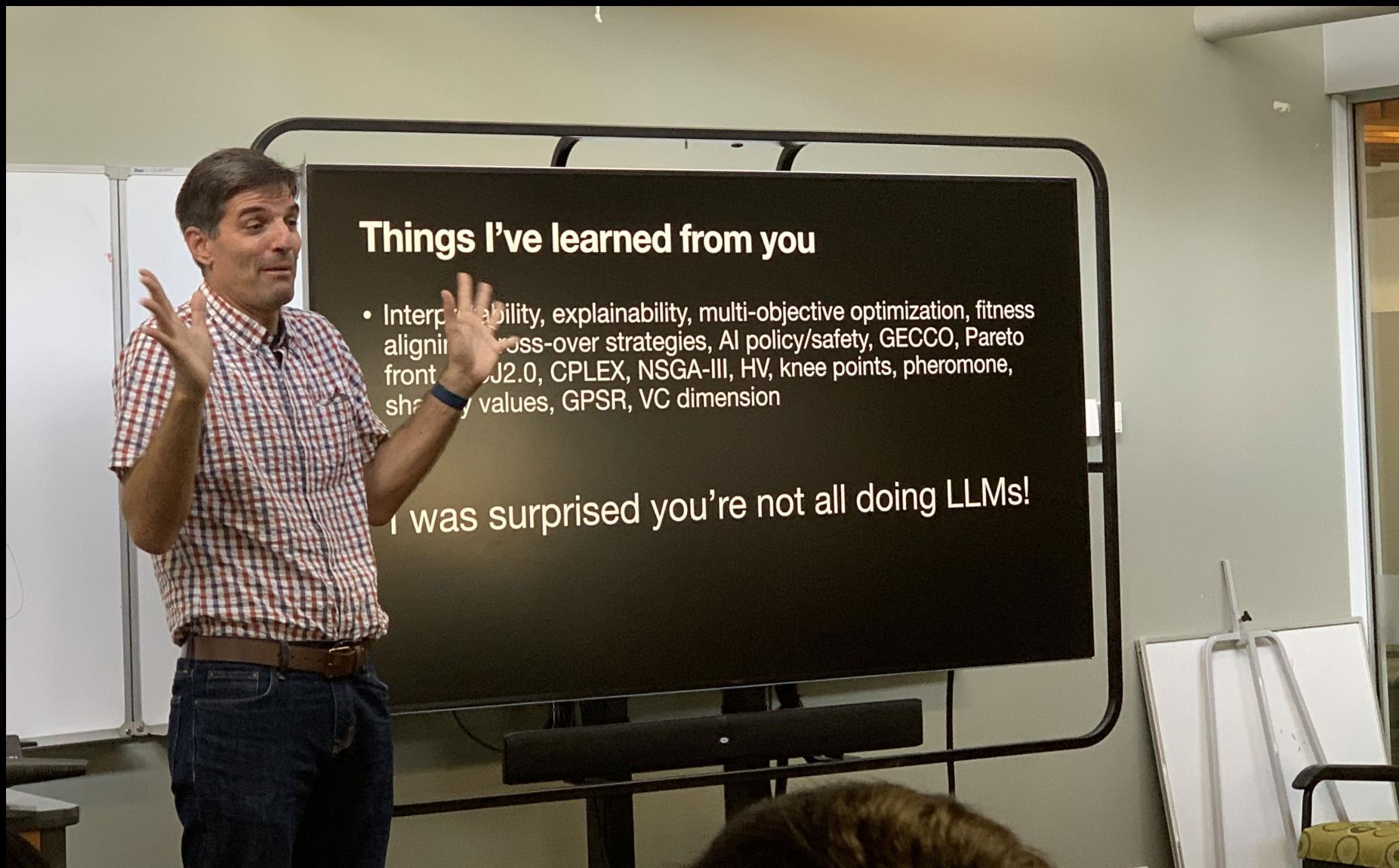


It works well.





## Diffusion Models



**Thank you for your listening!**

Bocheng Lin 2024.3.13

Thank you, Prof. Tom!

# Any Questions?

Thank you for your listening!