



# Introduction to Foundation Models

*Generative AI Foundations on AWS*

Emily Webber, Principal ML Specialist SA at AWS

Lesson 1 – Level 300



# Topics in our journey together

1. Primer on foundation models

2. How to pick a foundation model

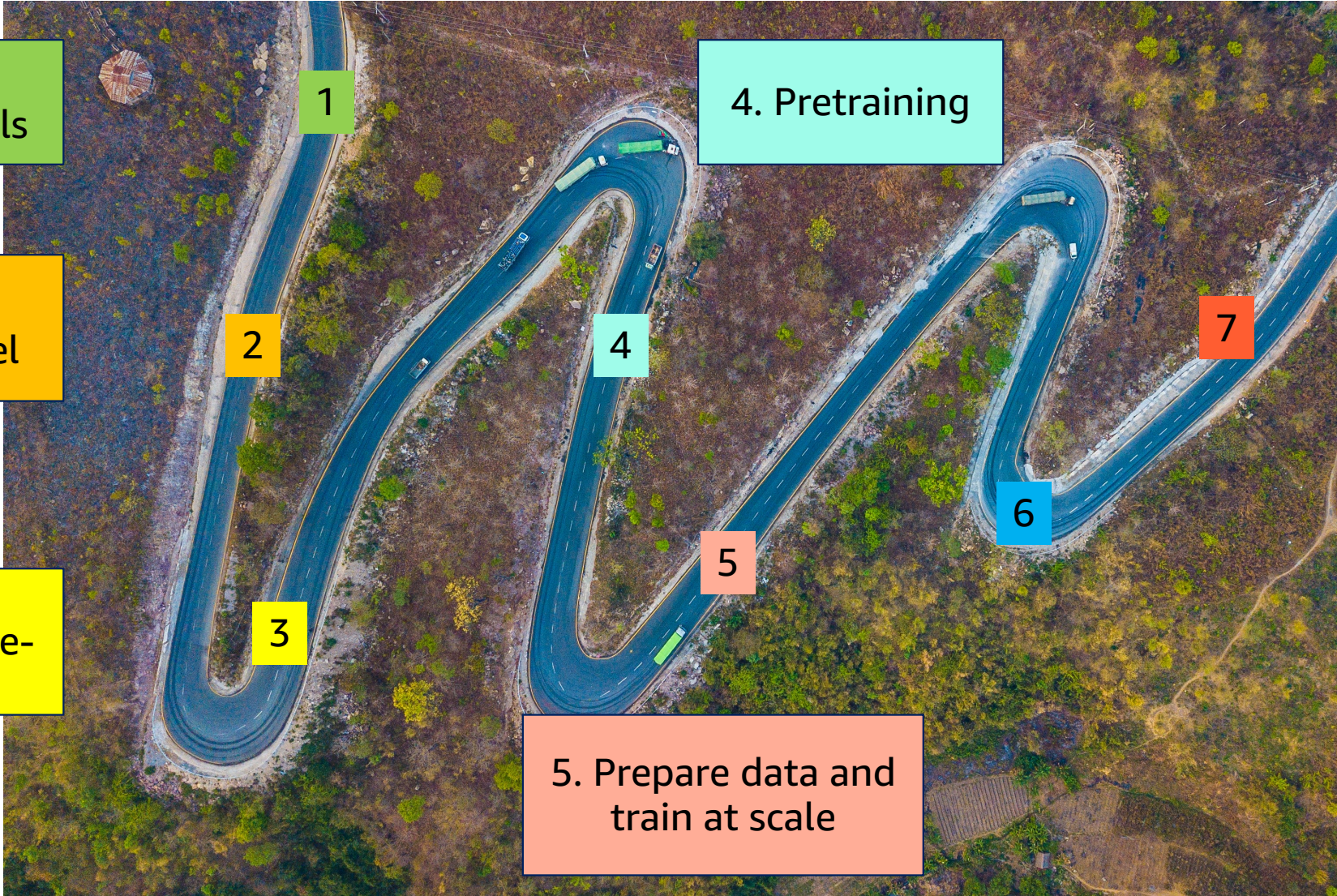
3. Prompt engineering & fine-tuning

4. Pretraining

5. Prepare data and train at scale

7. Distributed hosting

6. Reinforcement learning with human feedback





# Today's activities



- What foundation models are
- Where they come from
- How they impact Generative AI
- Foundation model lifecycle
- Hands-on walk through: Foundation models on Amazon SageMaker

Let's say I asked you to learn **everything** on the internet.  
How would you do it?



Structure



Storage



Time

5.74 B pages x 52 seconds  
= ~83,000 hours

=> **~41,000 human years**

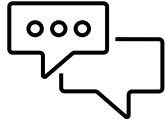
A foundation model can  
do this in a few months.

# You can do a lot with foundation models!

---



Text  
generation



Q&A



Text  
summarization



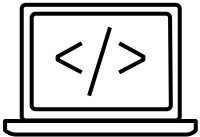
Text  
extraction



Paraphrase  
rephrase



Search



Code  
generation



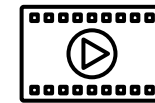
Image  
generation



Image  
classification



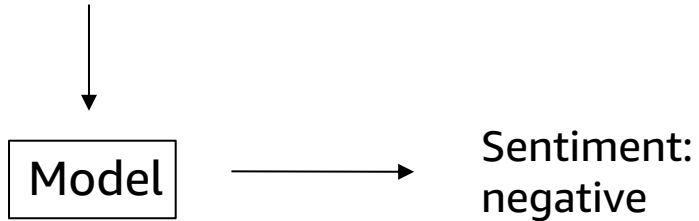
Audio  
generation



Video  
generation

# Many ML tasks can now be **re-cast as generative**, and most will benefit from foundation models.

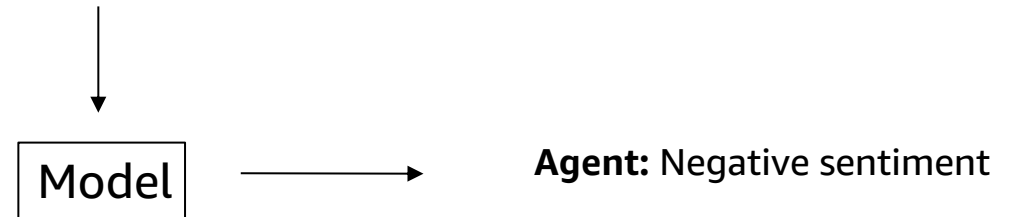
**Text:** I am not into this house; it's way too expensive and too far from the train line!



Traditional classification

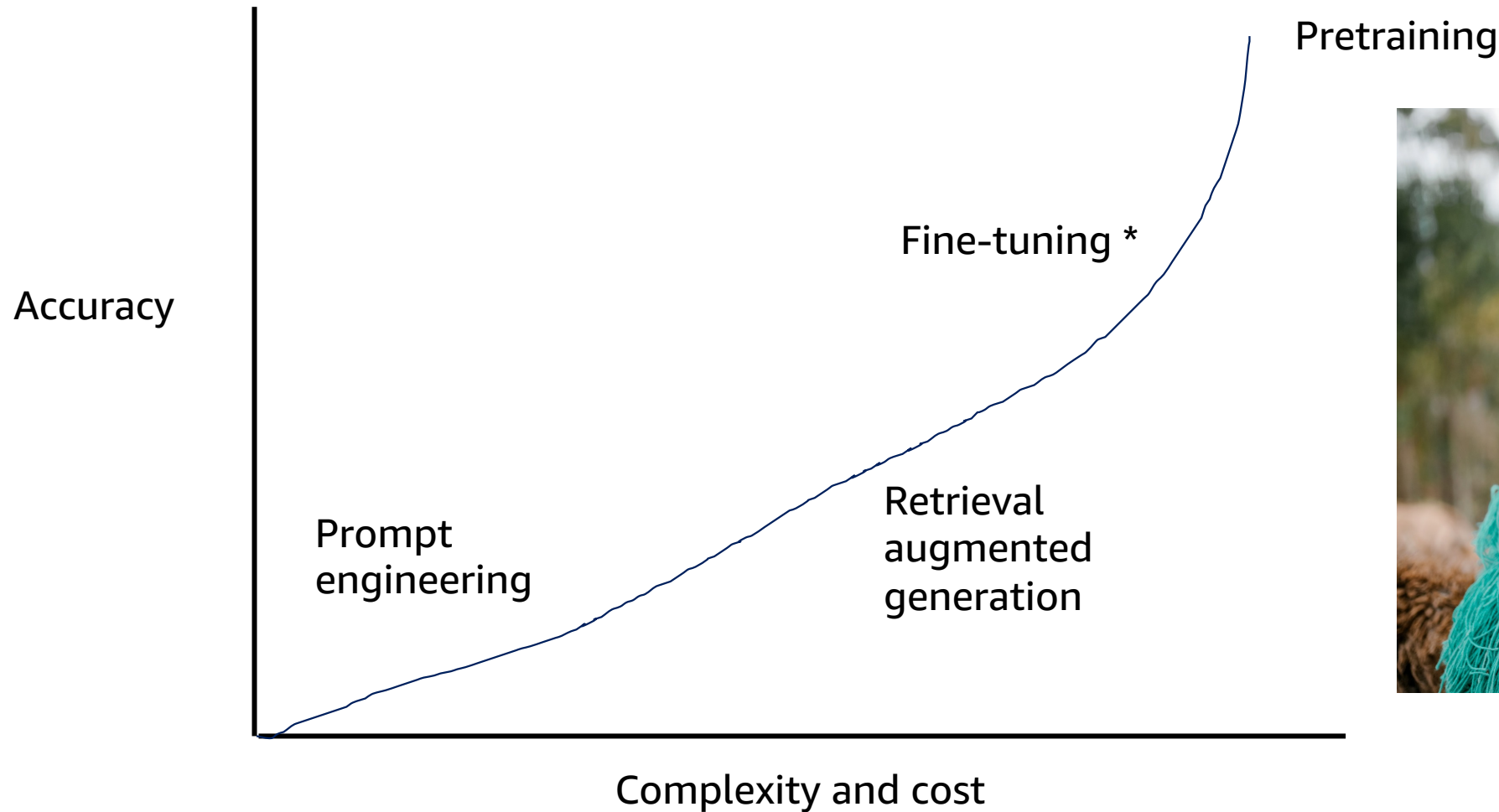
**Text:** I am not into this house; it's way too expensive and too far from the train line!

Classify this sentence into positive or negative sentiment:

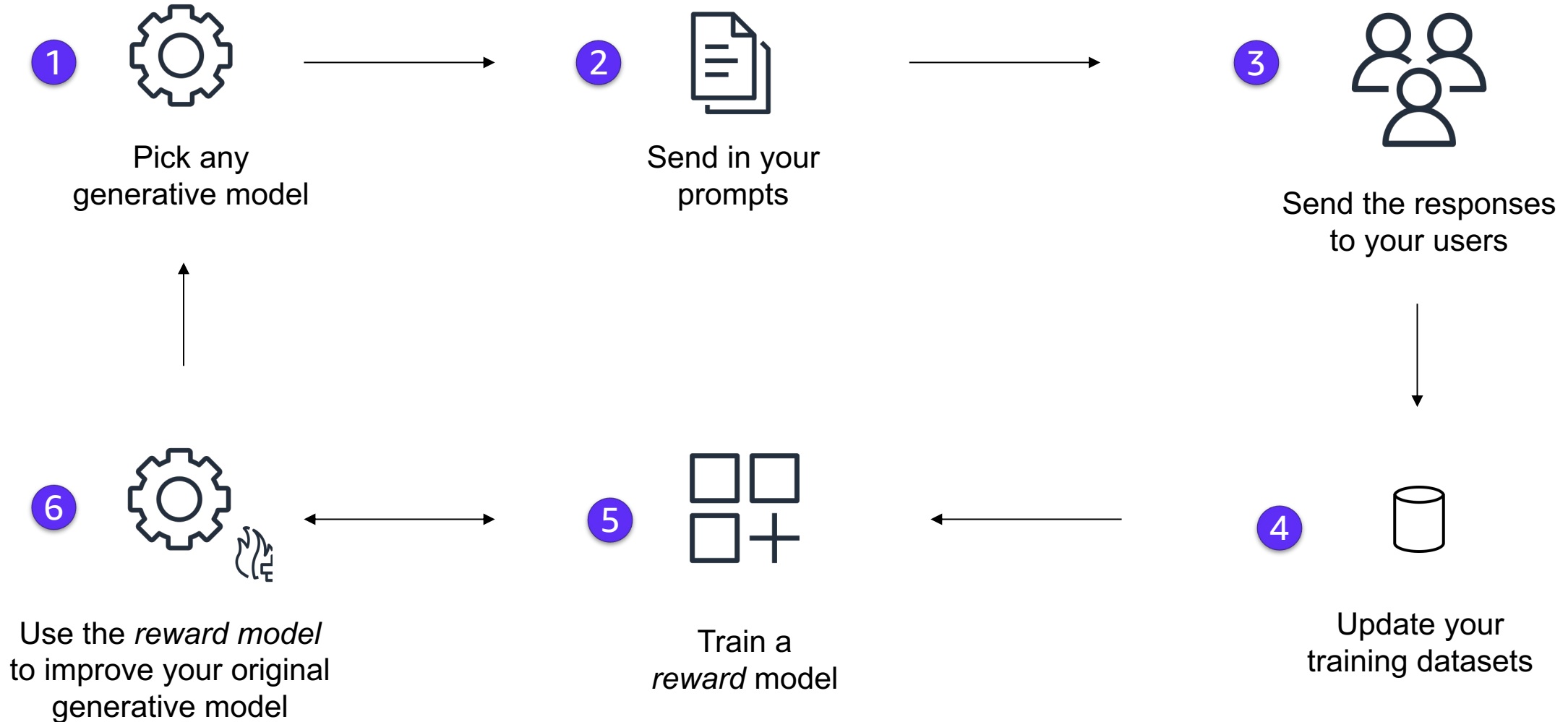


Using generation to classify text

# There are many ways to customize a foundation model



# The best generative models are built on human feedback





# Foundation model spotlight: Stable Diffusion

**Prompt:** landscape of the beautiful city of paris rebuilt near the pacific ocean in sunny california, amazing weather, sandy beach, palm trees, splendid haussmann architecture, digital painting, highly detailed, intricate, without duplication, concept art, matte painting, trending on artstation

**Negative prompts:** trees, green

▼ General info

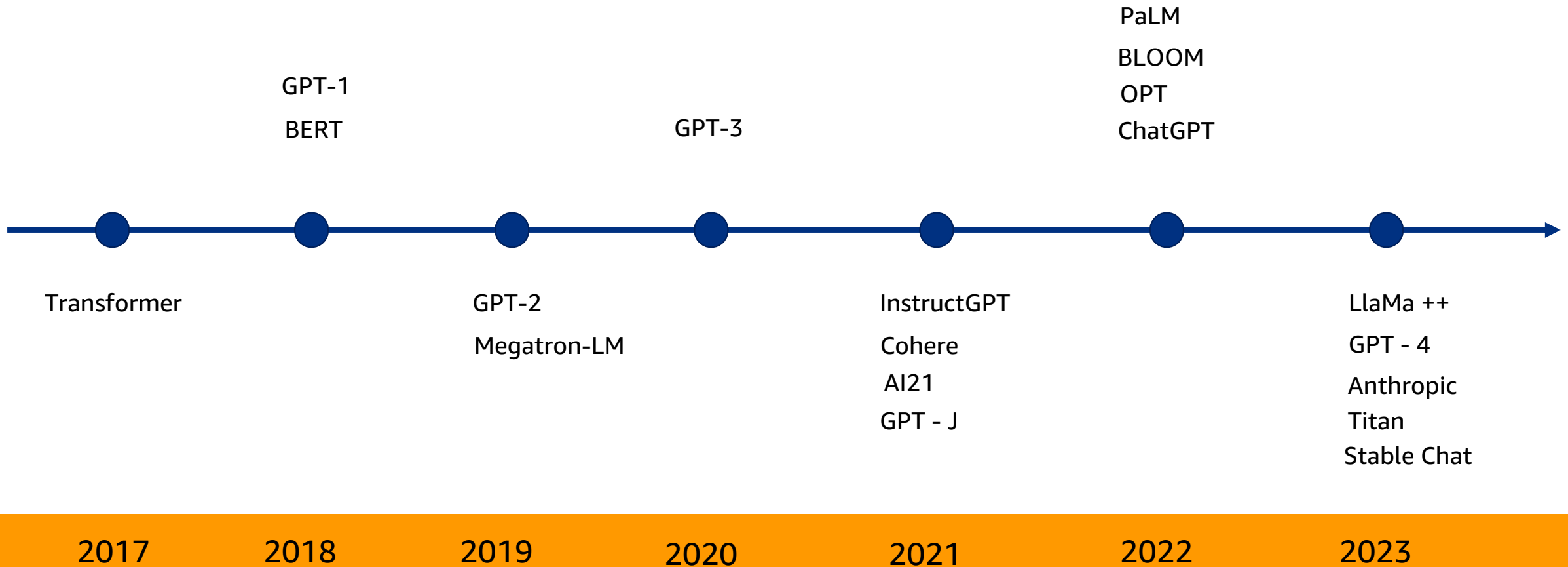
Width	720
Height	512
Number of inference steps	50
Guidance scale	7
Seed	1



**Stable Diffusion-generated image**

# Timeline of language foundation models

GPT = Generative Pretrained Transformers



# Foundation model spotlight: AI21 Jurassic-2 Jumbo Instruct

## Prompt

Tell me a story about a dog running down the street.

▶ Generate text

## Output

Once upon a time, there was a dog named Max. Max was a very happy dog, and he loved to run.

One day, Max was out for a walk with his favorite person, his owner, Mr. Jones. Mr. Jones was holding Max's leash, and Max was pulling him down the street.

Max was so excited, he couldn't stand still. He just wanted to run and run.

Mr. Jones was having a hard time keeping up with Max. He was getting tired, but he didn't want to let Max go.

## Prompt

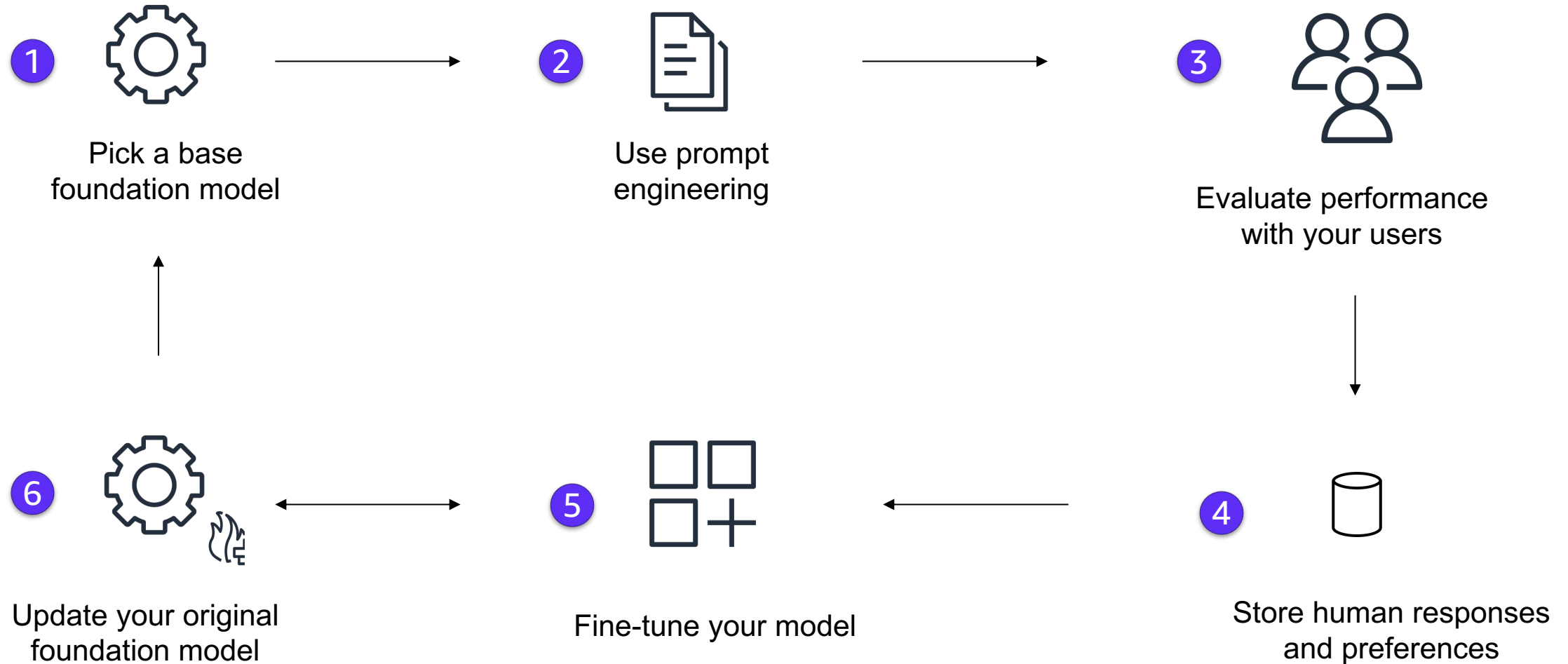
If you found two shoes, one for the right foot and one for the left foot, how many shoes would you have?

▶ Generate text

## Output

If you found two shoes, one for the right foot and one for the left foot, you would have two shoes.

# Foundation model lifecycle



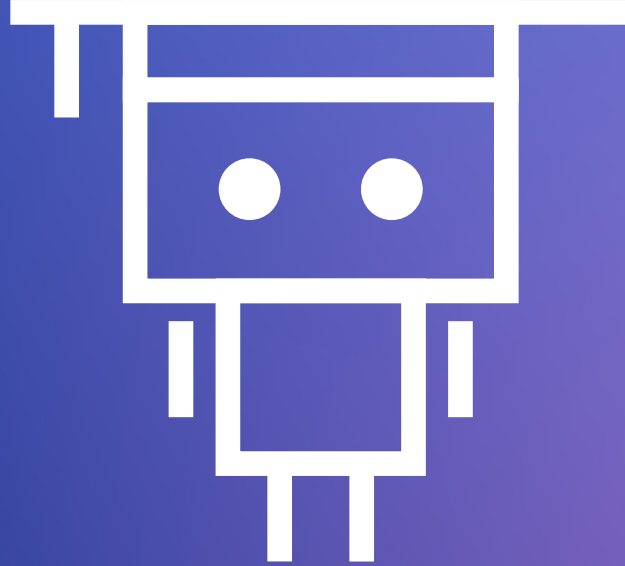




<https://bit.ly/sm-nb-1>

**Hands-on demo**





# Thank you!

Type: Corrections, feedback, or other questions?  
Contact us at <https://support.awsamazon.com/#/contacts/aws-academy>.  
All trademarks are the property of their owners.