



# Generative AI with Amazon Bedrock

Ganesh Gella

Director of Engineering, Amazon Lex &  
Amazon Bedrock Agents, AWS

तोहय छिवा वारय

सुवाठें

मीरु एला क्कनारु

भवान् कथमसि

सम कुशल मंगल

थुआढ़ा केहू हाल ऐ

आपोनाव केने

म बर यू डोंग

तमे केम छे

नीवु कॅगिद्धीरि

आप कैसे हैं

How are you?

तून कीैन आहीन

केमिति थाळति, केमिति थाळ

आप कैसे हो

सेतलेका मेनमा?

तू कसा आहेस

तिमीलाई कस्तो छ

आपुनि केनि आने?

आपनि केनन आज्जेन

उमी किव्वे हे

एप्पाटि इयुक्कीर्ग्कलं

कएम डबबरी बे गे



The diagram illustrates the flow of execution between memory locations and assembly code. Key components include:

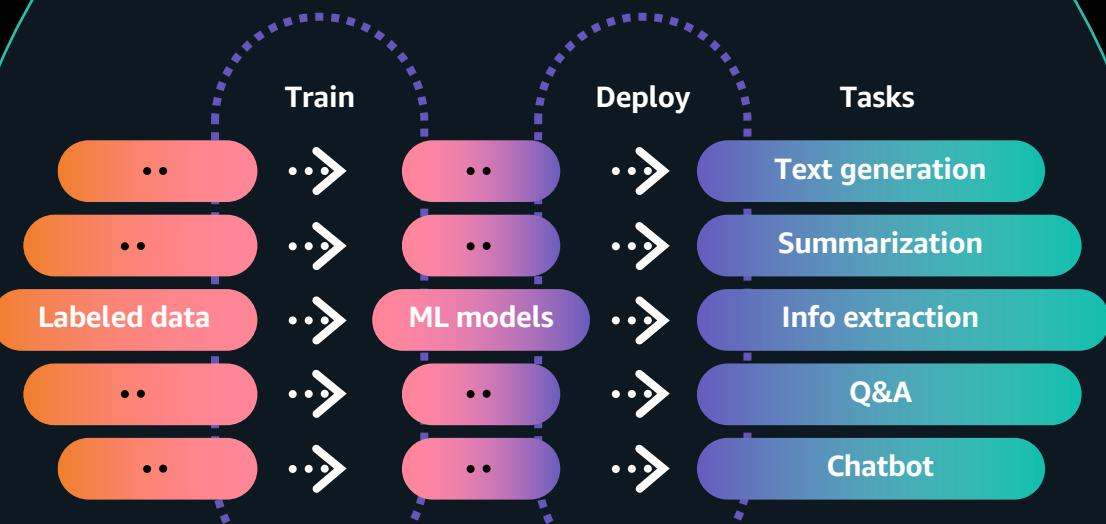
- Memory Locations:** Labeled with addresses like #6, #12, #8, and #0.
- Assembly Code Snippets:** Such as `STA`, `LDA`, `JMP`, `RTS`, `BNE`, `BEQ`, `CMP`, `BCC`, `CPY`, `BCS`, `JSR`, `SBC`, `SEC`, `STC`, `AND`, `OR`, `XOR`, `NOT`, `SHL`, `SHR`, `SWAP`, `SWAPB`, `SWAPW`, `SWAPD`, `SWAPQ`, `SWAPH`, `SWAPL`, `SWAPB8`, `SWAPW8`, `SWAPD8`, `SWAPQ8`, `SWAPH8`, `SWAPL8`, `SWAPB16`, `SWAPW16`, `SWAPD16`, `SWAPQ16`, `SWAPH16`, `SWAPL16`, `SWAPB32`, `SWAPW32`, `SWAPD32`, `SWAPQ32`, `SWAPH32`, `SWAPL32`, `SWAPB64`, `SWAPW64`, `SWAPD64`, `SWAPQ64`, `SWAPH64`, `SWAPL64`, `SWAPB128`, `SWAPW128`, `SWAPD128`, `SWAPQ128`, `SWAPH128`, `SWAPL128`, `SWAPB256`, `SWAPW256`, `SWAPD256`, `SWAPQ256`, `SWAPH256`, `SWAPL256`, `SWAPB512`, `SWAPW512`, `SWAPD512`, `SWAPQ512`, `SWAPH512`, `SWAPL512`, `SWAPB1024`, `SWAPW1024`, `SWAPD1024`, `SWAPQ1024`, `SWAPH1024`, `SWAPL1024`, `SWAPB2048`, `SWAPW2048`, `SWAPD2048`, `SWAPQ2048`, `SWAPH2048`, `SWAPL2048`, `SWAPB4096`, `SWAPW4096`, `SWAPD4096`, `SWAPQ4096`, `SWAPH4096`, `SWAPL4096`, `SWAPB8192`, `SWAPW8192`, `SWAPD8192`, `SWAPQ8192`, `SWAPH8192`, `SWAPL8192`, `SWAPB16384`, `SWAPW16384`, `SWAPD16384`, `SWAPQ16384`, `SWAPH16384`, `SWAPL16384`, `SWAPB32768`, `SWAPW32768`, `SWAPD32768`, `SWAPQ32768`, `SWAPH32768`, `SWAPL32768`, `SWAPB65536`, `SWAPW65536`, `SWAPD65536`, `SWAPQ65536`, `SWAPH65536`, `SWAPL65536`, `SWAPB131072`, `SWAPW131072`, `SWAPD131072`, `SWAPQ131072`, `SWAPH131072`, `SWAPL131072`, `SWAPB262144`, `SWAPW262144`, `SWAPD262144`, `SWAPQ262144`, `SWAPH262144`, `SWAPL262144`, `SWAPB524288`, `SWAPW524288`, `SWAPD524288`, `SWAPQ524288`, `SWAPH524288`, `SWAPL524288`, `SWAPB1048576`, `SWAPW1048576`, `SWAPD1048576`, `SWAPQ1048576`, `SWAPH1048576`, `SWAPL1048576`, `SWAPB2097152`, `SWAPW2097152`, `SWAPD2097152`, `SWAPQ2097152`, `SWAPH2097152`, `SWAPL2097152`, `SWAPB4194304`, `SWAPW4194304`, `SWAPD4194304`, `SWAPQ4194304`, `SWAPH4194304`, `SWAPL4194304`, `SWAPB8388608`, `SWAPW8388608`, `SWAPD8388608`, `SWAPQ8388608`, `SWAPH8388608`, `SWAPL8388608`, `SWAPB16777216`, `SWAPW16777216`, `SWAPD16777216`, `SWAPQ16777216`, `SWAPH16777216`, `SWAPL16777216`, `SWAPB33554432`, `SWAPW33554432`, `SWAPD33554432`, `SWAPQ33554432`, `SWAPH33554432`, `SWAPL33554432`, `SWAPB67108864`, `SWAPW67108864`, `SWAPD67108864`, `SWAPQ67108864`, `SWAPH67108864`, `SWAPL67108864`, `SWAPB134217728`, `SWAPW134217728`, `SWAPD134217728`, `SWAPQ134217728`, `SWAPH134217728`, `SWAPL134217728`, `SWAPB268435456`, `SWAPW268435456`, `SWAPD268435456`, `SWAPQ268435456`, `SWAPH268435456`, `SWAPL268435456`, `SWAPB536870912`, `SWAPW536870912`, `SWAPD536870912`, `SWAPQ536870912`, `SWAPH536870912`, `SWAPL536870912`, `SWAPB1073741840`, `SWAPW1073741840`, `SWAPD1073741840`, `SWAPQ1073741840`, `SWAPH1073741840`, `SWAPL1073741840`, `SWAPB2147483680`, `SWAPW2147483680`, `SWAPD2147483680`, `SWAPQ2147483680`, `SWAPH2147483680`, `SWAPL2147483680`, `SWAPB4294967360`, `SWAPW4294967360`, `SWAPD4294967360`, `SWAPQ4294967360`, `SWAPH4294967360`, `SWAPL4294967360`, `SWAPB8589934720`, `SWAPW8589934720`, `SWAPD8589934720`, `SWAPQ8589934720`, `SWAPH8589934720`, `SWAPL8589934720`, `SWAPB17179869440`, `SWAPW17179869440`, `SWAPD17179869440`, `SWAPQ17179869440`, `SWAPH17179869440`, `SWAPL17179869440`, `SWAPB34359738880`, `SWAPW34359738880`, `SWAPD34359738880`, `SWAPQ34359738880`, `SWAPH34359738880`, `SWAPL34359738880`, `SWAPB68719477760`, `SWAPW68719477760`, `SWAPD68719477760`, `SWAPQ68719477760`, `SWAPH68719477760`, `SWAPL68719477760`, `SWAPB137438955200`, `SWAPW137438955200`, `SWAPD137438955200`, `SWAPQ137438955200`, `SWAPH137438955200`, `SWAPL137438955200`, `SWAPB274877910400`, `SWAPW274877910400`, `SWAPD274877910400`, `SWAPQ274877910400`, `SWAPH274877910400`, `SWAPL274877910400`, `SWAPB549755820800`, `SWAPW549755820800`, `SWAPD549755820800`, `SWAPQ549755820800`, `SWAPH549755820800`, `SWAPL549755820800`, `SWAPB1099511641600`, `SWAPW1099511641600`, `SWAPD1099511641600`, `SWAPQ1099511641600`, `SWAPH1099511641600`, `SWAPL1099511641600`, `SWAPB2199023283200`, `SWAPW2199023283200`, `SWAPD2199023283200`, `SWAPQ2199023283200`, `SWAPH2199023283200`, `SWAPL2199023283200`, `SWAPB4398046566400`, `SWAPW4398046566400`, `SWAPD4398046566400`, `SWAPQ4398046566400`, `SWAPH4398046566400`, `SWAPL4398046566400`, `SWAPB8796093132800`, `SWAPW8796093132800`, `SWAPD8796093132800`, `SWAPQ8796093132800`, `SWAPH8796093132800`, `SWAPL8796093132800`, `SWAPB17592186265600`, `SWAPW17592186265600`, `SWAPD17592186265600`, `SWAPQ17592186265600`, `SWAPH17592186265600`, `SWAPL17592186265600`, `SWAPB35184372531200`, `SWAPW35184372531200`, `SWAPD35184372531200`, `SWAPQ35184372531200`, `SWAPH35184372531200`, `SWAPL35184372531200`, `SWAPB70368745062400`, `SWAPW70368745062400`, `SWAPD70368745062400`, `SWAPQ70368745062400`, `SWAPH70368745062400`, `SWAPL70368745062400`, `SWAPB140737490124800`, `SWAPW140737490124800`, `SWAPD140737490124800`, `SWAPQ140737490124800`, `SWAPH140737490124800`, `SWAPL140737490124800`, `SWAPB281474980249600`, `SWAPW281474980249600`, `SWAPD281474980249600`, `SWAPQ281474980249600`, `SWAPH281474980249600`, `SWAPL281474980249600`, `SWAPB56294996049600`, `SWAPW56294996049600`, `SWAPD56294996049600`, `SWAPQ56294996049600`, `SWAPH56294996049600`, `SWAPL56294996049600`, `SWAPB112589992099200`, `SWAPW112589992099200`, `SWAPD112589992099200`, `SWAPQ112589992099200`, `SWAPH112589992099200`, `SWAPL112589992099200`, `SWAPB225179984198400`, `SWAPW225179984198400`, `SWAPD225179984198400`, `SWAPQ225179984198400`, `SWAPH225179984198400`, `SWAPL225179984198400`, `SWAPB450359968396800`, `SWAPW450359968396800`, `SWAPD450359968396800`, `SWAPQ450359968396800`, `SWAPH450359968396800`, `SWAPL450359968396800`, `SWAPB900719936793600`, `SWAPW900719936793600`, `SWAPD900719936793600`, `SWAPQ900719936793600`, `SWAPH900719936793600`, `SWAPL900719936793600`, `SWAPB1801439873587200`, `SWAPW1801439873587200`, `SWAPD1801439873587200`, `SWAPQ1801439873587200`, `SWAPH1801439873587200`, `SWAPL1801439873587200`, `SWAPB3602879747174400`, `SWAPW3602879747174400`, `SWAPD3602879747174400`, `SWAPQ3602879747174400`, `SWAPH3602879747174400`, `SWAPL3602879747174400`, `SWAPB7205759494348800`, `SWAPW7205759494348800`, `SWAPD7205759494348800`, `SWAPQ7205759494348800`, `SWAPH7205759494348800`, `SWAPL7205759494348800`, `SWAPB1441151898869600`, `SWAPW1441151898869600`, `SWAPD1441151898869600`, `SWAPQ1441151898869600`, `SWAPH1441151898869600`, `SWAPL1441151898869600`, `SWAPB2882303797739200`, `SWAPW2882303797739200`, `SWAPD2882303797739200`, `SWAPQ2882303797739200`, `SWAPH2882303797739200`, `SWAPL2882303797739200`, `SWAPB5764607595478400`, `SWAPW5764607595478400`, `SWAPD5764607595478400`, `SWAPQ5764607595478400`, `SWAPH5764607595478400`, `SWAPL5764607595478400`, `SWAPB1152921519095200`, `SWAPW1152921519095200`, `SWAPD1152921519095200`, `SWAPQ1152921519095200`, `SWAPH1152921519095200`, `SWAPL1152921519095200`, `SWAPB2305843038190400`, `SWAPW2305843038190400`, `SWAPD2305843038190400`, `SWAPQ2305843038190400`, `SWAPH2305843038190400`, `SWAPL2305843038190400`, `SWAPB4611686076380800`, `SWAPW4611686076380800`, `SWAPD4611686076380800`, `SWAPQ4611686076380800`, `SWAPH4611686076380800`, `SWAPL4611686076380800`, `SWAPB9223372152761600`, `SWAPW9223372152761600`, `SWAPD9223372152761600`, `SWAPQ9223372152761600`, `SWAPH9223372152761600`, `SWAPL9223372152761600`, `SWAPB18446744305523200`, `SWAPW18446744305523200`, `SWAPD18446744305523200`, `SWAPQ18446744305523200`, `SWAPH18446744305523200`, `SWAPL18446744305523200`, `SWAPB36893488611046400`, `SWAPW36893488611046400`, `SWAPD36893488611046400`, `SWAPQ36893488611046400`, `SWAPH36893488611046400`, `SWAPL36893488611046400`, `SWAPB73786977222092800`, `SWAPW73786977222092800`, `SWAPD73786977222092800`, `SWAPQ73786977222092800`, `SWAPH73786977222092800`, `SWAPL73786977222092800`, `SWAPB147573954444185600`, `SWAPW147573954444185600`, `SWAPD147573954444185600`, `SWAPQ147573954444185600`, `SWAPH147573954444185600`, `SWAPL147573954444185600`, `SWAPB295147908888371200`, `SWAPW295147908888371200`, `SWAPD295147908888371200`, `SWAPQ295147908888371200`, `SWAPH295147908888371200`, `SWAPL295147908888371200`, `SWAPB590295817776742400`, `SWAPW590295817776742400`, `SWAPD590295817776742400`, `SWAPQ590295817776742400`, `SWAPH590295817776742400`, `SWAPL590295817776742400`, `SWAPB1180591635553484800`, `SWAPW1180591635553484800`, `SWAPD1180591635553484800`, `SWAPQ1180591635553484800`, `SWAPH1180591635553484800`, `SWAPL1180591635553484800`, `SWAPB2361183271106969600`, `SWAPW2361183271106969600`, `SWAPD2361183271106969600`, `SWAPQ2361183271106969600`, `SWAPH2361183271106969600`, `SWAPL2361183271106969600`, `SWAPB4722366542213939200`, `SWAPW4722366542213939200`, `SWAPD4722366542213939200`, `SWAPQ4722366542213939200`, `SWAPH4722366542213939200`, `SWAPL4722366542213939200`, `SWAPB9444733084427878400`, `SWAPW9444733084427878400`, `SWAPD9444733084427878400`, `SWAPQ9444733084427878400`, `SWAPH9444733084427878400`, `SWAPL9444733084427878400`, `SWAPB18889466168855756800`, `SWAPW18889466168855756800`, `SWAPD18889466168855756800`, `SWAPQ18889466168855756800`, `SWAPH18889466168855756800`, `SWAPL18889466168855756800`, `SWAPB37778932337711513600`, `SWAPW37778932337711513600`, `SWAPD37778932337711513600`, `SWAPQ37778932337711513600`, `SWAPH37778932337711513600`, `SWAPL37778932337711513600`, `SWAPB75557864675423027200`, `SWAPW75557864675423027200`, `SWAPD75557864675423027200`, `SWAPQ75557864675423027200`, `SWAPH75557864675423027200`, `SWAPL75557864675423027200`, `SWAPB151115729350860554400`, `SWAPW151115729350860554400`, `SWAPD151115729350860554400`, `SWAPQ151115729350860554400`, `SWAPH151115729350860554400`, `SWAPL151115729350860554400`, `SWAPB302231458701721108800`, `SWAPW302231458701721108800`, `SWAPD302231458701721108800`, `SWAPQ302231458701721108800`, `SWAPH302231458701721108800`, `SWAPL302231458701721108800`, `SWAPB604462917403442217600`, `SWAPW604462917403442217600`, `SWAPD604462917403442217600`, `SWAPQ604462917403442217600`, `SWAPH604462917403442217600`, `SWAPL604462917403442217600`, `SWAPB1208925834806884435200`, `SWAPW1208925834806884435200`, `SWAPD1208925834806884435200`, `SWAPQ1208925834806884435200`, `SWAPH1208925834806884435200`, `SWAPL1208925834806884435200`, `SWAPB2417851669613768870400`, `SWAPW2417851669613768870400`, `SWAPD2417851669613768870400`, `SWAPQ2417851669613768870400`, `SWAPH2417851669613768870400`, `SWAPL2417851669613768870400`, `SWAPB4835703339227537740800`, `SWAPW4835703339227537740800`, `SWAPD4835703339227537740800`, `SWAPQ4835703339227537740800`, `SWAPH4835703339227537740800`, `SWAPL4835703339227537740800`, `SWAPB9671406678455075481600`, `SWAPW9671406678455075481600`, `SWAPD9671406678455075481600`, `SWAPQ9671406678455075481600`, `SWAPH9671406678455075481600`, `SWAPL9671406678455075481600`, `SWAPB19342813356910150963200`, `SWAPW19342813356910150963200`, `SWAPD19342813356910150963200`, `SWAPQ19342813356910150963200`, `SWAPH19342813356910150963200`, `SWAPL19342813356910150963200`, `SWAPB38685626713820301926400`, `SWAPW38685626713820301926400`, `SWAPD38685626713820301926400`, `SWAPQ38685626713820301926400`, `SWAPH38685626713820301926400`, `SWAPL38685626713820301926400`, `SWAPB77371253427640603852800`, `SWAPW77`



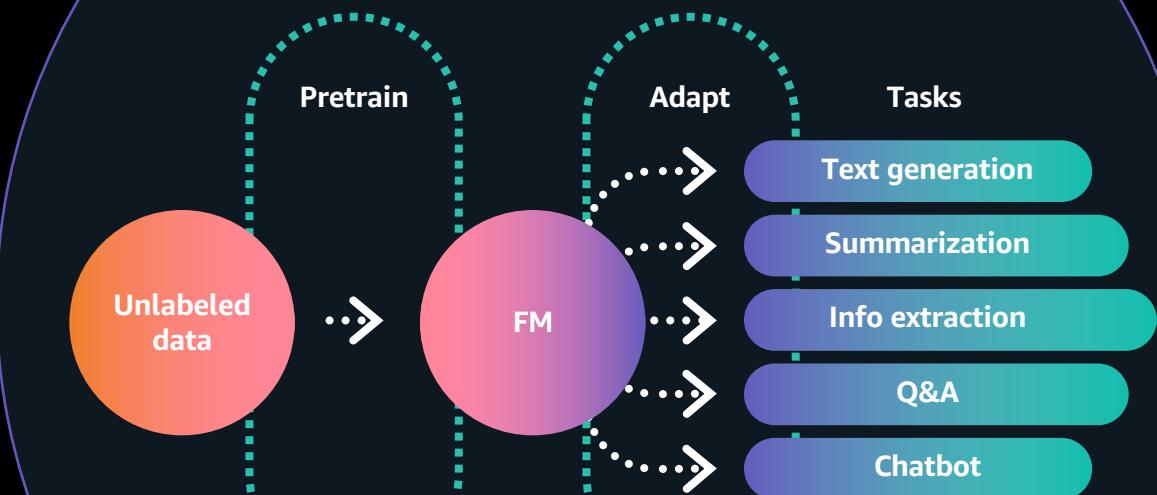




## TRADITIONAL ML MODELS



## FOUNDATION MODELS



# Essentials for building a **generative AI application**

**Purpose-built**  
ML infrastructure

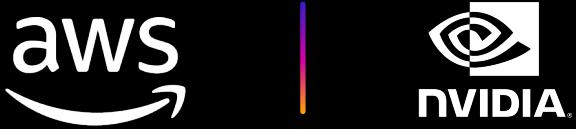
**Access to a variety**  
of foundation models

**Private environment**  
to leverage your data

**Easy-to-use tools** to build  
and deploy applications

# Essentials for building a **generative AI application**

Purpose-built  
ML infrastructure



2010

**CG1**

NVIDIA Tesla  
M2050 "Fermi"  
GPUs

**G2**

NVIDIA GRID  
GK104 "Kepler"  
GPUs

**P2**

NVIDIA  
K80  
GPUs

**G3**

NVIDIA  
Tesla M60  
GPUs

**P3**

NVIDIA V100  
Tensor Core  
GPUs

**G4**

NVIDIA T4  
Tensor Core  
GPUs

2020

**P4**

NVIDIA A100  
Tensor Core  
GPUs

**G5**

NVIDIA A10G  
Tensor Core  
GPUs

**G5g**

NVIDIA T4G  
Tensor Core  
GPUs

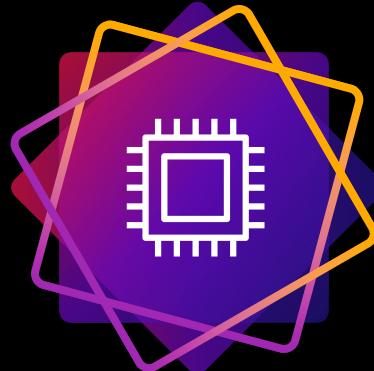
2023

**P5**

NVIDIA H100  
Tensor Core  
GPUs

# Innovating at the silicon level

AWS Trainium



AWS Inferentia

# Inferentia2

4x

HIGHER THROUGHPUT



10x

LOWER LATENCY

# Essentials for building a **generative AI application**

**Purpose-built**  
ML infrastructure

**Access to a variety**  
of foundation models

**Private environment**  
to leverage your data

**Easy-to-use tools** to build  
and deploy applications

# Essentials for building a **generative AI application**

**Access to a variety**  
of foundation models



No one model  
**will rule them all**

# Amazon **Bedrock**

Broad choice of models

**AI21labs**

**amazon**

**ANTHROPIC**

**cohere**

**Meta**

**stability.ai**

**Jurassic**

Contextual answers,  
summarization,  
paraphrasing

**Amazon Titan**

Text summarization,  
generation, Q&A, search

**Claude**

Summarization, complex  
reasoning, writing, coding

**Command + Embed**

Text generation, search,  
classification

**Llama 2**

Q&A and reading  
comprehension

**Stable Diffusion**

High-quality images and  
art

NEW

# Now Available in Amazon **Bedrock**

## ANTHROPIC

### Claude 2.1

Industry-leading 200K  
token context window

2x reduction in model  
hallucination rate

Reduces cost of prompts and  
completions on Bedrock by 25%

NEW

Now Available in Amazon **Bedrock**

∞Meta

**Llama 2 70B**

Fine-tuned for chat use cases

Supports large-scale tasks

NEW



## Amazon **Titan Text Lite**

Summarization, copywriting,  
and ideal for fine-tuning

GENERALLY AVAILABLE

NEW



## Amazon **Titan Text Express**

Open-ended text generation,  
conversational chat, and RAG support

GENERALLY AVAILABLE

NEW

# Amazon Titan Image Generator

Generate realistic, studio-quality images

AVAILABLE IN PREVIEW TODAY



Generate studio-quality images using natural language prompts



Customize images with proprietary data to match your brand style



Higher scores for text-to-image alignment compared to several leading models

# Invisible watermarks support responsible AI



AWS Services Search for services, features, blogs, docs, and more [Alt+S] N. Virginia MyRole/AWSUser @ 0123-4567-8901

Amazon Bedrock > Image Playground

## Image Playground

**Titan Image Generator** Change

### IMAGE GENERATION

Show me an image of an iguana

Configurations

Mode: Generate Edit

Negative prompt: Add system prompt

Reference image: (Placeholder)

Response image: Select style: Landscape Orientation: Landscape Portrait

Size: 512 x 512

Number of images: 3

Run Advanced configurations Info

CloudShell Feedback Language © 2023, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences



AWS Services Search for services, features, blogs, docs, and more [Alt+S] N. Virginia MyRole/AWSUser @ 0123-4567-8901

Amazon Bedrock > Image Playground

## Image Playground

**Titan Image Generator** Change

**OUTPAINTING**

Image of an iguana in a rainforest

Configurations X

Mode  Generate  Edit

Negative prompt  Add system prompt

Reference image  Info X

Response image  Info

Style Select style

Orientation  Landscape  Portrait

Size 512 x 512

Number of images 3

Run Advanced configurations Info

CloudShell Feedback Language © 2023, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences



AWS Services Search for services, features, blogs, docs, and more [Alt+S] N. Virginia MyRole/AWSUser @ 0123-4567-8901

Amazon Bedrock > Image Playground

## Image Playground

**a** Titan Image Generator Change



Configurations X

Mode  Generate  Edit

Negative prompt  Add system prompt

Reference image  Info

Response image  Info

Style  Select style

Orientation  Landscape  Portrait

Size  512 x 512

Number of images  3

Run Advanced configurations Info

CloudShell Feedback Language © 2023, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences



AWS Services Search for services, features, blogs, docs, and more [Alt+5] N. Virginia MyRole/AWSUser @ 0123-4567-8901

Amazon Bedrock > Image Playground

### Image Playground

**a** Titan Image Generator Change

**IMAGE VARIATION**

orange iguana facing right in a rainforest

Configurations

Mode

Generate Edit

Negative prompt Info

Add system prompt

Response image Info

Select style

Orientation

Landscape  Portrait

Size

512 x 512

Number of images

3

Run Advanced configurations Info

CloudShell Feedback Language

© 2023, Amazon Web Services, Inc. or its affiliates. Privacy Terms Cookie preferences

A screenshot of the Amazon Bedrock Image Playground interface. On the left, there's a preview area showing a smaller version of the generated image. A pink oval highlights the text input field containing the prompt "orange iguana facing right in a rainforest". To the right of the preview is a larger, detailed image of a vibrant orange iguana with a white face and blue stripes, resting on a wooden log in a lush green background. The interface includes various configuration options like mode, negative prompt, response image, and advanced settings.

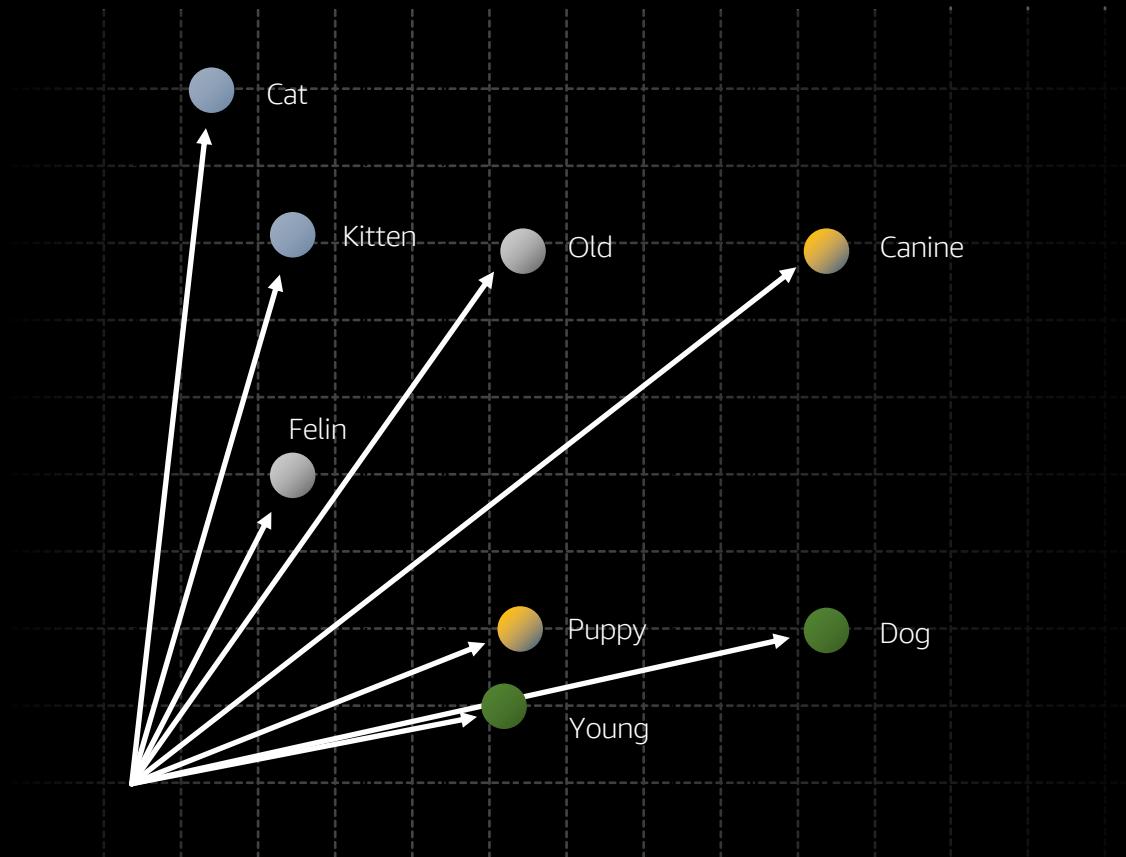
Vector embeddings are **numerical representations of text**

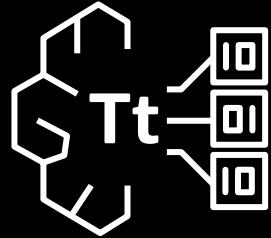
#### HUMAN TEXT

You may return or exchange your order within 30 days of when the order was shipped. Full refunds will be processed back to the original payment method.

After 30 days, we offer a store credit only.  
Customers will be responsible for the cost to ship returns or exchanges after 30 days.

Vectors are critical for  
**customizing generative  
AI applications**





Amazon Titan  
**Text Embeddings**

Translates text into numerical representations

Supports semantic search, text retrieval, and clustering

Supports a context length of up to 8k tokens

Works with 25+ languages

bright color golf shoes

## WITHOUT VECTOR EMBEDDINGS

### Result 1

- 1 Not applicable  text: Genuine FOOTJOY Golf Shoes Bag Zipped Sports Bag Shoe Caase -Black Color. Genuine FOOT JOY Golf Shoes Bag & Sports Bag Descriptions: Product Name: FOOTJOY Golf Shoes Bag Material: PU Color: Black Size: 24 (W) x 35 (L) x 11 (H) (cm) Side mesh type.
- 2 Not applicable  text: NIKE Genuine Shoe Bag Zipped Shoe Bag Golf Shoe Case Sports Gym Pouch. Genuine NIKE Shoes Bag Information: Model: Nike golf shoe bag (Nike Golf Shoes Pouch) Color: Black Size: 36 x 23 x 15 (cm) Material: Polyester
- 3 Not applicable  text: 1 Pairs Shoes Shapers Stretcher for Golf Shoes This pair of shoes stretchers can keep your golf shoes in style! It is made of high quality material, solid and durable in use. It is mainly used inside the shoe to preserve its shape, stop it from
- 4 Not applicable  text: Nike Control Junior Golf Shoes (3 Y, White/Metallic Silver/Bright Crimson). <p>#1.5 Swoosh design trademark embroidery at left front chest and Nike Athletic Department logo clip label at front lower right hem. Dark grey heather jersey at
- 5 Not applicable  text: 100Pcs 70mm Four Corner Mixed Color Plastic Golf Tees. Made of plastic, light and strong;Packed in reseal able bag for ease of use;Innovative design, four corners to support;Flat end even can clean up the mud on shoes or golf;Material: ABS

## WITH VECTOR EMBEDDINGS

### Result 2

- 1 Not applicable  text: FootJoy Ladies Summer Series Golf Shoes 98847 White/Tan Womens Closeout. The FootJoy Womens Superlite Spikeless golf shoes are designed to provide versatility and style on and off the course for all lady golfers. This new spikeless sole design
- 2 Not applicable  text: FootJoy FTJ54252-13 Wide Contour Casual Mens Golf Shoes Dark Brown - 13 Wide. Pro Am Golf has specialized in the sale of golfing products for over 35 years Our products are designed to make sure that it helps you to improve the game Utmost care is
- 3 Not applicable  text: FootJoy Superlites Athletic Spikeless Golf Shoes 2016 CLOSEOUT White/Grey/Black Medium 8.5. Pro Am Golf has specialized in the sale of golfing products for over 35 years Our products are designed to make sure that it helps you to improve the game
- 4 Not applicable  text: FootJoy Superlites Athletic Spikeless Golf Shoes 2016 CLOSEOUT Dark Blue/Lime Medium 10.5. Pro AM Golf has specialized in the sale of golfing products for over 35 years Our products are designed to make sure that it helps you to improve the game
- 5 Not applicable  text: FootJoy Superlites Athletic Spikeless Golf Shoes 2016 CLOSEOUT White/Grey/Black Medium 11.5. Pro AM Golf has specialized in the sale of golfing products for over 35 years Our products are designed to make sure that it helps you to improve the game

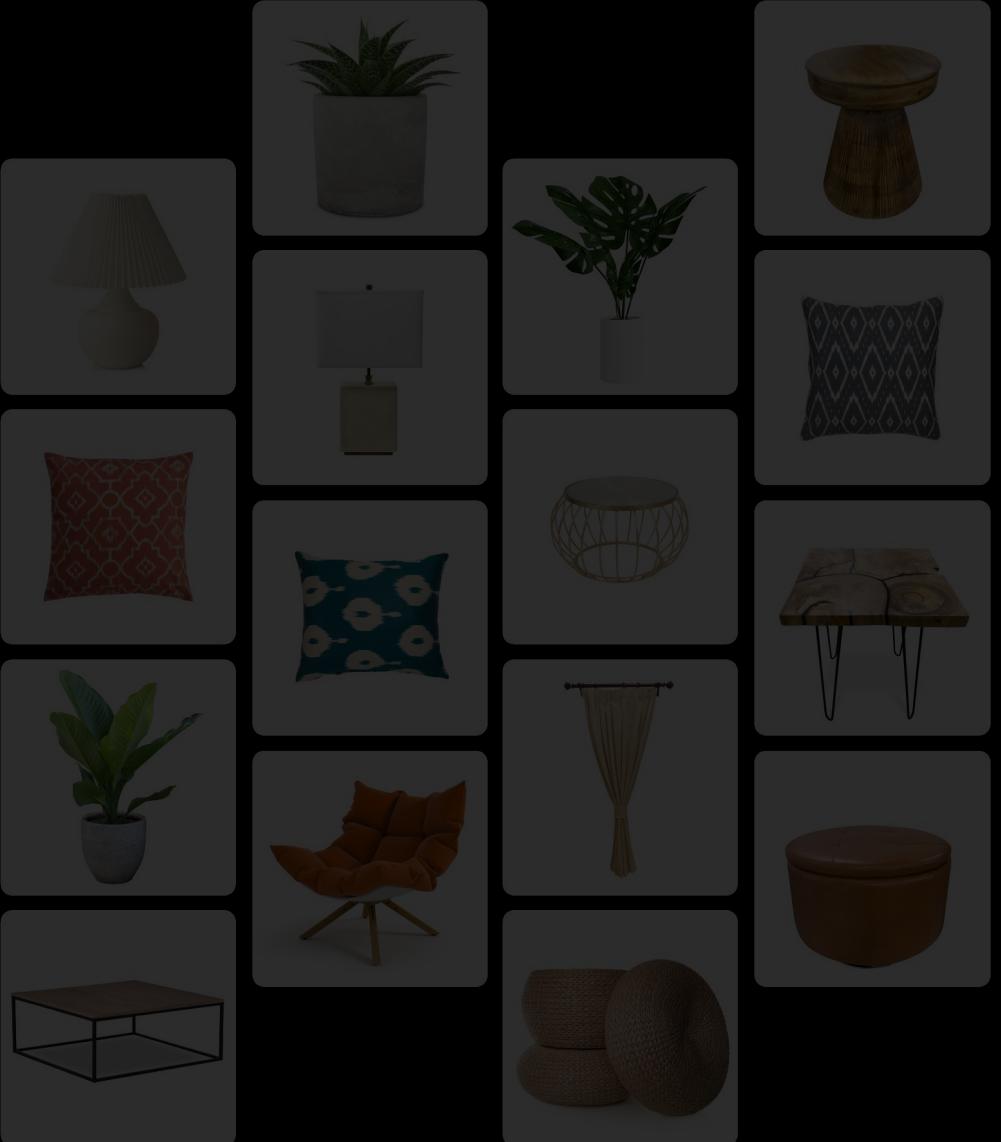
## IMAGE SEARCH



Upload your image



Show me what works well with my sofa



NEW

# Amazon Titan **Multimodal Embeddings**

Search, recommendation, and personalization

GENERALLY AVAILABLE



Accepts text, image, or a combination of text-image to generate embeddings

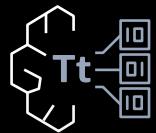


Adapts to unique and proprietary data



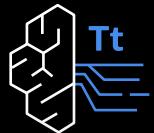
Built-in mitigation to help reduce biased search results

# Amazon Titan Foundation Models



## TITAN TEXT EMBEDDINGS

Translates text into numerical representations



## TITAN TEXT LITE

Summarization, copywriting, fine-tuning



## TITAN TEXT EXPRESS

Open-ended text generation, conversational chat, RAG support



## TITAN MULTIMODAL EMBEDDINGS

Search, recommendation, personalization



## TITAN IMAGE GENERATOR

Generate realistic, studio-quality images

# Amazon **Bedrock**

Broad choice of models

AI21labs

amazon

ANTHROPIC

cohere

Meta

stability.ai



Jurassic-2 Ultra

Jurassic-2 Mid

Titan Text Embeddings

Titan Multimodal Embeddings

Titan Text Lite

Titan Text Express

Titan Image Generator

Claude 2

Claude 2.1

Claude Instant

Command + Embed

Cohere Command Light

Cohere Embed English

Cohere Embed Multilingual

Llama 2

Llama 2 13B

Llama 2 70B

Stable Diffusion XL1.0

More than **10,000 customers** are using Amazon Bedrock



# Essentials for building a **generative AI application**

**Purpose-built**  
ML infrastructure

**Access to a variety**  
of foundation models

**Private environment**  
to leverage your data

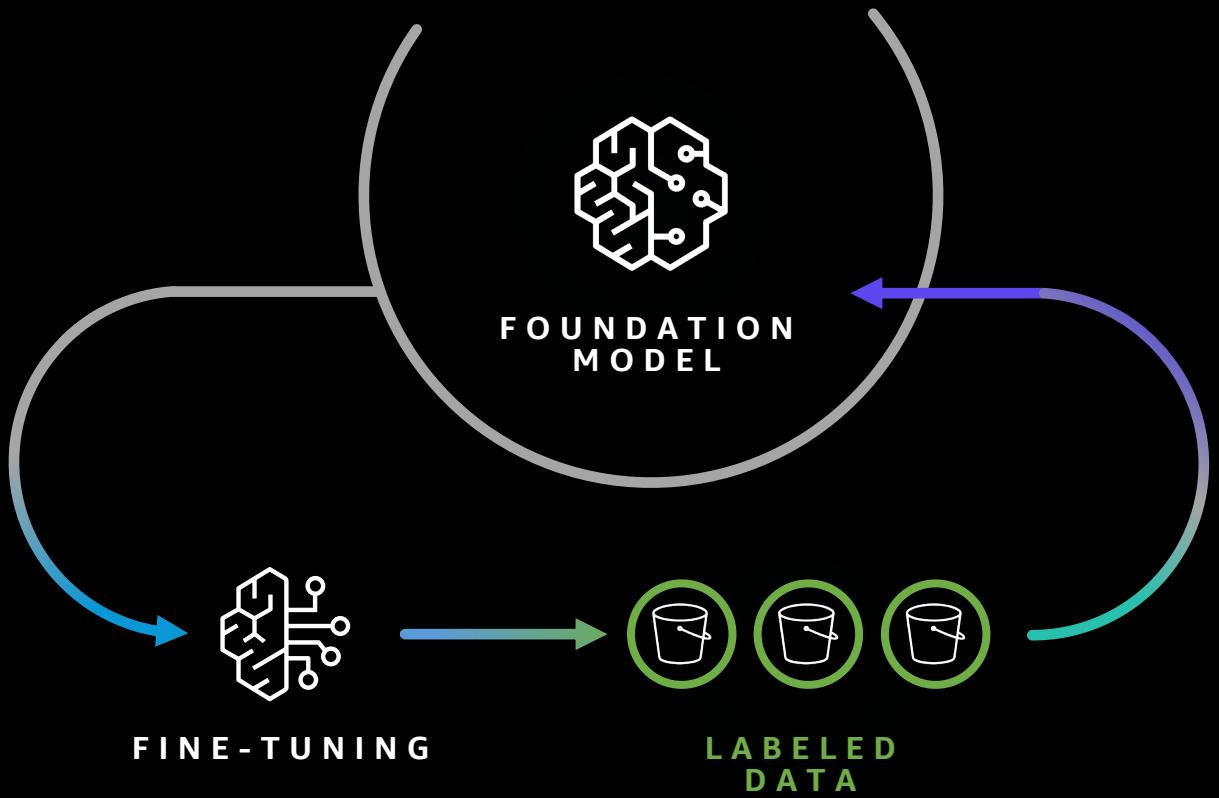
**Easy-to-use tools** to build  
and deploy applications

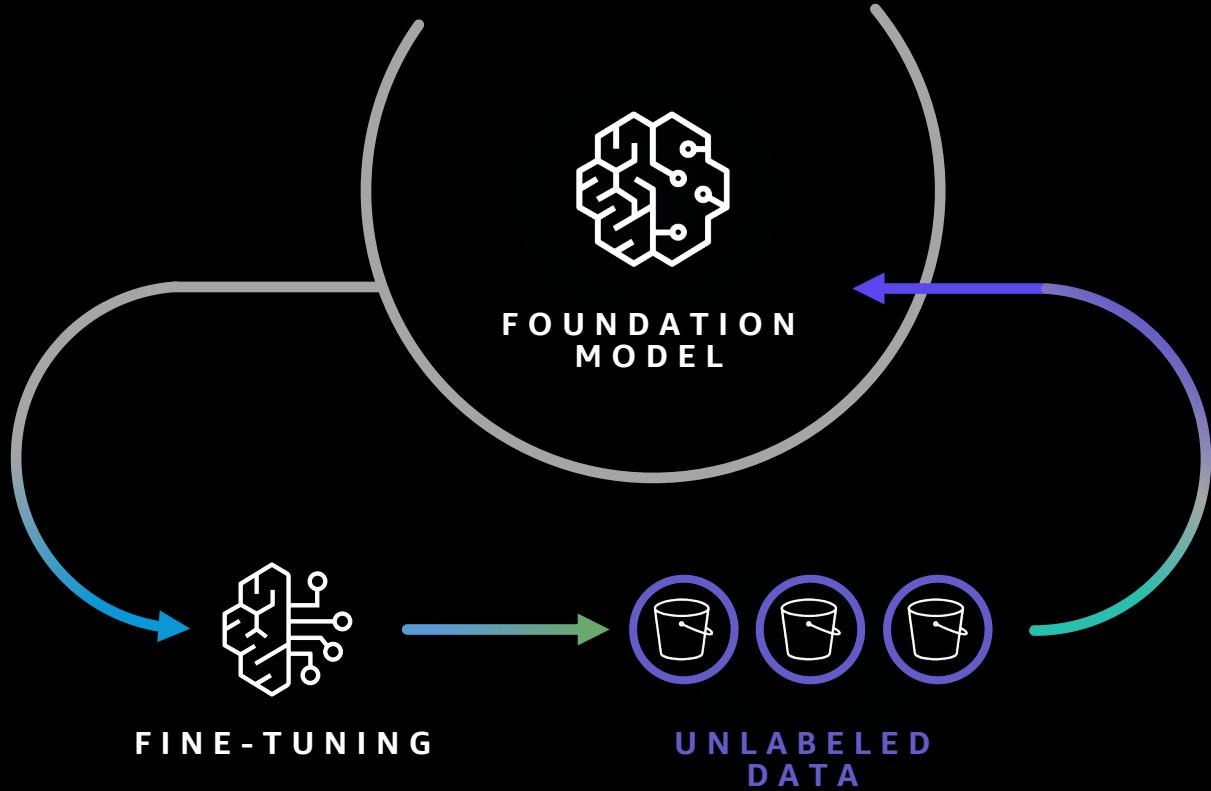
# Essentials for building a **generative AI application**

**Private environment**  
to leverage your data

**Data is the differentiator**  
for generative AI applications

Adapt models for  
your use case with  
**fine tuning**





Update your models  
through **continued**  
**pre-training**

# Fine tune additional models in Amazon Bedrock

COMING SOON

ANTHROPIC

cohere

Meta

amazon

# Essentials for building a **generative AI application**

**Easy-to-use tools**

to build and deploy  
applications

NEW

# Model evaluation on Amazon Bedrock

Evaluate, compare, and select the best foundation model for your use case

PREVIEW



COST

QUALITY



TRADE OFF



LATENCY

Getting started

**Overview**

Examples

Providers

Foundation models

Base models

Custom models

Playgrounds

Text

Chat

Image

Orchestration

Knowledge base

Agents

Assessment &amp; deployment

Model Evaluation [Preview](#)

Provisioned Throughput

Model access [12 new](#)

Settings

# Overview

[Explore & Learn](#)[Build & Test](#)

## Foundation models

Amazon Bedrock supports foundation models from industry-leading providers. Choose the model that is best suited to achieving your unique goals.



Jurassic-2 series  
By AI21 Labs



Titan  
By Amazon



Claude  
By Anthropic



Command  
By Cohere

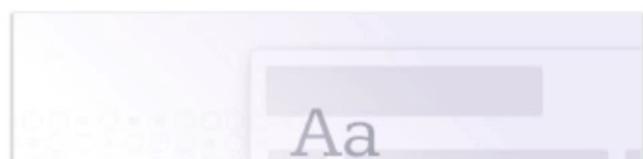


Llama 2  
By Meta



Stable Diffusion  
By Stability AI

## Playgrounds



Is Finetuning the only option to make  
Large Language Models work with my Data ?

Implementing RAG can  
**be resource intensive**  
**and time consuming**

- **CONVERT**  
data into embeddings
- **STORE**  
integrations with vector databases
- **RETRIEVE**  
find relevant results of from vector database  
based on user's query
- **AUGMENT**  
add above results along with user's query to  
augment the prompt
- **GENERATE**  
instruct LLMs to generate responses based on  
contextual data

# Knowledge Bases for **Amazon Bedrock**

Fully-managed native support for  
retrieval augmented generation



Automatically converts text documents into embeddings



Stores embeddings in your vector database



Retrieves embeddings and augments prompts



1



2



3

## AnyHotel FAQ

### 1. What are the check-in and check-out times?

Check-in time is 3pm and check-out time is 11am at all AnyHotel locations. Early check-in and late check-out may be available upon request and availability. Please inquire at the front desk upon arrival.

### 2. Do you offer airport shuttles?

Airport shuttles are available at the following AnyHotel locations:

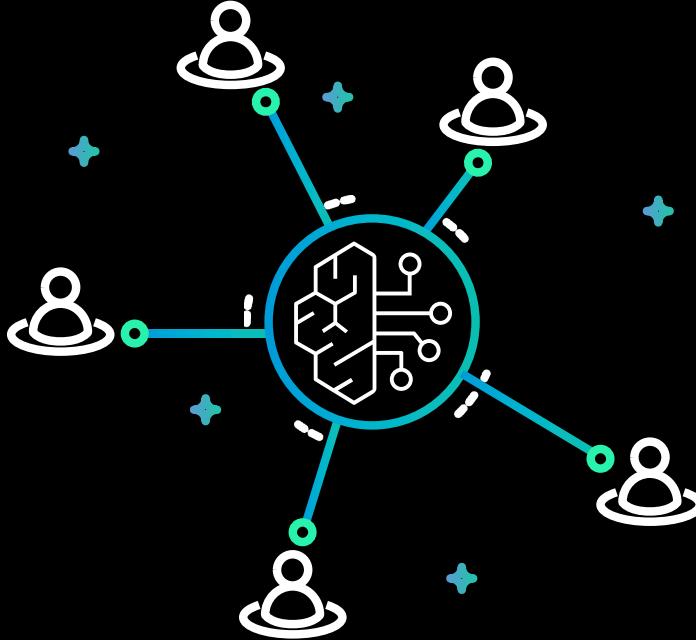
- AnyHotel Dallas: Complimentary airport shuttle available to and from Dallas/Fort Worth International Airport. Shuttle runs every 30 minutes from 5am-11pm.

- AnyHotel Chicago: Complimentary airport shuttle available to and from O'Hare International Airport and Chicago Midway Airport. Shuttle runs every hour from 5am-11pm.

- AnyHotel San Francisco: Complimentary airport shuttle available to and from San Francisco International Airport. Shuttle runs every 30 minutes from 5am-11pm.

- AnyHotel New York: Complimentary shuttle available to and from LaGuardia Airport and JFK Airport. Shuttle runs every hour from 5am-11pm.

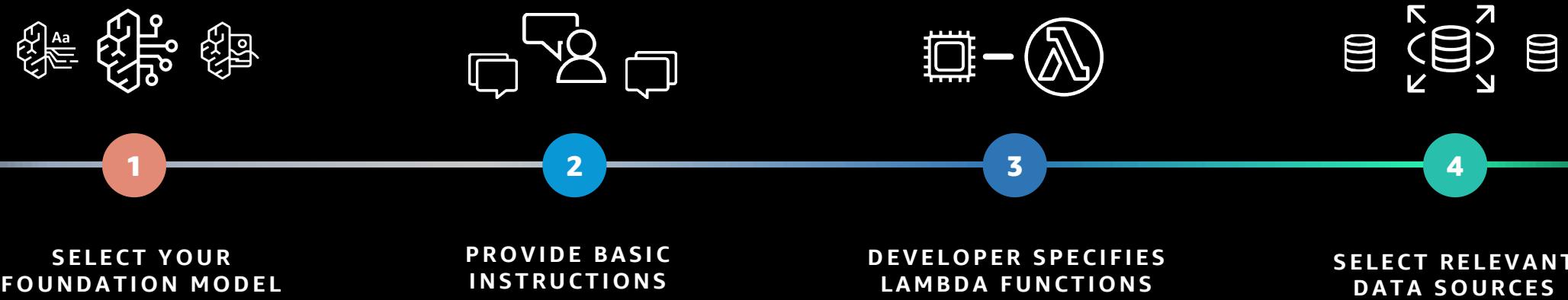
What if I have more dynamic data or  
complex tasks for my Application needs ?



# Agents for **Amazon Bedrock**

Enable generative AI applications to complete tasks in just a few clicks

# Enabling foundation models to execute tasks





**KreditBee**



KreditBee

Shellcode

+

# Essentials for building a **generative AI application**

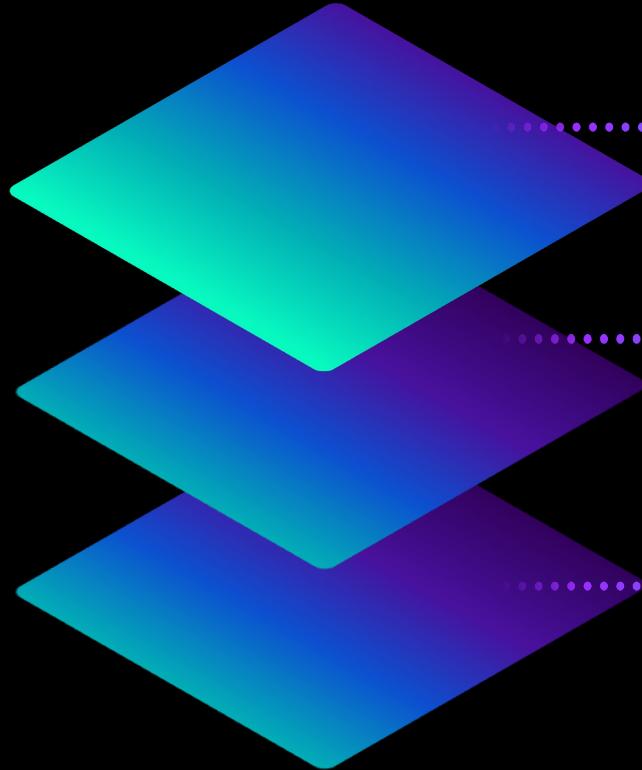
**Purpose-built**  
ML infrastructure

**Access to a variety**  
of foundation models

**Private environment**  
to leverage your data

**Easy-to-use tools** to build  
and deploy applications

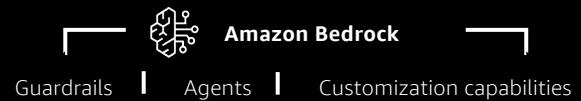
# Generative AI Stack



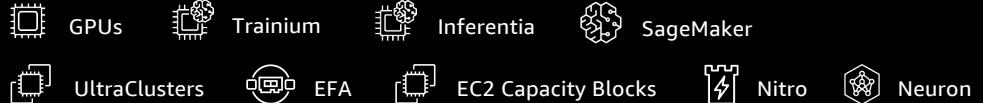
## APPLICATIONS THAT LEVERAGE FMs



## TOOLS TO BUILD WITH FMS AND LLMS



## INFRASTRUCTURE FOR FM TRAINING & INFERENCE



# Amazon Q accelerates productivity



AMAZON Q is  
YOUR AWS  
EXPERT



AMAZON Q  
IN AMAZON  
QUICKSIGHT



AMAZON Q is  
YOUR BUSINESS  
EXPERT



AMAZON Q  
IN AMAZON  
CONNECT



Everything you need to accelerate your  
**generative AI journey**



# Go Create with AWS

