



Event ends in 2 days 15 hours 13 minutes.



[Event dashboard](#) > Image and MultiModal applications

Image and MultiModal applications

Overview

Image generation can be a tedious task for artists, designers and content creators who illustrate their thoughts with the help of images. With the help of Foundation Models (FMs) this tedious task can be streamlined to just a single line of text that expresses the thoughts of the artist, FMs can be used for creating realistic and artistic images of various subjects, environments, and scenes from language prompts.

Image indexing and searching is another tedious enterprise task. With the help of FMs, enterprise can build multimodal image indexing, searching and recommendation applications quickly.

In this lab, we will explore how to use FMs available in Amazon Bedrock to generate images as well as modify existing images, and how to use FMs to do multimodal image indexing and searching.

Relevance

Amazon Titan Image Generator G1 is an image generation model. It generates images from text, and allows users to upload and edit an existing image. Users can edit an image with a text prompt (without a mask) or parts of an image with an image mask. You can extend the boundaries of an image with outpainting, and fill in an image with inpainting. It can also generate variations of an image based on an optional text prompt. Amazon Titan Image Generator G1 Generator includes watermarking on the output files.

Text-to-image (T2I) generation – Input a text prompt and generate a new



Inpainting – Uses an image and a segmentation mask as input (either from the user or estimated by the model) and reconstructs the region within the mask. Use inpainting to remove masked elements and replace them with background pixels.

Outpainting – Uses an image and a segmentation mask as input (either from the user or estimated by the model) and generates new pixels that seamlessly extend the region. Use precise outpainting to preserve the pixels of the masked image when extending the image to the boundaries. Use default outpainting to extend the pixels of the masked image to the image boundaries based on segmentation settings.

Image variation – Uses an image and an optional prompt as input. It generates a new image that preserves the content of the input

Audience

- Marketing teams which need to generate images
- Multiple business units needing data from various internal documents which have a mix of tables and images and charts.
- Research units which need access to multi-form data
- Senior leadership who need access to crystalized summary from multiple forms of data.

Challenges

Data is coming in multiple forms which include images, charts, tables and text. All LLM's are great at textual data and hence the struggle is to create data from the disparate sources which can be used by LLM's to generate meaningful responses to the query. Maintaining relevant answers as conversations grow in complexity and scope.

Sub-patterns

This lab explores the following patterns:

1. [Convert text and prompts into Images](#) - Bedrock provides access to variety of image generation models like StableDiffusion and Titan Image generation.
2. [Titan Image Generator-Text to Image](#) - Bedrock Titan image generator model provides ability to generate high quality images from the input text.



This model adds an invisible watermark to all generated images to reduce the spread of misinformation, assist with copyright protection, and track content usage

3. [Titan Multimodal Images](#) - Amazon Titan Multimodal Embedding Models can be used for enterprise tasks such as image search and similarity based recommendation from documents which include dis-parate information items like text, images, graphs.

[Previous](#)[Next](#)