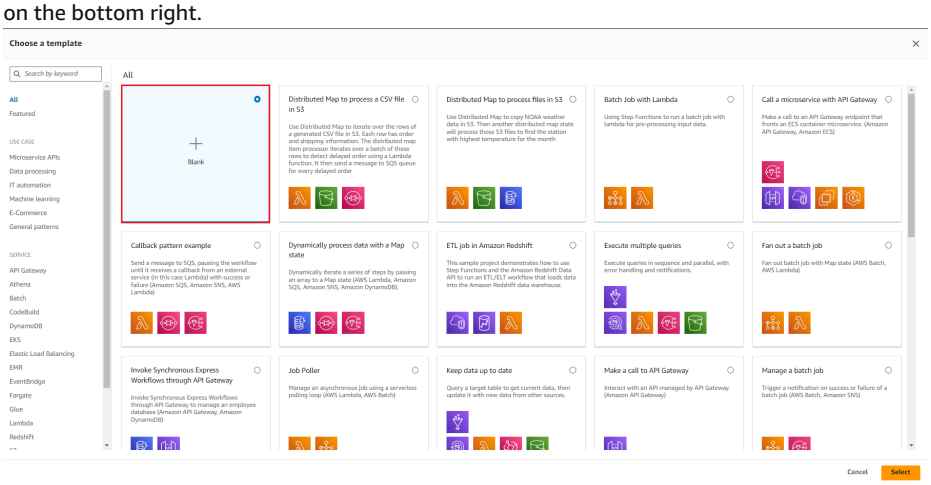
Invoking a model in Amazon Bedrock using AWS Step Functions

In this section, you will build a Step Function workflow to invoke an Anthropic Claude model in Amazon Bedrock and test in the Step Functions console. Then, you will run the same AWS Step Functions workflow to trigger the model from a user-facing AppSync API.

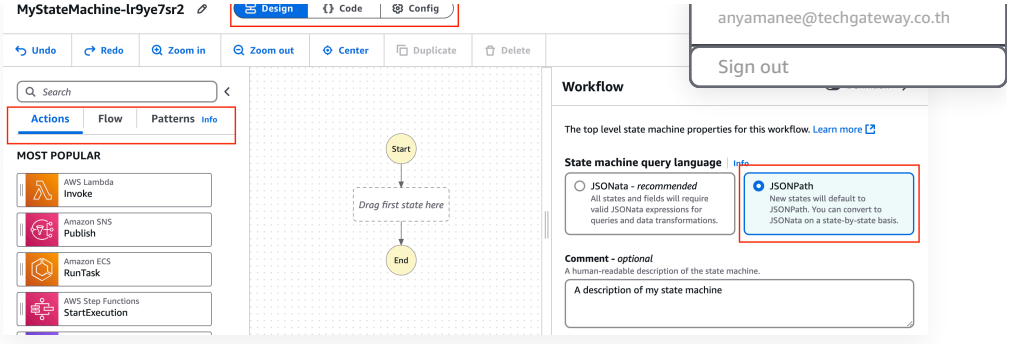
Create the Step Functions workflow

Console changes and new features!
Step Functions has launched a new feature to support JSONata and Variables on 11/22 along with console changes. While the new features simplify the creation of workflow, they are not integrated in to the workshop yet. Also you will see slight changes between the screenshots and the actual console.

1. Open the Step Functions console
2. Select **Create state machine** at the top right and choose the **Blank** template and choose **Select** on the bottom right.



3. You are now in the Step Functions visual designer studio.
4. Explore the visual designer studio. On the left, you can see tabs named `Actions`, `Flow`, and `Patterns`. You can search for actions using the search bar, drag and drop them in the graphical designer layout which is the middle part of the page. On the right of the page, you can see the options to configure the steps as you work on them.
5. In the top navigation bar, you will see tabs named `Design`, `Code`, and `Config`. You can edit the configurations of the workflow using those tabs



Add Bedrock integration

1. In the left window, type `Bedrock InvokeModel` in the search box

AWS account access

Open AWS console (us-west-2)

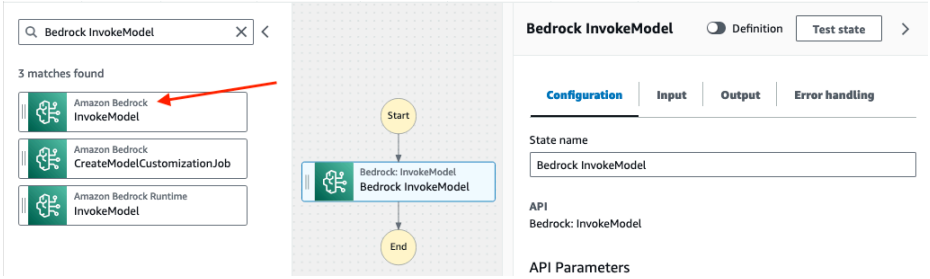
Get AWS CLI credentials

Exit event

2. Drag and drop the **InvokeModel** task under the `Amazon Bedrock` category to the designer. Notice how the search returned multiple results. *Amazon Bedrock Runtime InvokeModel* uses direct SDK/API integration, while the *Amazon Bedrock InvokeModel* uses an optimized integration. See the banner below for more information.

> **Do you know?**
> Step Functions integrates with AWS services in multiple ways. One common integration pattern is to invoke the AWS SDK API. Another integration pattern is the optimized call to the service with additional capabilities than offered in the SDK. For example, Bedrock optimized integration supports optionally retrieving and storing request/response data from S3. If you have use cases where you want to generate an image using LLM, you can use the native optimized integration to store the large image response from LLM in S3.

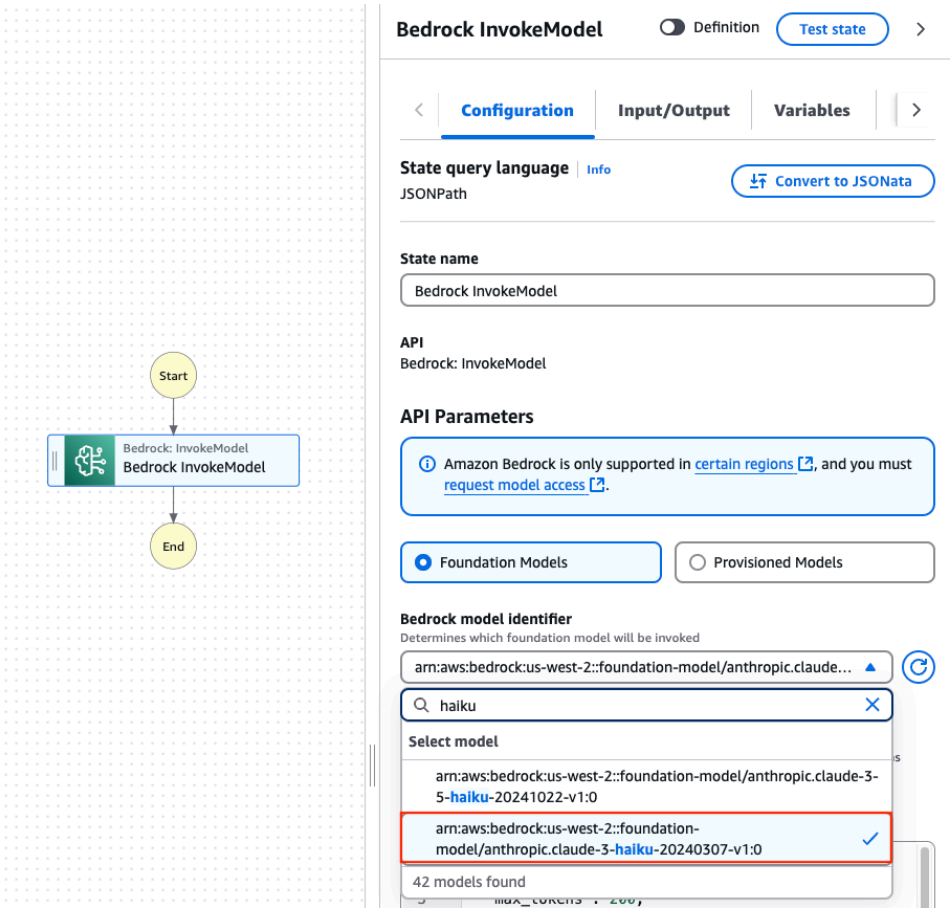
3. Select the **Amazon Bedrock InvokeModel** task



Configure Bedrock Integration

1. Locate the configuration section on the right side of the studio. Type **haiku** under `Bedrock model identifier` and select `anthropic.claude-3-haiku`.

> Select **claude-3-haiku**, **not claude-3-5-haiku**.



> **Important**
> As you are creating the workflow, you will see a red banner at the top saying **workflow not created**. You may ignore this warning as it will resolve once you save the workflow

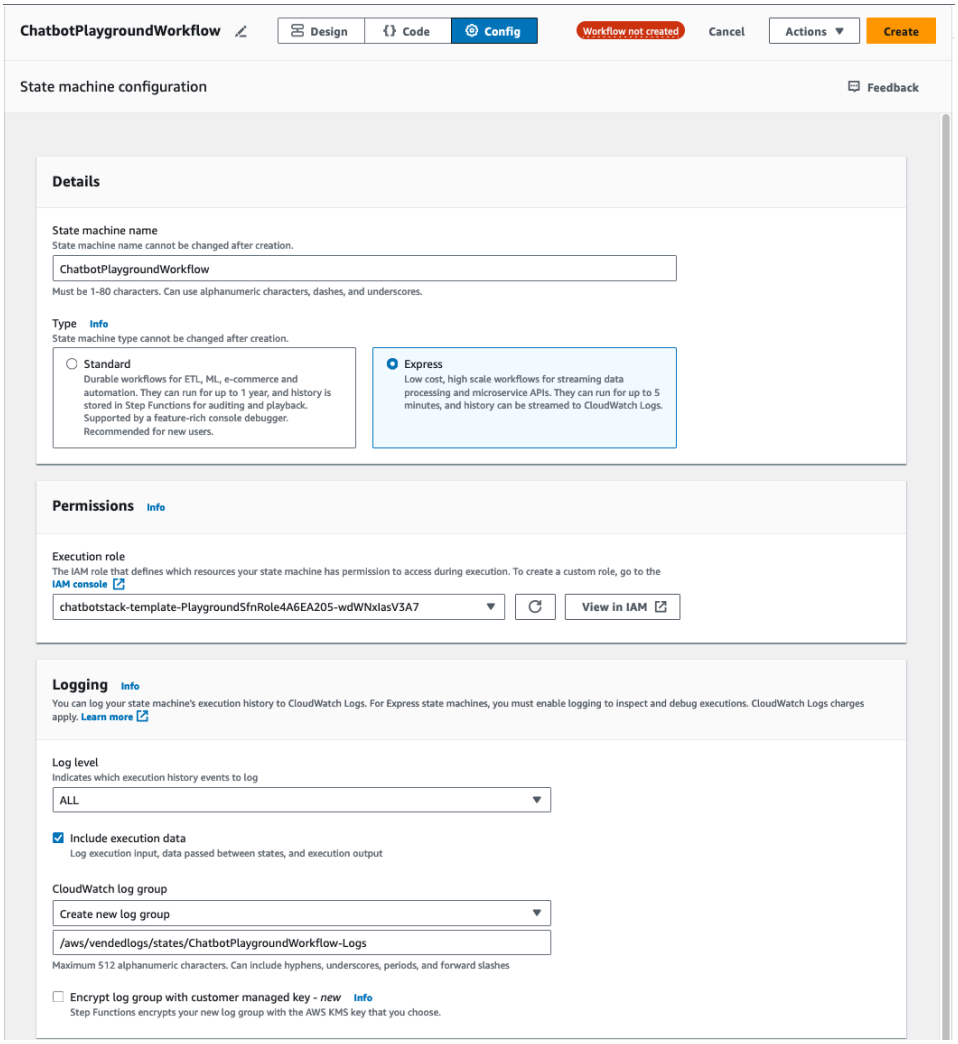
2. Move down to **Bedrock Model Parameters** section and replace the text in the text area with the json below, then click **Save**.

> **What is $ for?**
> `$` symbol in the request below indicates that the value is evaluated at run time. `text.$` expects a runtime variable by the name `prompt`.

```
{
  "anthropic_version": "bedrock-2023-05-31",
  "max_tokens": 200,
  "messages": [
    {
      "role": "user",
      "content": [
        {
          "type": "text",
          "text.$": "$.prompt"
        }
      ]
    }
  ]
}
```

Configure the State machine settings

1. Select the `Config` tab at the top of the State Machine visual studio
2. Name your State Machine `ChatbotPlaygroundWorkflow`
3. Select the `Type` as **Express**. You will be using this same state machine as a data source for a synchronous API call in a few steps. Express workflows are suitable for high burst and low latency use cases which can complete within 5 minutes, while Standard workflows can run for 365 days. Learn more about the difference here.
4. Select the IAM role starting with `chatbotstack-PlaygroundSfnRole` for your Step Function Execution Role. This role is pre-created with the policy to invoke the Bedrock model.
5. Review that your settings look as the screenshot below and select **Create**

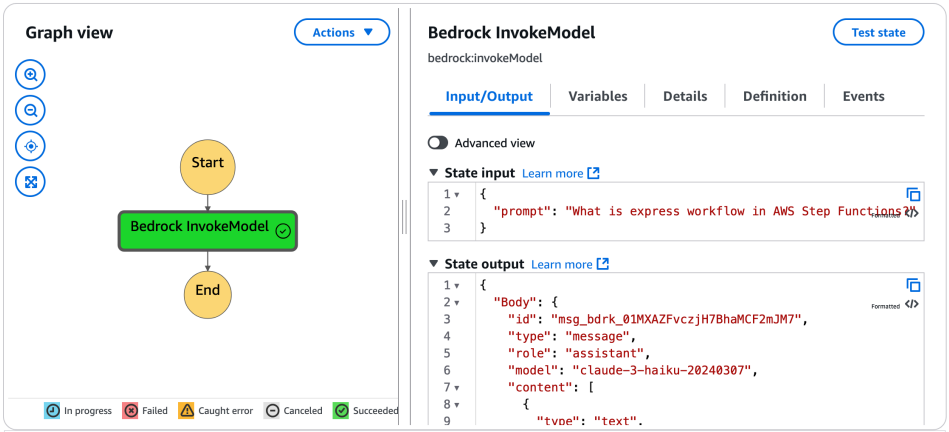


Run the workflow

1. Select `Execute` from the next page. This will open a new Start Execution tab in your browser
2. Insert the below JSON in the input field

```
{
  "prompt": "What is express workflow in AWS Step Functions?"
}
```

3. Select **Start execution**
4. Select the `Bedrock InvokeModel` task in the execution and explore the `Input/Output` tab on the right side.
5. If you do not see any input, turn on **Advanced view**



6. You may now close this browser tab to return to the State Machine design editor

Congratulations!
You have successfully invoked a foundation model in Amazon Bedrock using Step Functions.

Previous Next