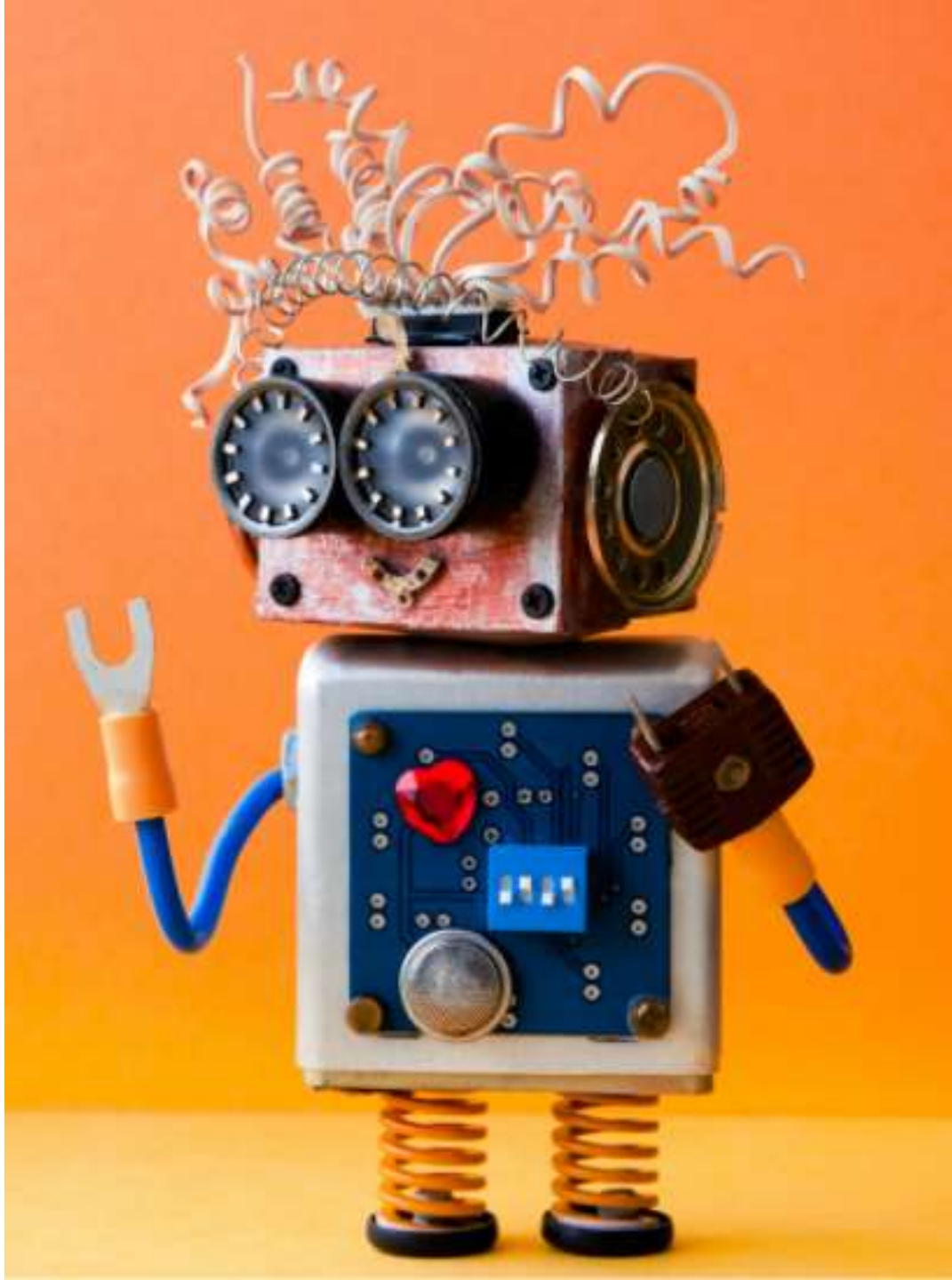
PROJECT OVERVIEW

EASY

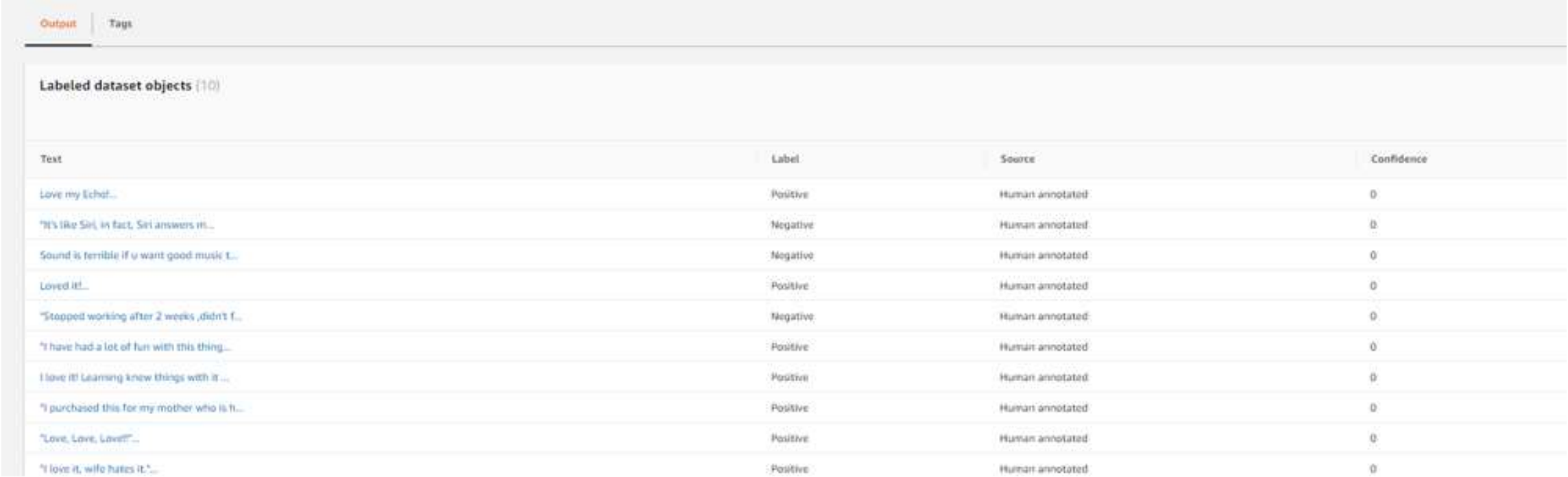
ADVANCED



PROJECT OVERVIEW

- In this project, we will learn how to (1) label images for object detection by drawing bounding boxes, (2) to label text data to perform sentiment analysis, and (3) perform pixel level semantic segmentation using Amazon SageMaker GroundTruth.
- These are the key learning outcomes:
    1. Learn how to define a bounding box labeling job using Amazon SageMaker Groundtruth
    2. Learn how to define a text labeling job using Amazon SageMaker Groundtruth
    3. Learn how to define semantic segmentation labeling job using Amazon SageMaker Groundtruth
    4. Perform text labeling job
    5. Draw bounding boxes and pixel-level semantics in Amazon SageMaker Groundtruth

PROJECT OVERVIEW

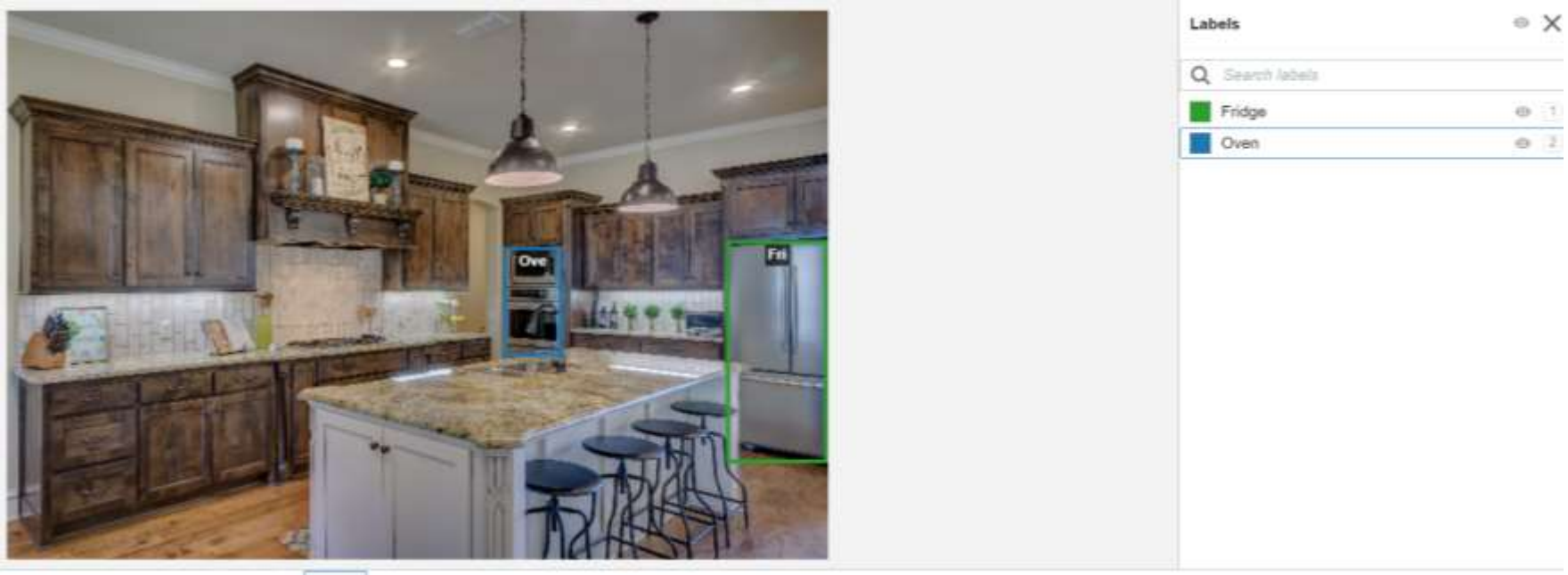


Output | Tags

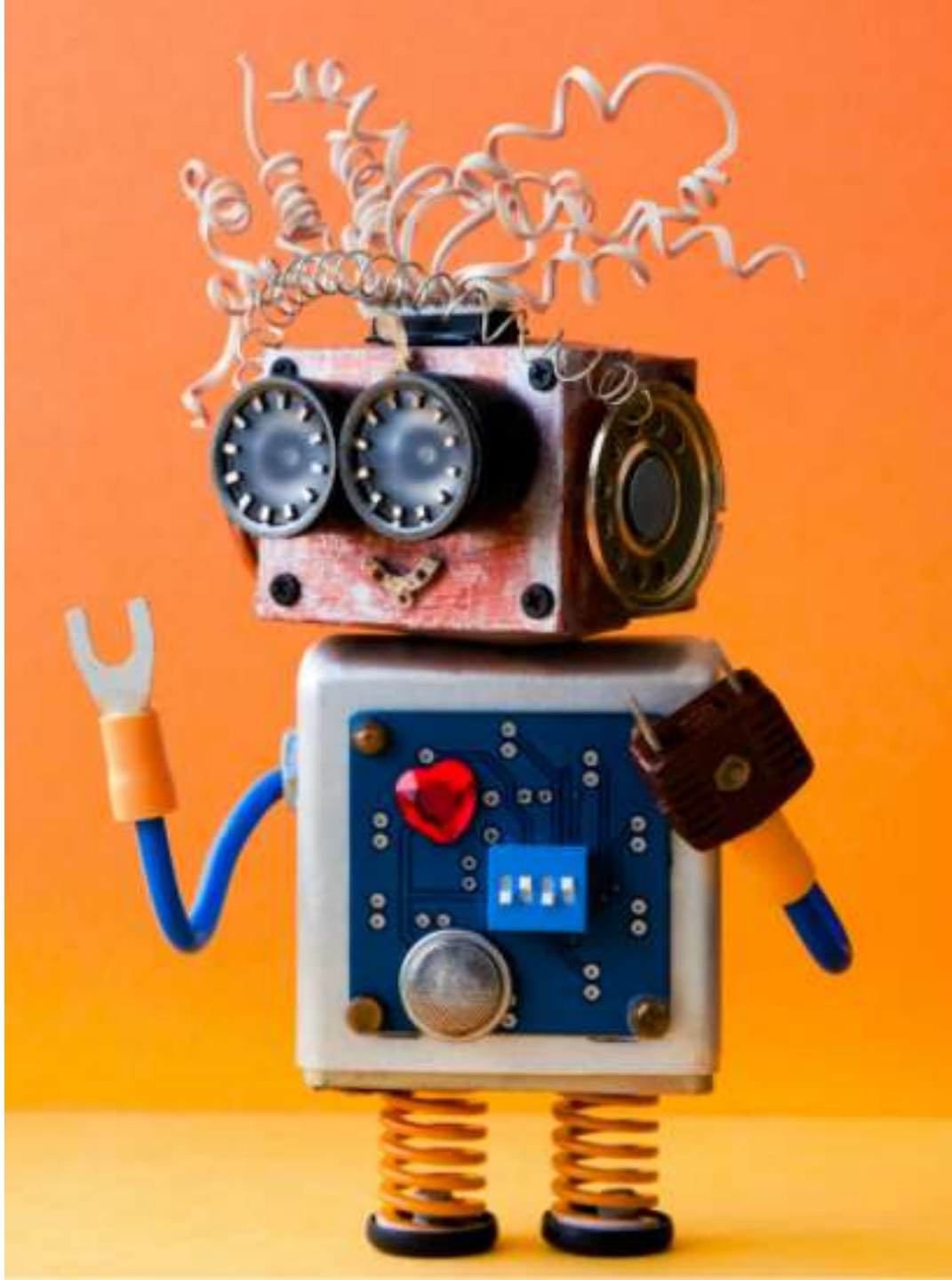
Labeled dataset objects (10)

Text	Label	Source	Confidence
Love my Echo!...	Positive	Human annotated	0
"It's like Siri, in fact, Siri answers m...	Negative	Human annotated	0
Sound is terrible if u want good music t...	Negative	Human annotated	0
Loved it!...	Positive	Human annotated	0
"Stopped working after 2 weeks ,didn't f...	Negative	Human annotated	0
"I have had a lot of fun with this thing...	Positive	Human annotated	0
I love it! Learning knew things with it ...	Positive	Human annotated	0
"I purchased this for my mother who is h...	Positive	Human annotated	0
"Love, Love, Love!!"...	Positive	Human annotated	0
"I love it, wife hates it."...	Positive	Human annotated	0

PROJECT OVERVIEW

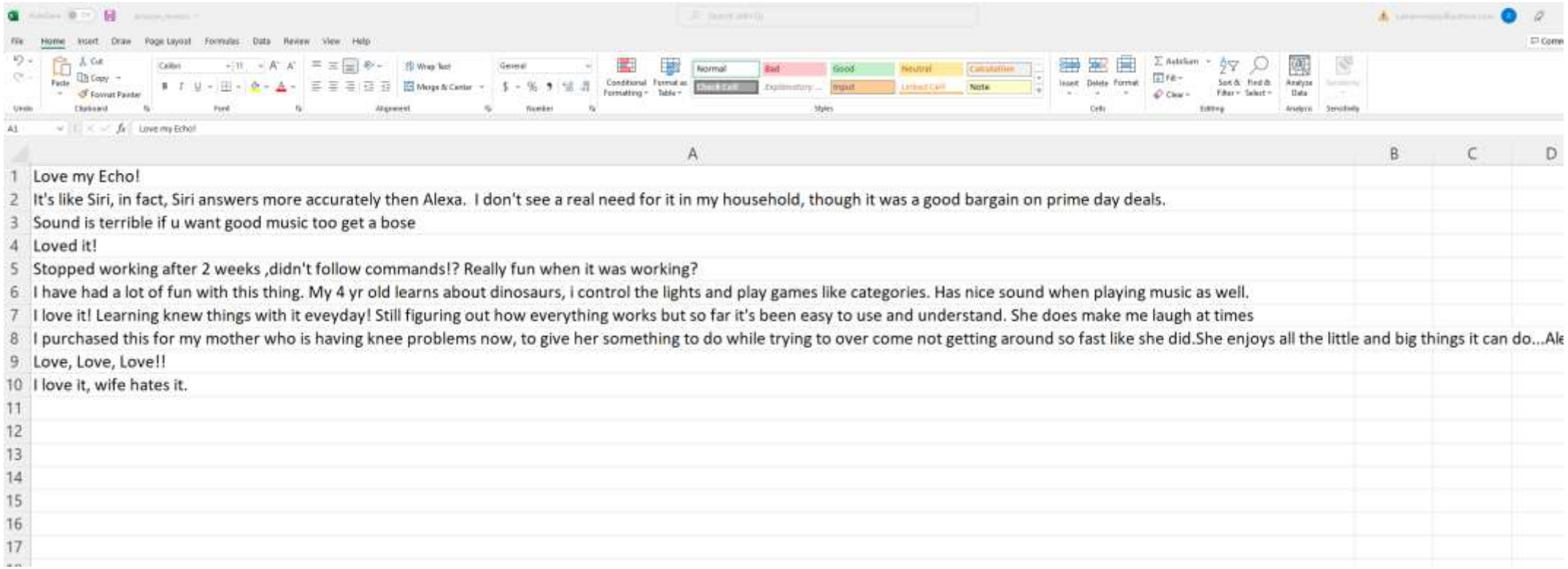


LABELING TEXT DATASET



LABELING TEXT DATASET

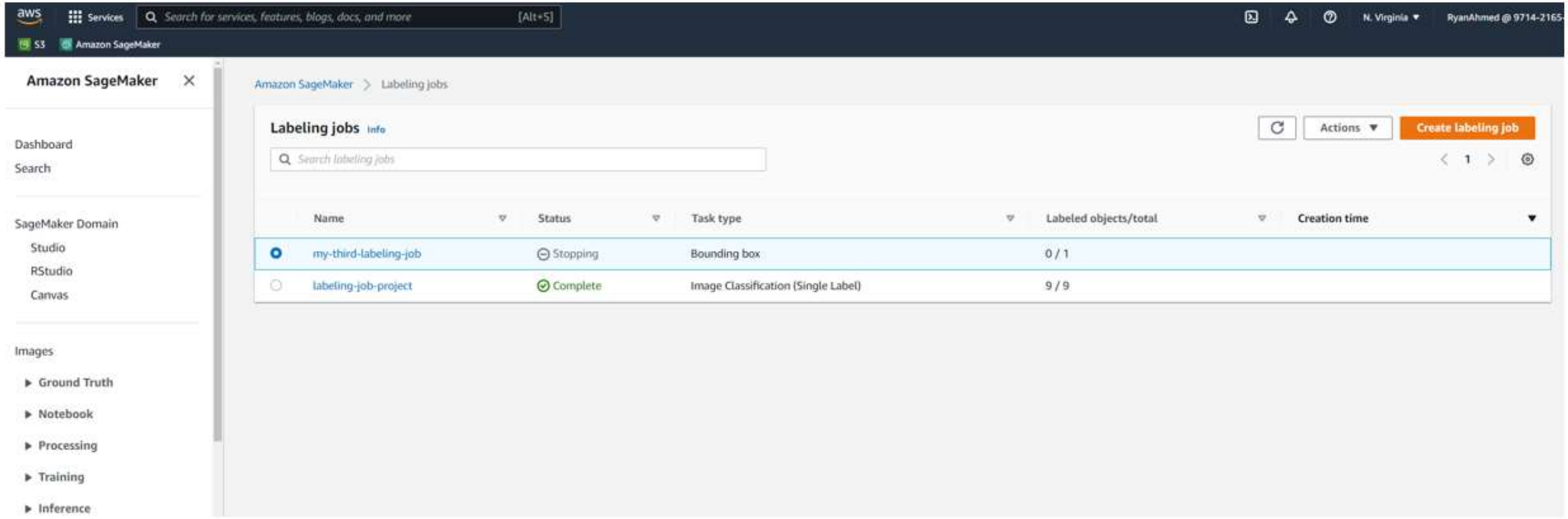
THE OBJECTIVE IS TO LABEL THE FOLLOWING 10 REVIEWS TO EITHER POSITIVE OR NEGATIVE



	A
1	Love my Echo!
2	It's like Siri, in fact, Siri answers more accurately then Alexa. I don't see a real need for it in my household, though it was a good bargain on prime day deals.
3	Sound is terrible if u want good music too get a bose
4	Loved it!
5	Stopped working after 2 weeks ,didn't follow commands!? Really fun when it was working?
6	I have had a lot of fun with this thing. My 4 yr old learns about dinosaurs, i control the lights and play games like categories. Has nice sound when playing music as well.
7	I love it! Learning knew things with it eveyday! Still figuring out how everything works but so far it's been easy to use and understand. She does make me laugh at times
8	I purchased this for my mother who is having knee problems now, to give her something to do while trying to over come not getting around so fast like she did.She enjoys all the little and big things it can do...Ale
9	Love, Love, Love!!
10	I love it, wife hates it.
11	
12	
13	
14	
15	
16	
17	

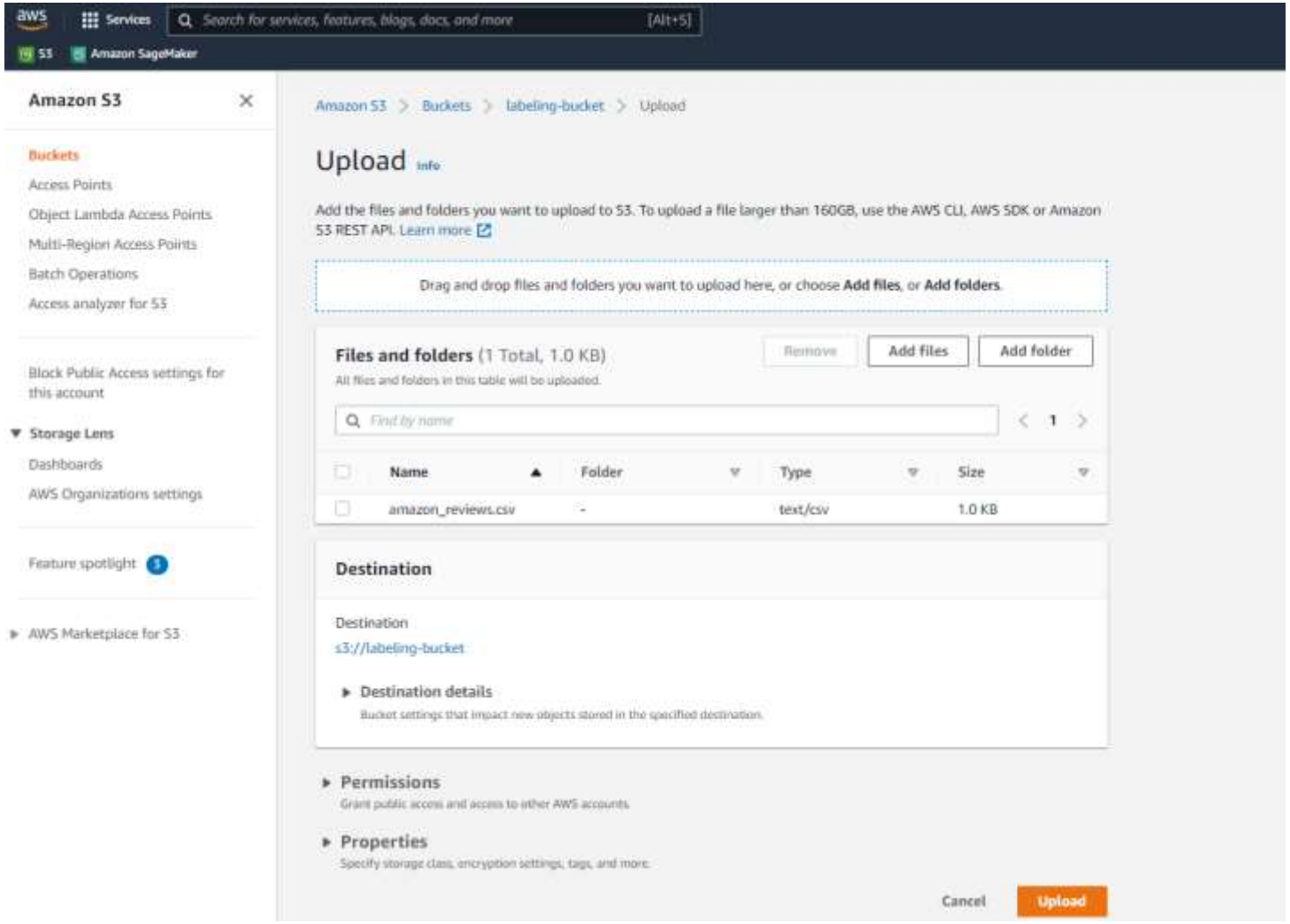
LABELING TEXT DATASET

NAVIGATE TO SAGEMAKER GROUNDTRUTH AND CLICK ON CREATE LABELING JOB



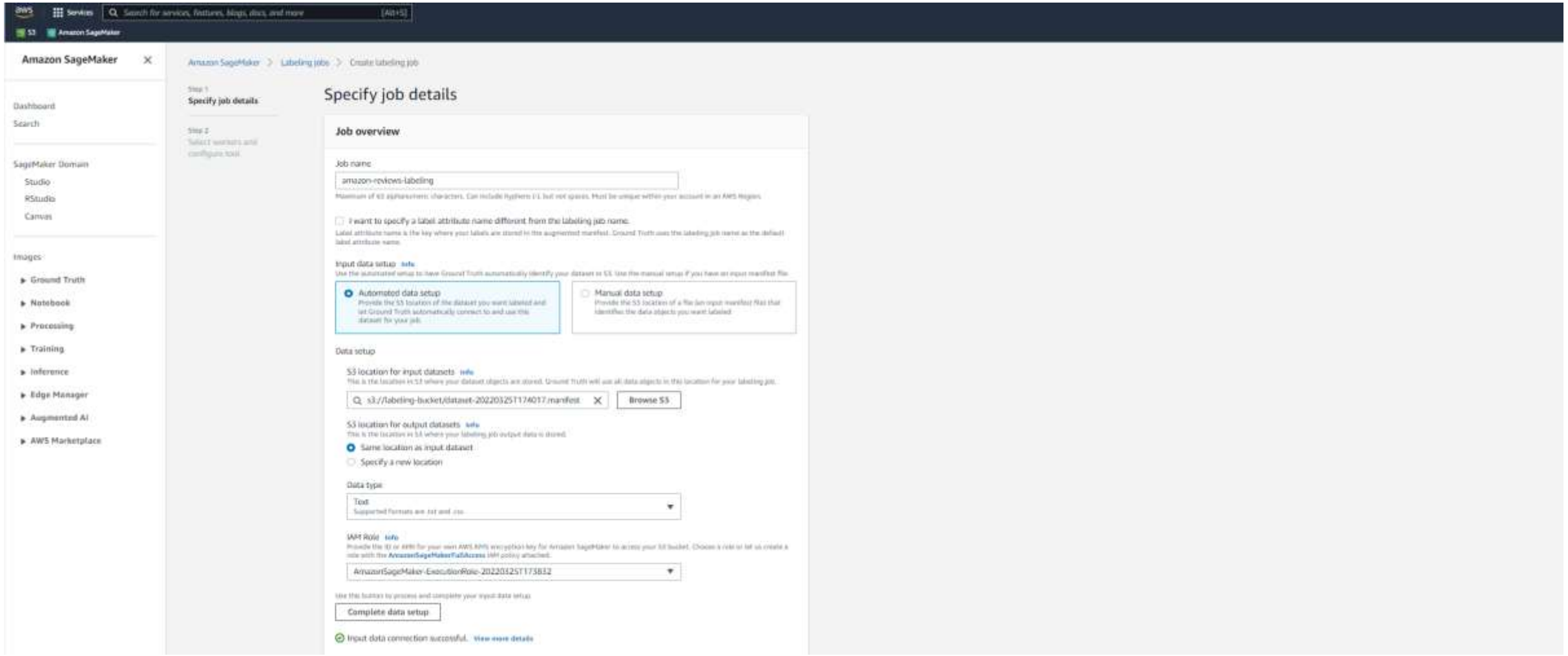
LABELING TEXT DATASET

UPLOAD AMAZON REVIEWS DATASET TO S3



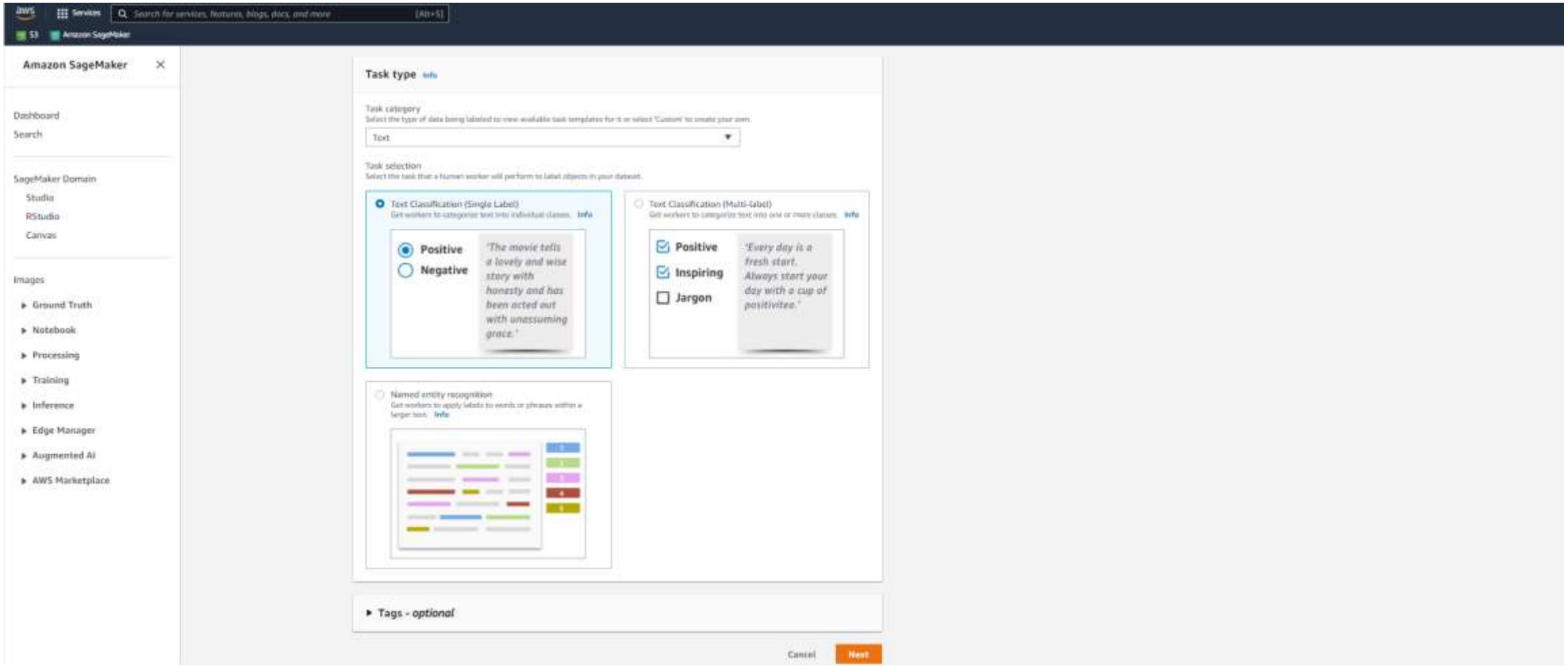
LABELING TEXT DATASET

PROVIDE A NAME TO THE LABELING JOB, SELECT AUTOMATED DATA SETUP AND PROVIDE THE S3 PATH. CLICK ON COMPLETE DATA SETUP TO GENERATE THE INPUT MANIFEST FILE.



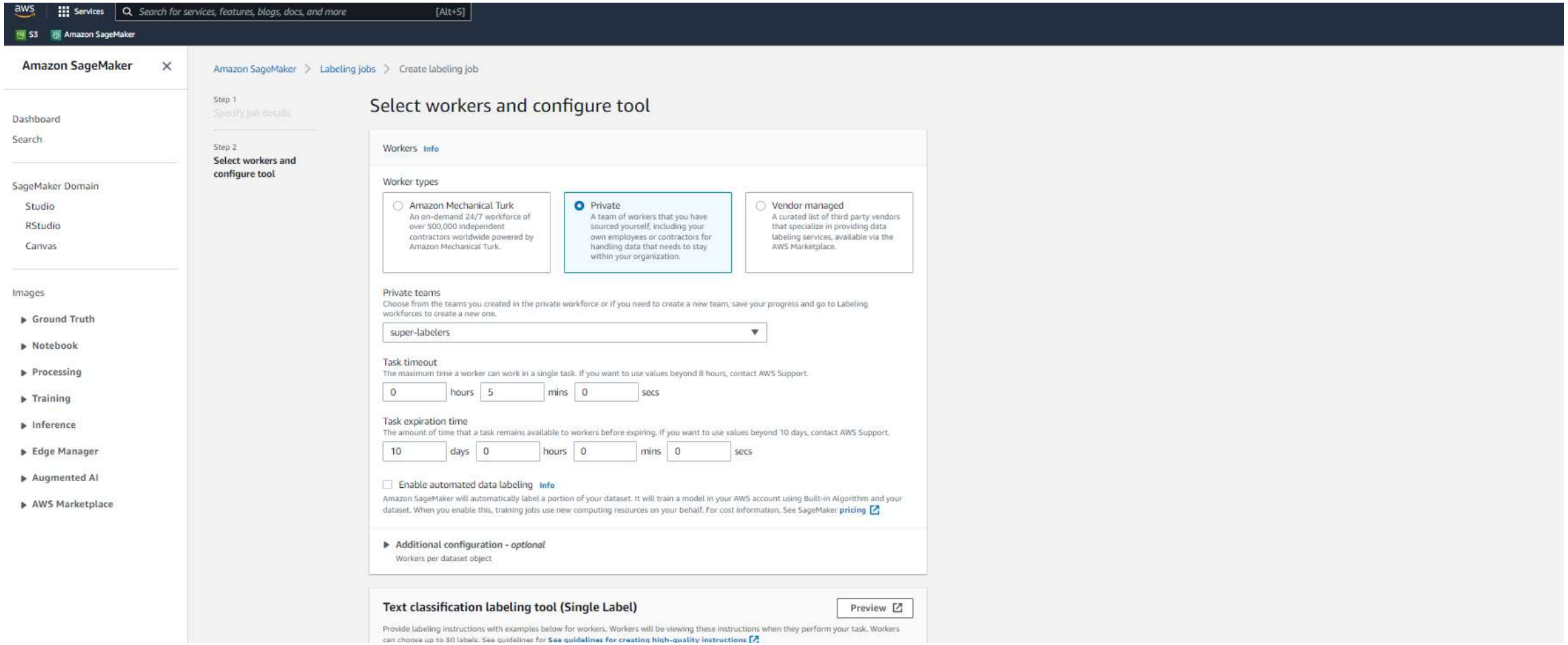
LABELING TEXT DATASET

SELECT TEXT AND TEXT CLASSIFICATION THEN CLICK NEXT



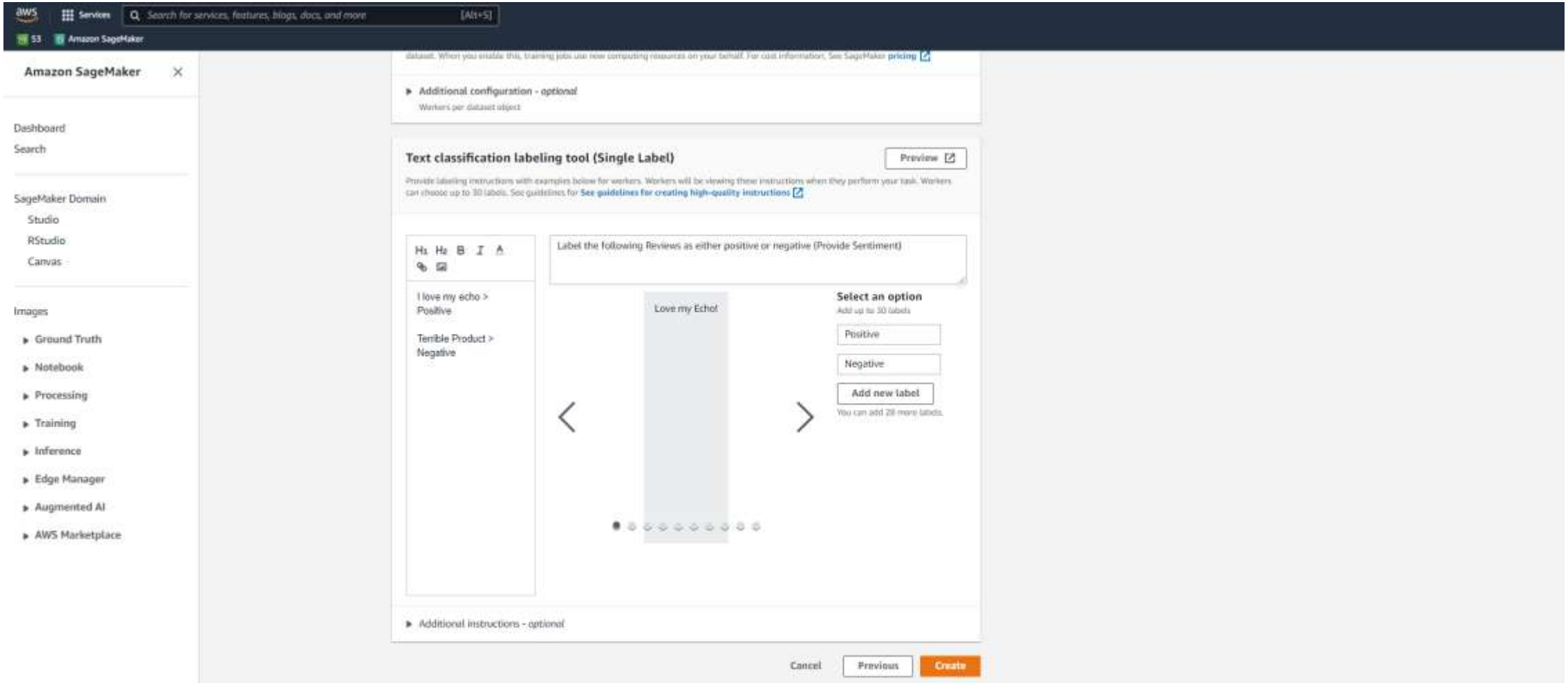
LABELING TEXT DATASET

CHOOSE PRIVATE AND SELECT YOUR TEAM



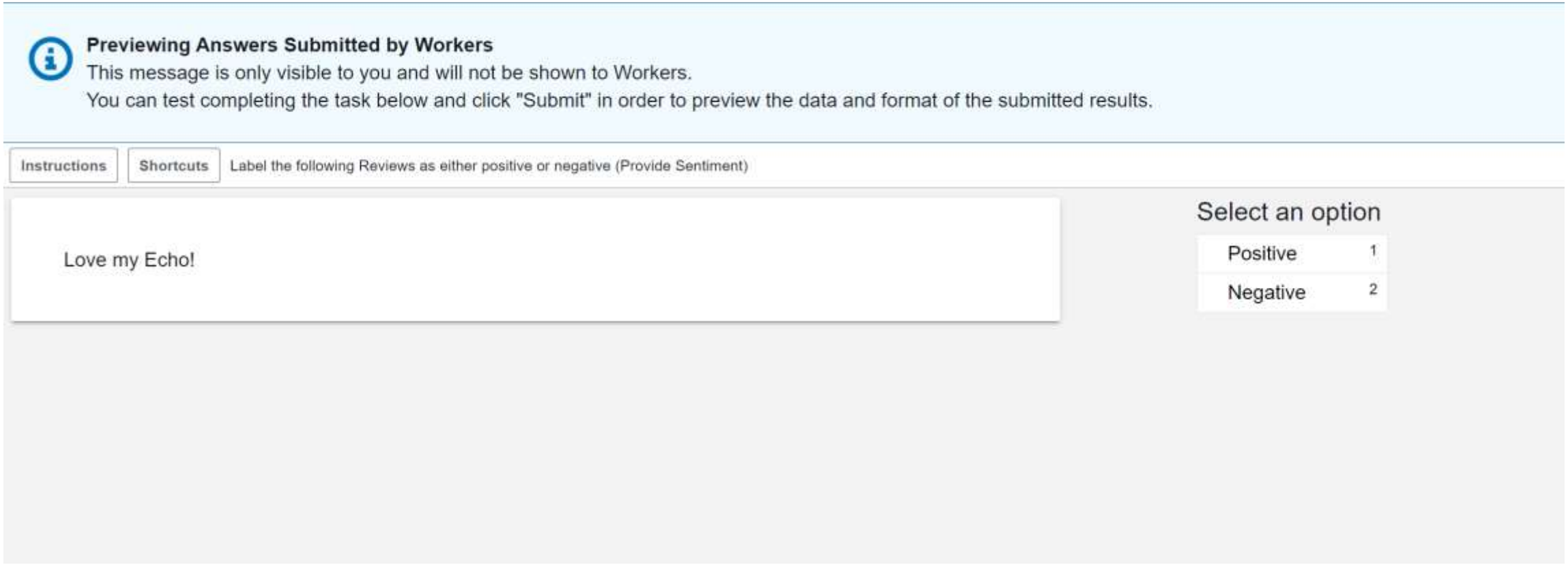
LABELING TEXT DATASET

PROVIDE GUIDANCE TO THE LABELING TEAM. CLICK PREVIEW TO SEE THE LABELING JOB!



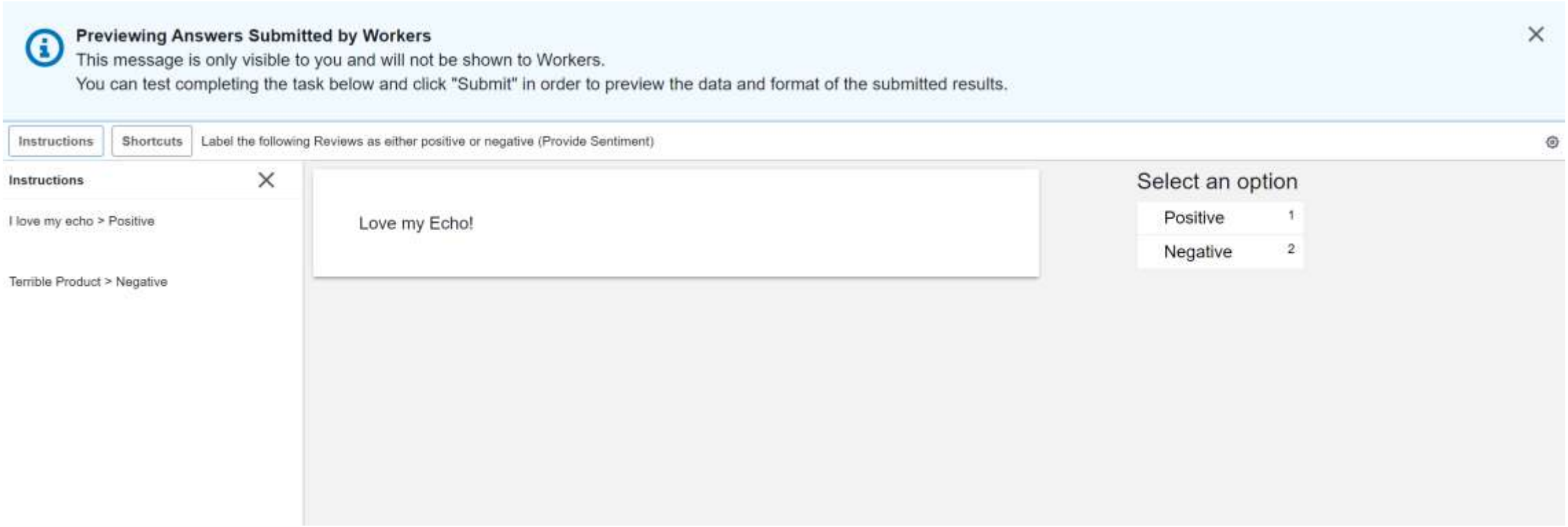
LABELING TEXT DATASET

VIEW OF THE LABELING JOB!



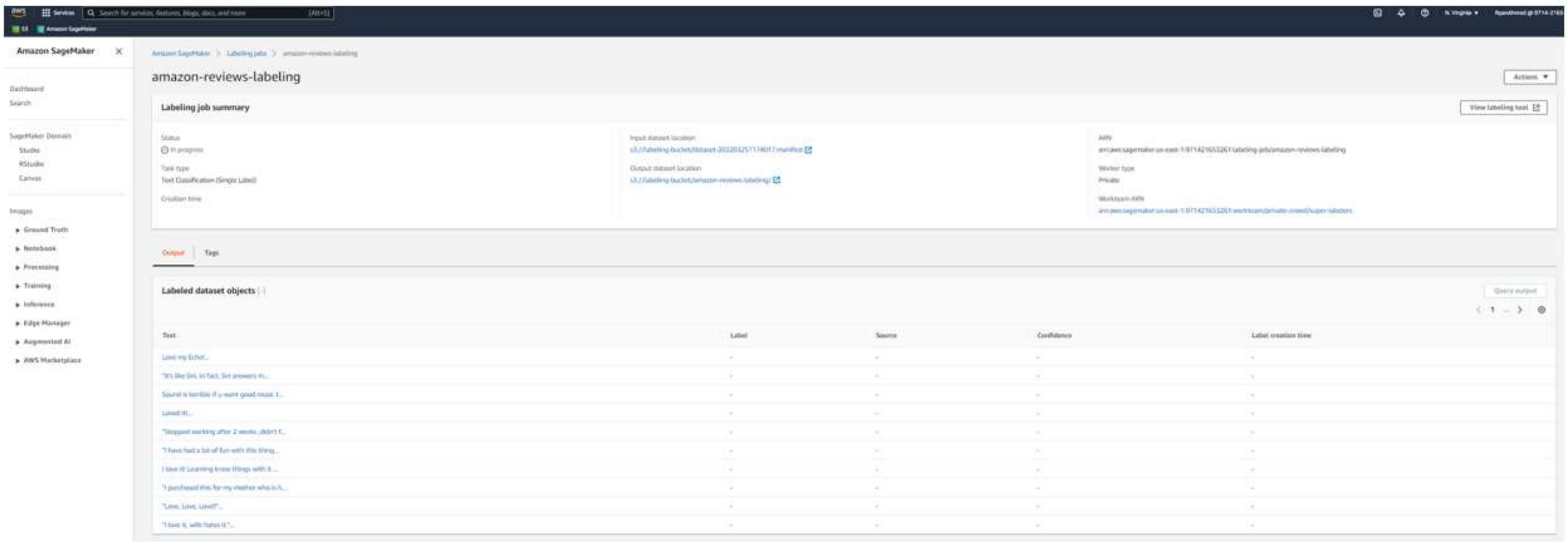
LABELING TEXT DATASET

CLICK ON INSTRUCTIONS. CLICK ON CREATE LABELING JOB ONCE YOU MADE SURE EVERYTHING LOOKS GOOD!



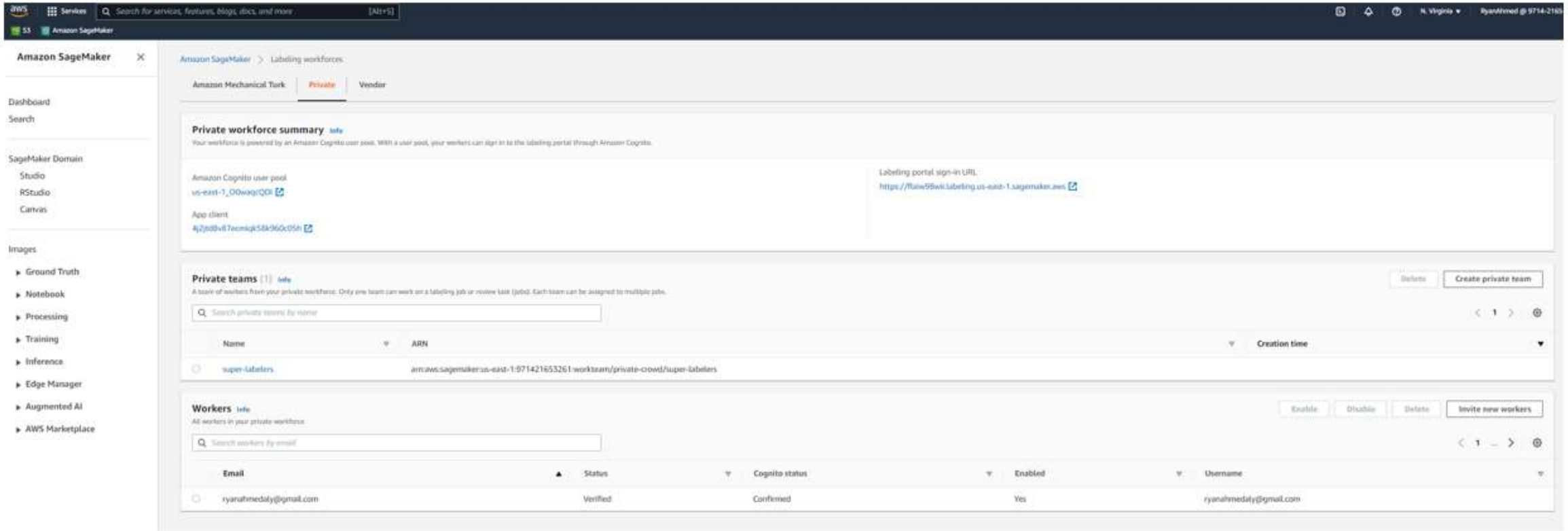
LABELING TEXT DATASET

CHECK OUT THE LABELING JOB DETAILS



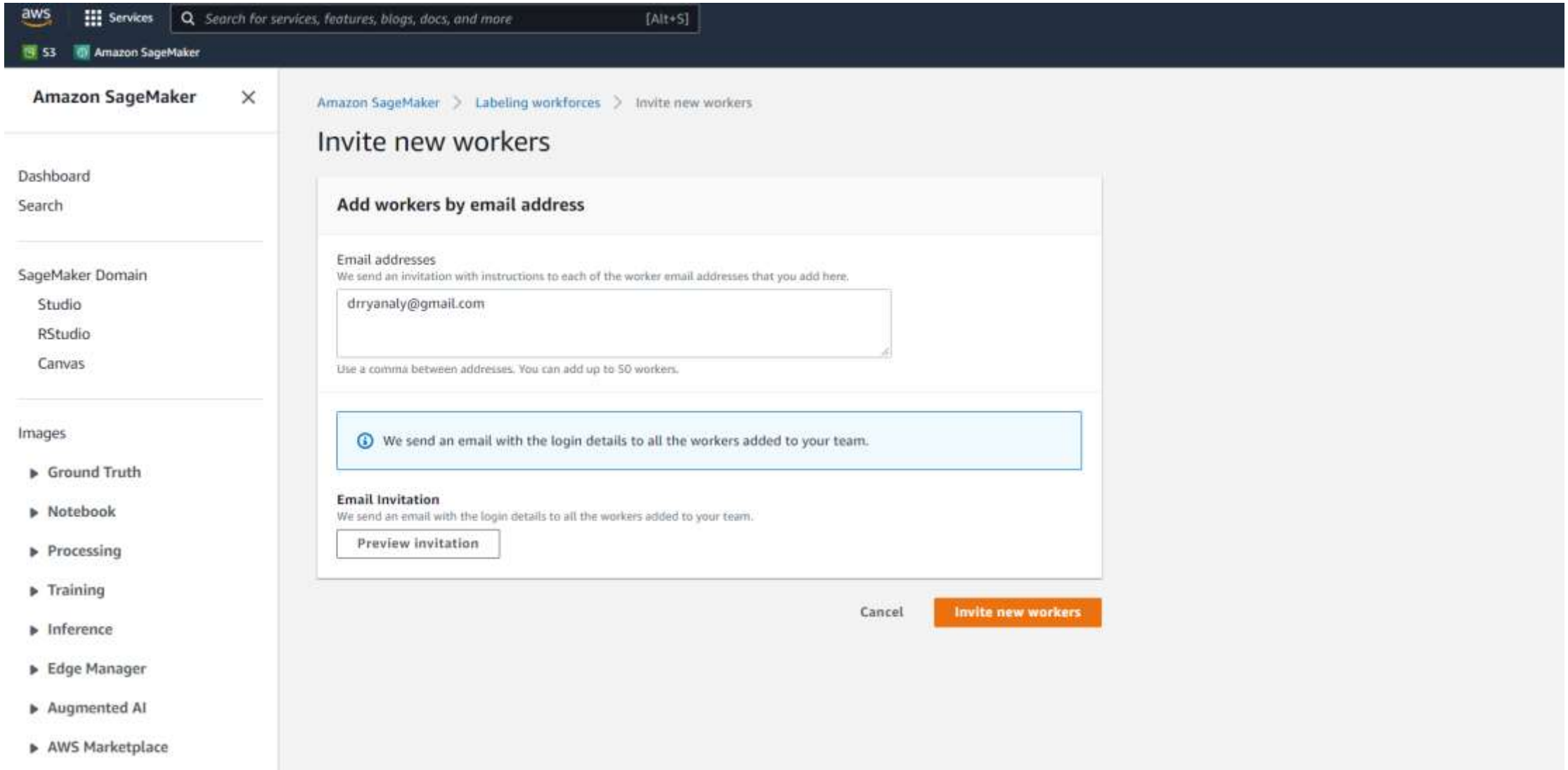
LABELING TEXT DATASET

NAVIGATE TO LABELING WORKFORCE AND INVITE TEAM MEMBERS



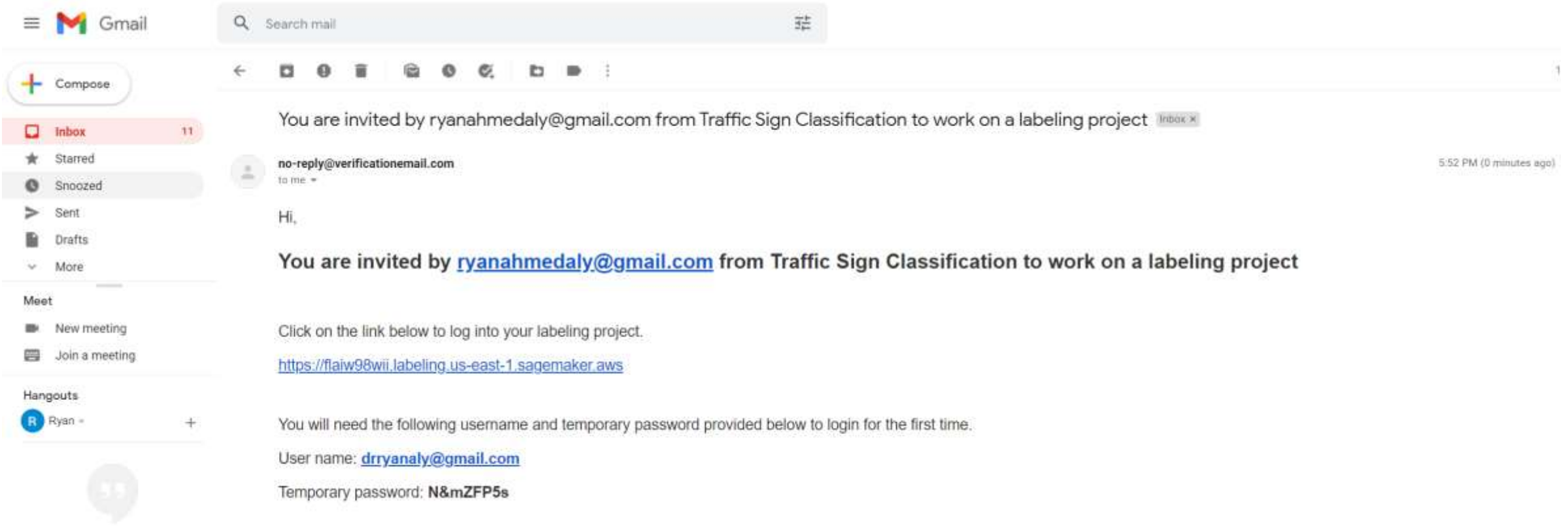
LABELING TEXT DATASET

ENTER THE EMAIL OF THE WORKER AND CLICK ON INVITE NEW WORKER



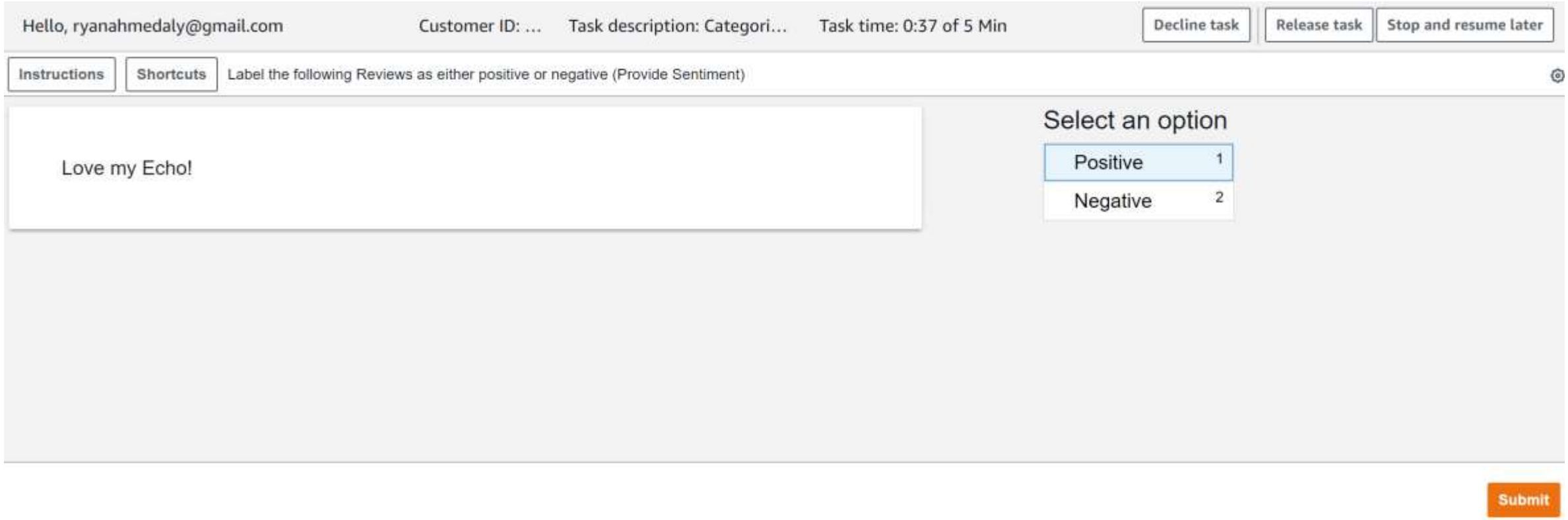
LABELING TEXT DATASET

YOU WILL RECEIVE THE FOLLOWING EMAIL, LOGIN WITH THE PROVIDED USERNAME AND PASSWORD



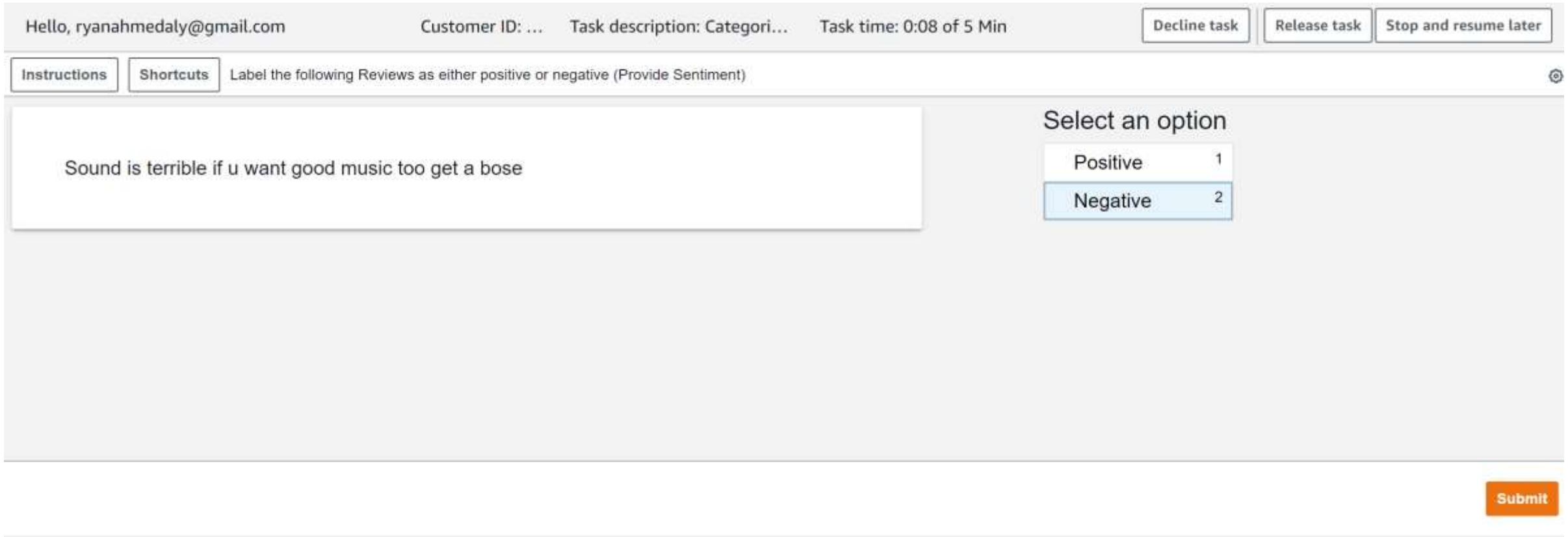
LABELING TEXT DATASET

START LABELING! HAVE FUN! THIS REVIEW IS POSITIVE



LABELING TEXT DATASET

THIS REVIEW IS NEGATIVE

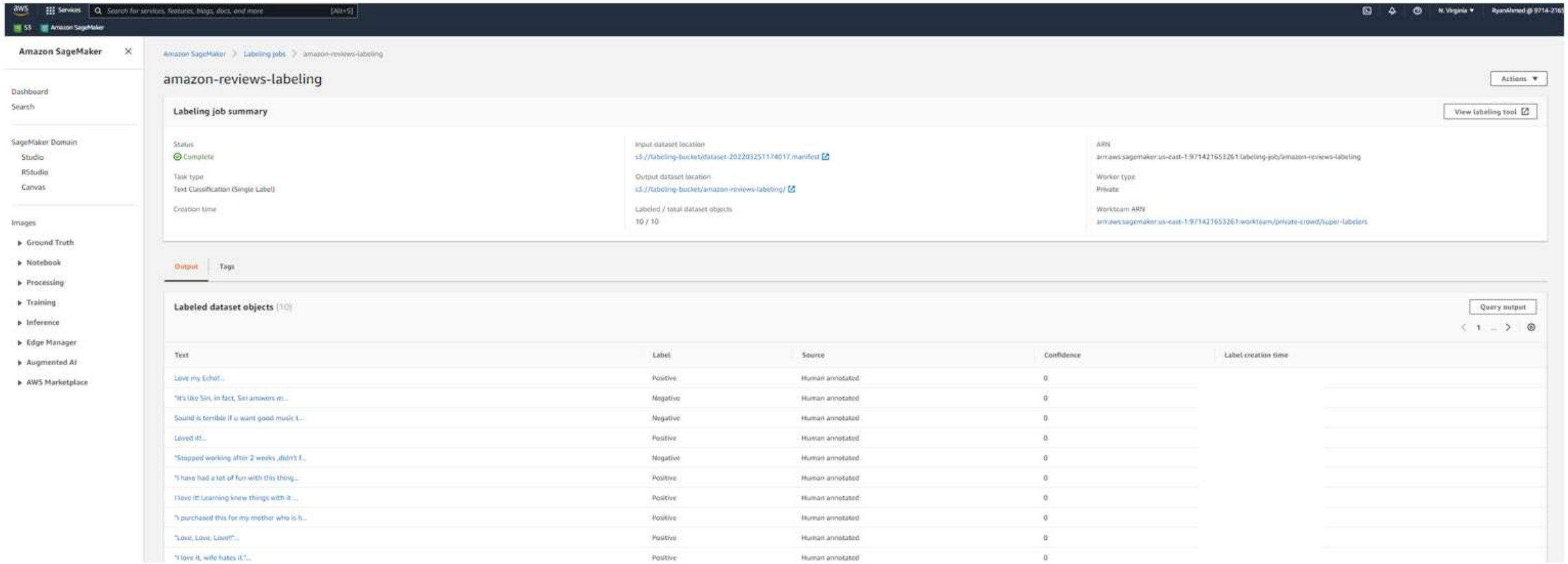


LABELING TEXT DATASET

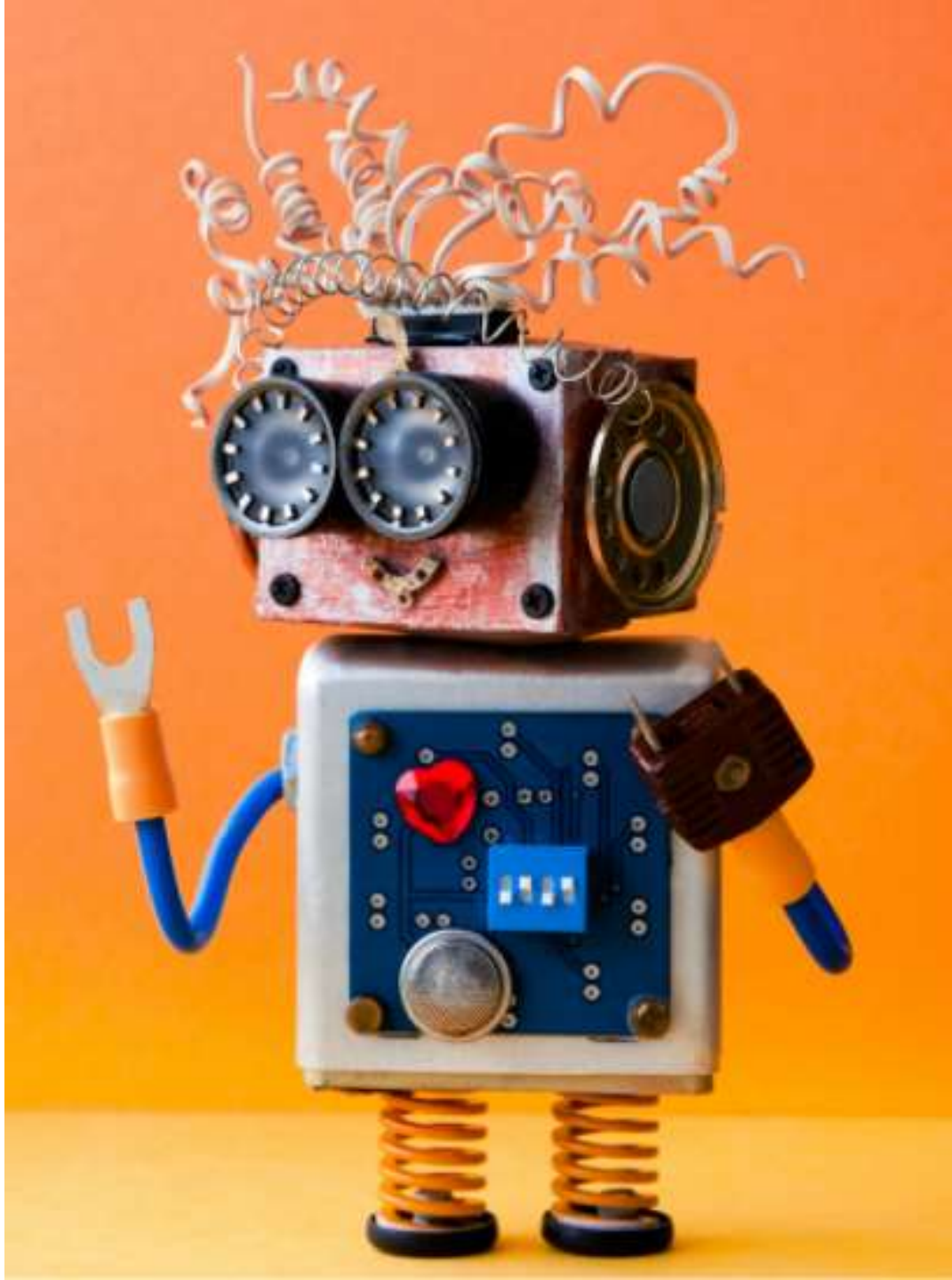
REVIEW THE OUTPUT MANIFEST FILE

LABELING TEXT DATASET

CHECK THE STATUS OF THE LABELING JOB

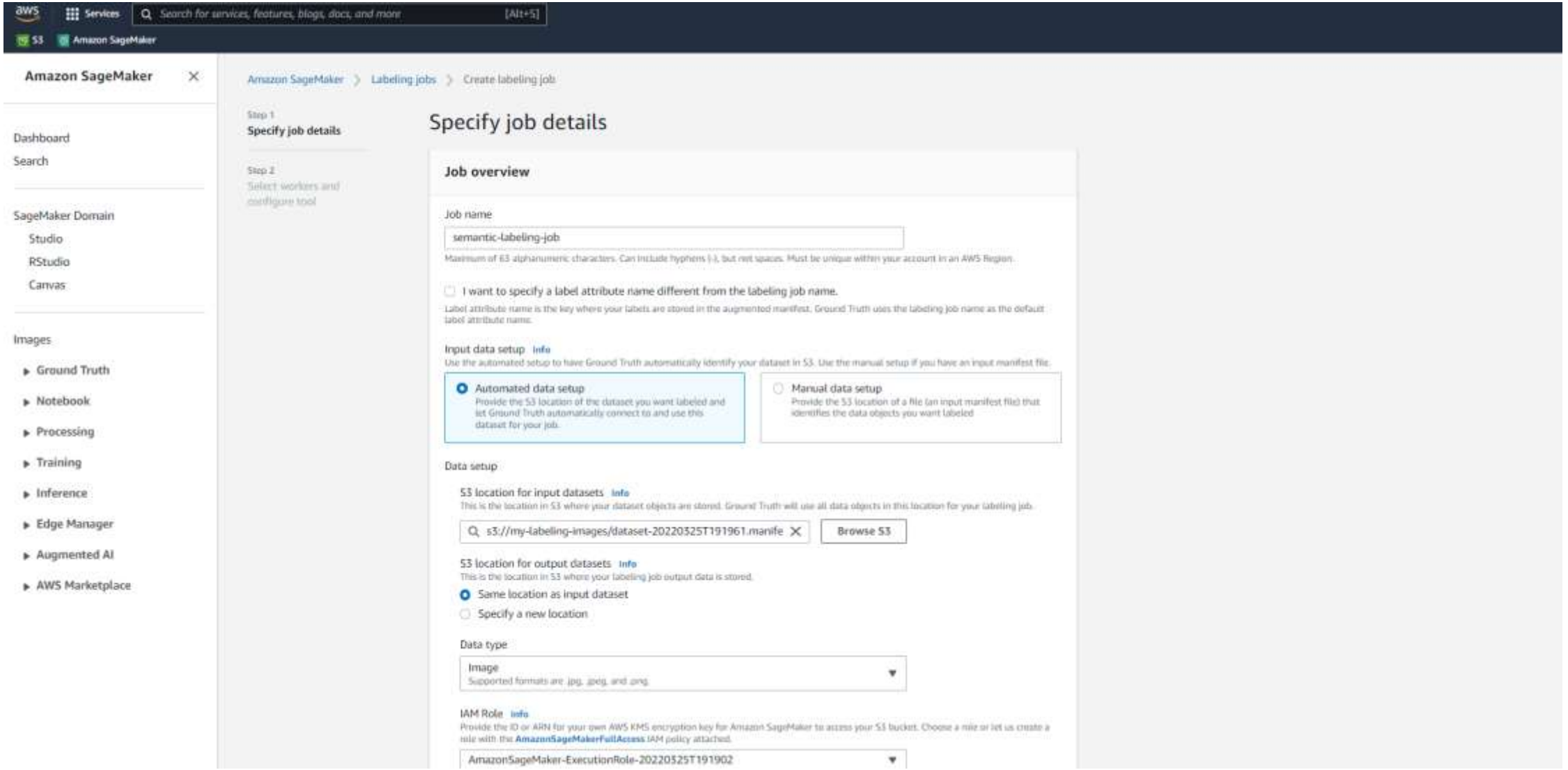


LABELING IMAGES: SEMANTIC SEGMENTATION



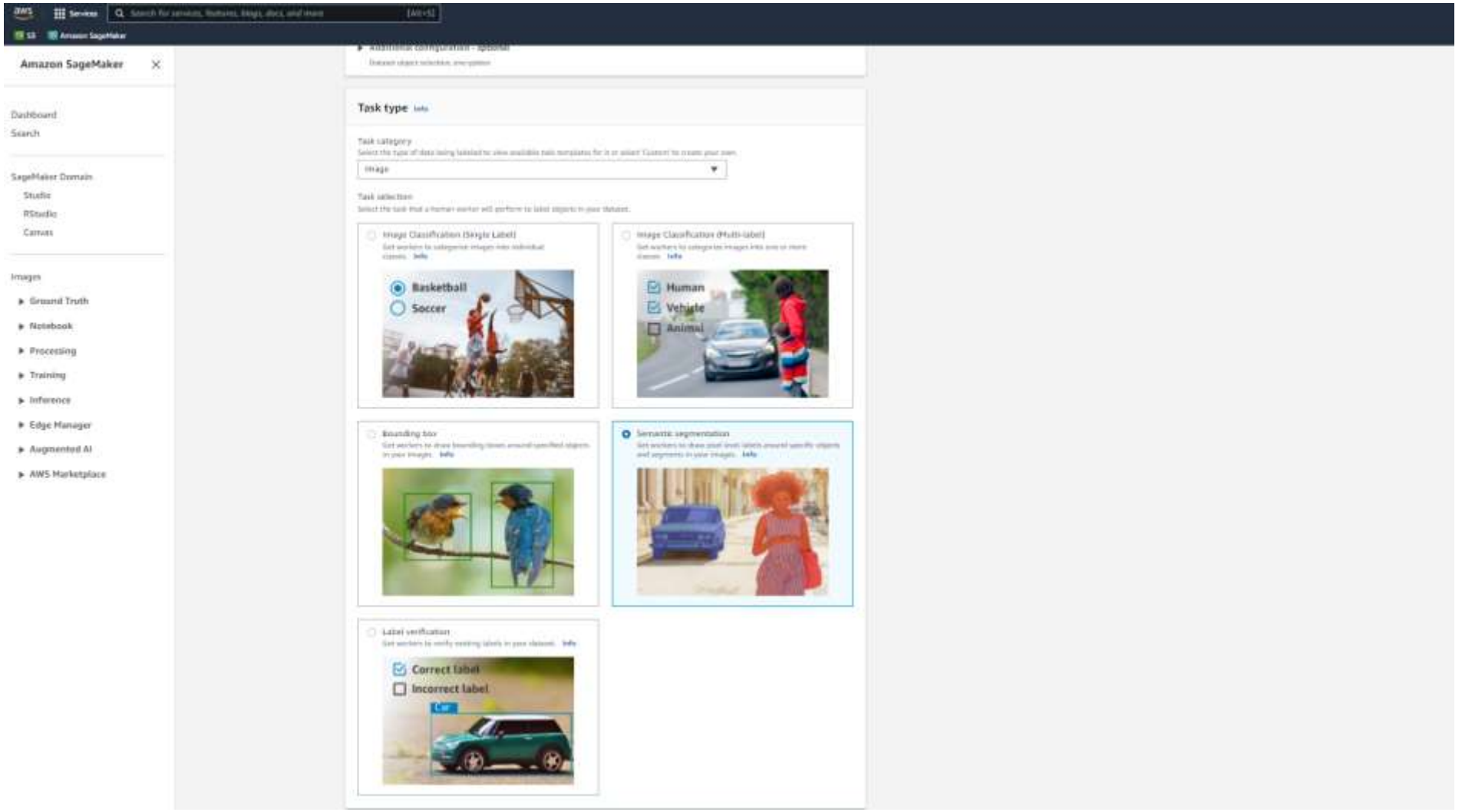
LABELING IMAGES: SEMANTIC SEGMENTATION

LET'S CREATE A SEMANTIC SEGMENTATION JOB



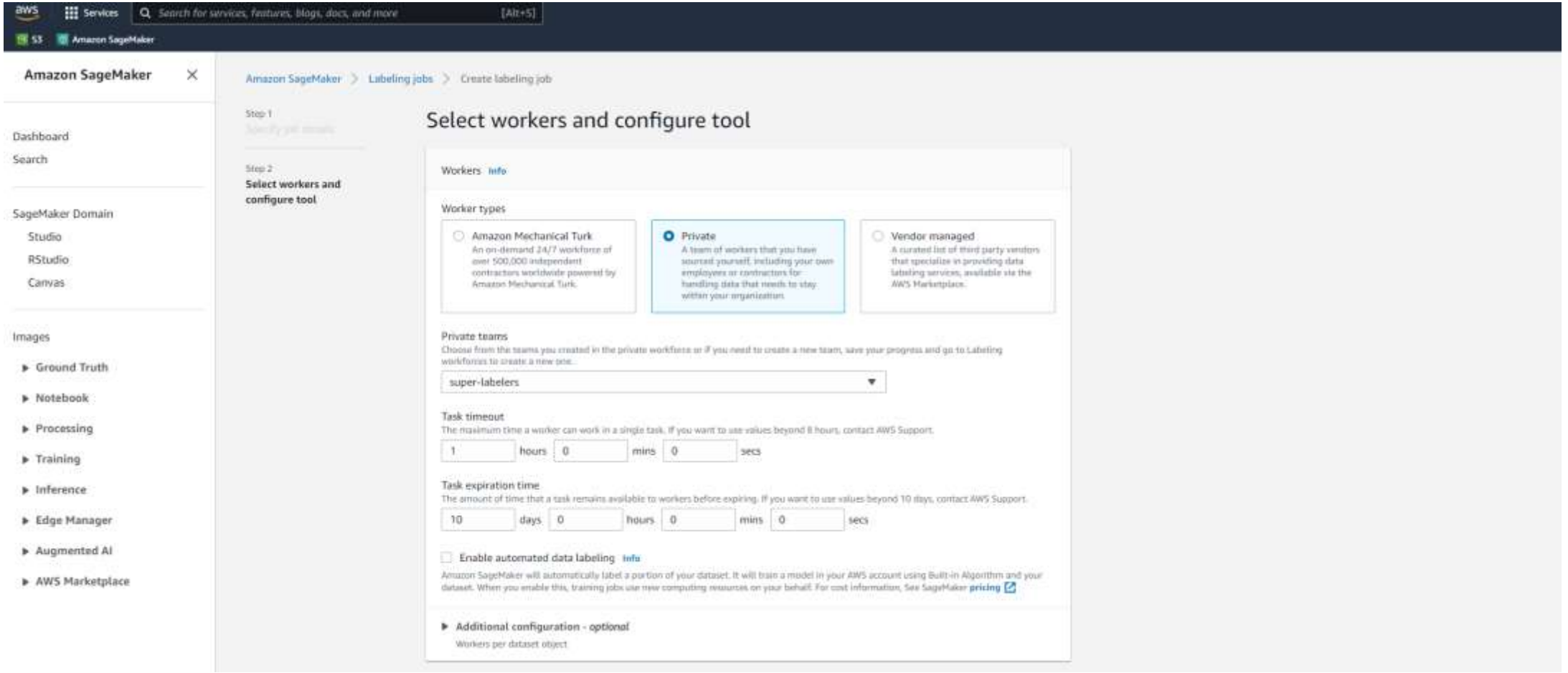
LABELING IMAGES: SEMANTIC SEGMENTATION

BOUNDING BOXES LABELING JOB IS NOW COMPLETE



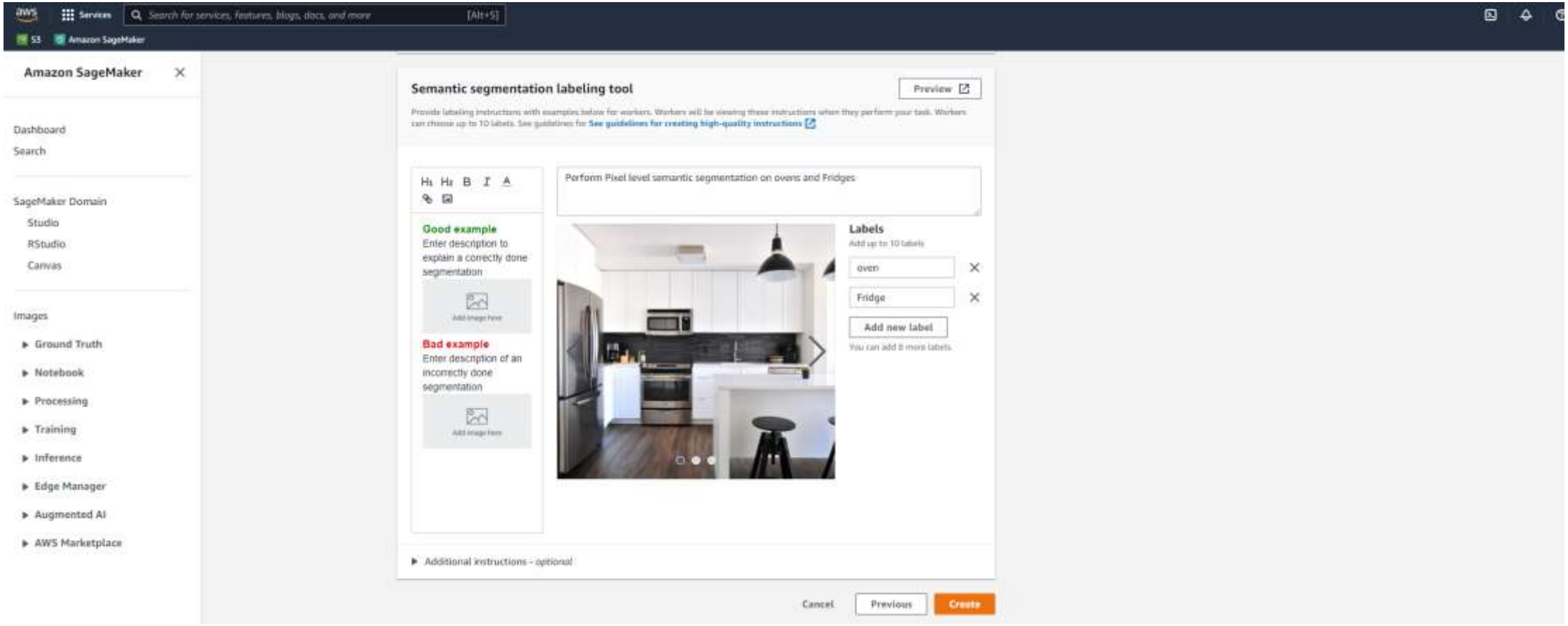
LABELING IMAGES: SEMANTIC SEGMENTATION

CHOOSE THE TEAM



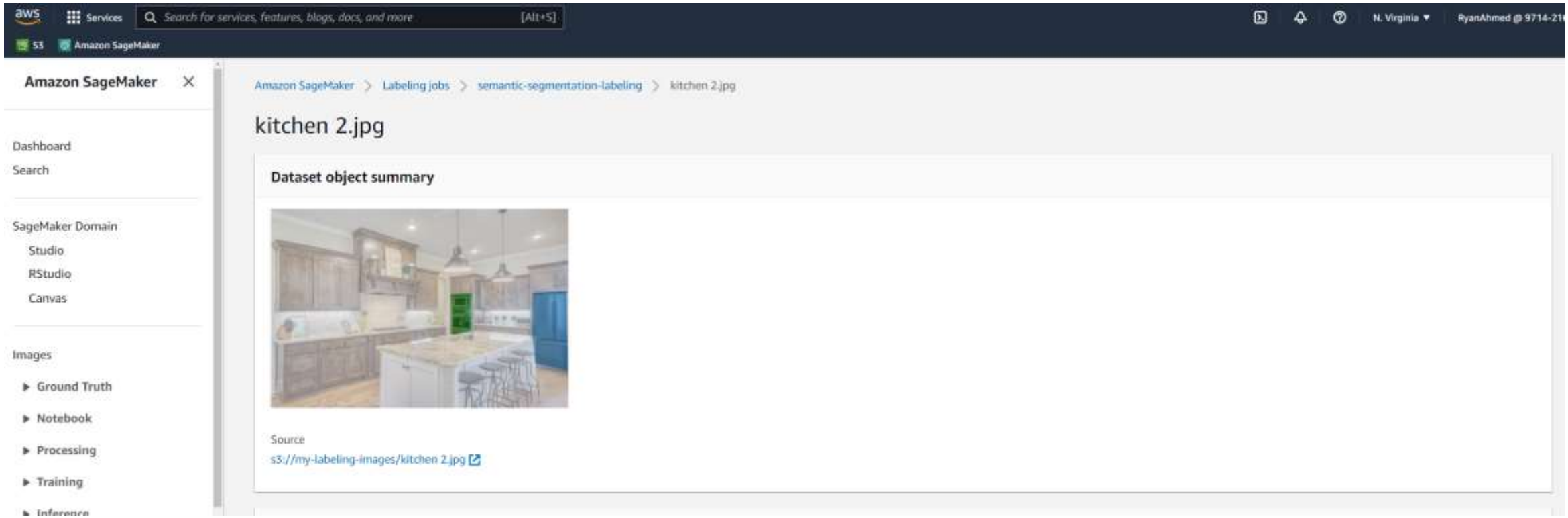
LABELING IMAGES: SEMANTIC SEGMENTATION

PROVIDE INSTRUCTIONS!



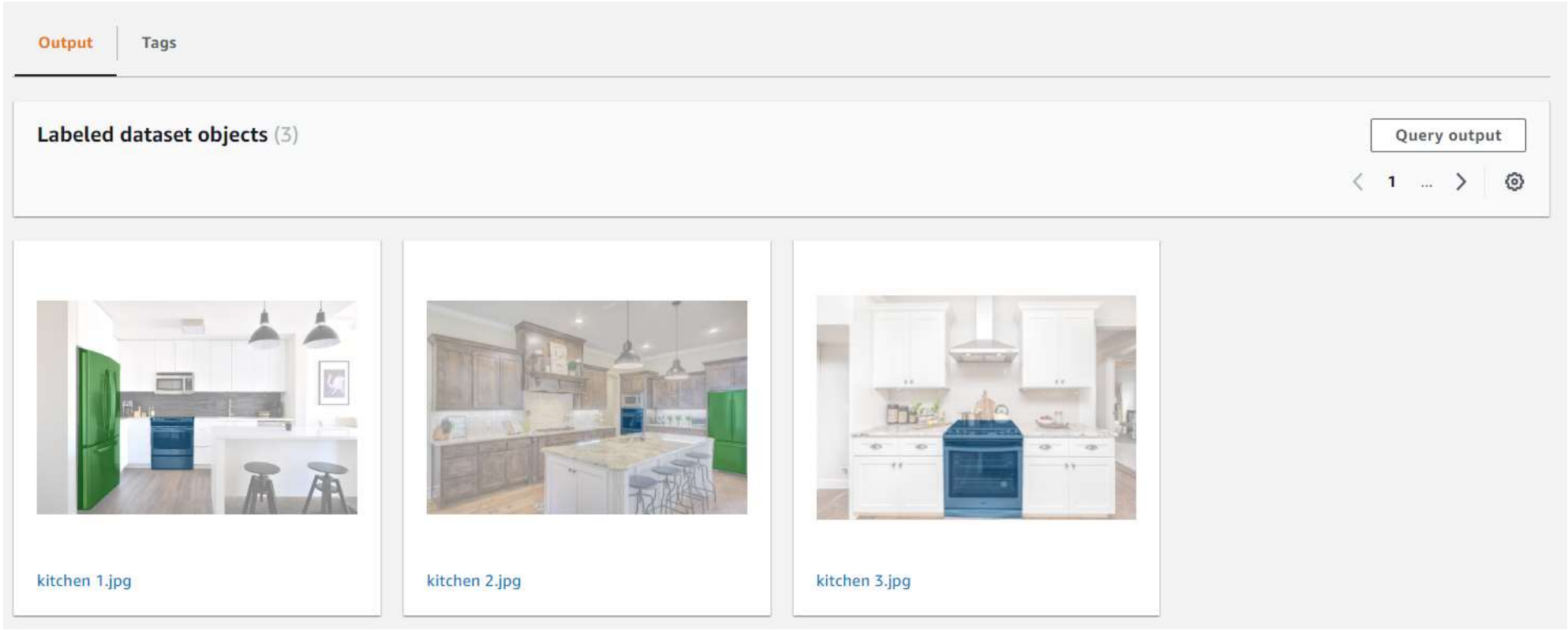
LABELING IMAGES: SEMANTIC SEGMENTATION

SELECT FRIDGE AND OVEN FROM THE IMAGES



LABELING IMAGES: SEMANTIC SEGMENTATION

LABELING JOB IS COMPLETE



LABELING IMAGES: SEMANTIC SEGMENTATION

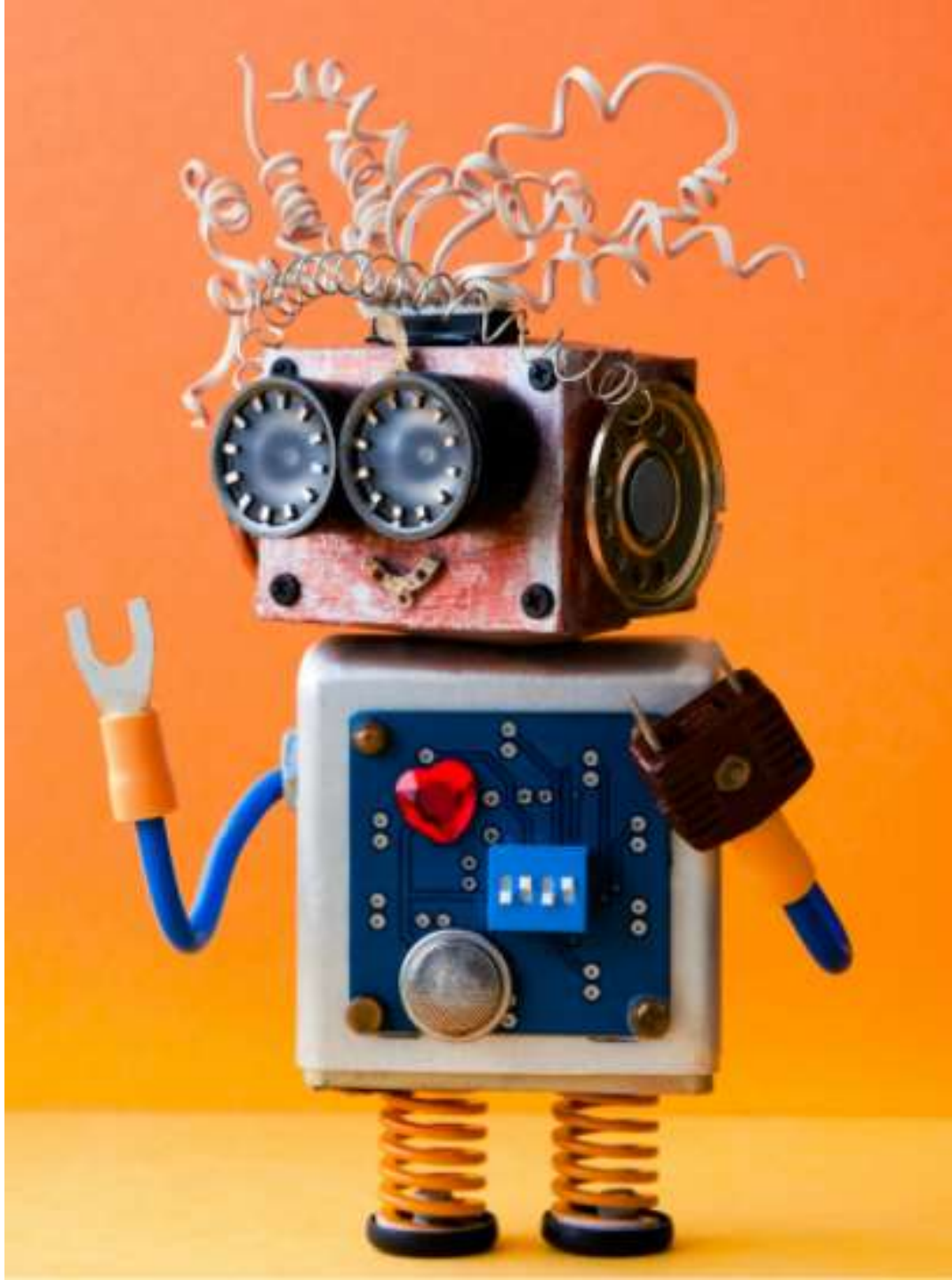
THE OUTPUT IS THE MANIFEST FILE ALONG WITH AN IMAGE CONTINAINING PIXEL LEVEL SEGMENTATION (WITH COLOR CODE THAT INDICATES THE CLASS)

output (5).manifest - Notepad

File Edit View


```
{"source-ref":"s3://my-labeling-images/kitchen 1.jpg","semantic-segmentation-labeling-ref":"s3://my-labeling-images/semantic-segmentation-
labeling/annotations/consolidated-annotation/output/0_2022-03-25T23:29:49.844419.png","semantic-segmentation-labeling-ref-metadata":{"internal-color-
map":{"0":{"class-name":"BACKGROUND","hex-color":"#ffffff","confidence":0},"1":{"class-name":"oven","hex-color":"#2ca02c","confidence":0},"2":{"class-
name":"Fridge","hex-color":"#1f77b4","confidence":0}},"type":"groundtruth/semantic-segmentation","human-annotated":"yes","creation-
date":                          ,"job-name":"labeling-job/semantic-segmentation-labeling"}}
{"source-ref":"s3://my-labeling-images/kitchen 2.jpg","semantic-segmentation-labeling-ref":"s3://my-labeling-images/semantic-segmentation-
labeling/annotations/consolidated-annotation/output/1_2022-03-25T23:32:45.839361.png","semantic-segmentation-labeling-ref-metadata":{"internal-color-
map":{"0":{"class-name":"BACKGROUND","hex-color":"#ffffff","confidence":0},"1":{"class-name":"oven","hex-color":"#2ca02c","confidence":0},"2":{"class-
name":"Fridge","hex-color":"#1f77b4","confidence":0}},"type":"groundtruth/semantic-segmentation","human-annotated":"yes","creation-
date":                          ,"job-name":"labeling-job/semantic-segmentation-labeling"}}
{"source-ref":"s3://my-labeling-images/kitchen 3.jpg","semantic-segmentation-labeling-ref":"s3://my-labeling-images/semantic-segmentation-
labeling/annotations/consolidated-annotation/output/2_2022-03-25T23:32:46.289645.png","semantic-segmentation-labeling-ref-metadata":{"internal-color-
map":{"0":{"class-name":"BACKGROUND","hex-color":"#ffffff","confidence":0},"1":{"class-name":"oven","hex-color":"#2ca02c","confidence":0},"2":{"class-
name":"Fridge","hex-color":"#1f77b4","confidence":0}},"type":"groundtruth/semantic-segmentation","human-annotated":"yes","creation-
date":                          ,"job-name":"labeling-job/semantic-segmentation-labeling"}}
```

GROUNDTRUTH PRICING



PRICING OF SAGEMAKER GROUNDTRUTH

- AWS Free Tier: For first two month, 500 objects labeled per month are free.
- Customers pay for each labeled object (image, audio, text) whether it's labeled automatically by SageMaker Ground Truth (autolabeling) or by humans.
- If Mechanical Turks or third-party vendors are used, an additional cost is added per labeled object.
- If customers leverage their own internal employees, no additional cost per labeled object is required.

REVIEWED OBJECTS (IMAGES, VIDEO FRAMES, TEXT, AUDIO FILES, ETC.)

NUMBER OF REVIEWED OBJECTS PER MONTH	PRICE PER REVIEWED OBJECT
Less than 50,000 objects	$0.08
50,000 to 1,000,000 objects	$0.04
Greater than 1,000,000 objects	$0.02

PRICING OF SAGEMAKER GROUNDTRUTH

- If Amazon Mechanical Turks are used for labeling, there is an additional charge per object per review instance.
- Multiple labelers per object are recommended to enhance the accuracy of the labels.

WORKFLOW	SUGGESTED PRICE PER LABELER
Image classification	$0.012
Text classification	$0.012
Named Entity Recognition (NER)	$0.024
Bounding box	$0.036
Semantic Segmentation	$0.84

PRICING OF SAGEMAKER GROUNDTRUTH: EXAMPLE #1

QUESTION

- Let's assume we have a company that is leveraging its own internal employees to label images.
- The total number of images they labeled is 40,000 images.
- The company uses Amazon SageMaker Groundtruth to manage the labeling process.
- Calculate the estimated cost of the labeling process.

SOLUTION

- Since the company used its own workforce and since the number of images is below 50,000. Then the cost of each image is $0.08.
- Total Cost = 40,000 human-labeled images x $0.08 per image = $3,200

AWS SageMaker Pricing Examples: https://aws.amazon.com/sagemaker/data-labeling/pricing/?nc=sn&loc=3&refid=ps_a134p0000078pqxaae&trkcampaign=acq_paid_search_brand

PRICING OF SAGEMAKER GROUNDTRUTH: EXAMPLE #2

QUESTION

- A company want to label 85,000 posts (text data) using Amazon SageMaker GroundTruth.
- The company set a custom workflow and set a payment of $0.036 per post.
- The company decided to label each post three times to improve the accuracy.
- Because the company used Mechanical Turk, the cost includes an additional charge of $0.036 for each human-labeled post to pay the labeler.

SOLUTION

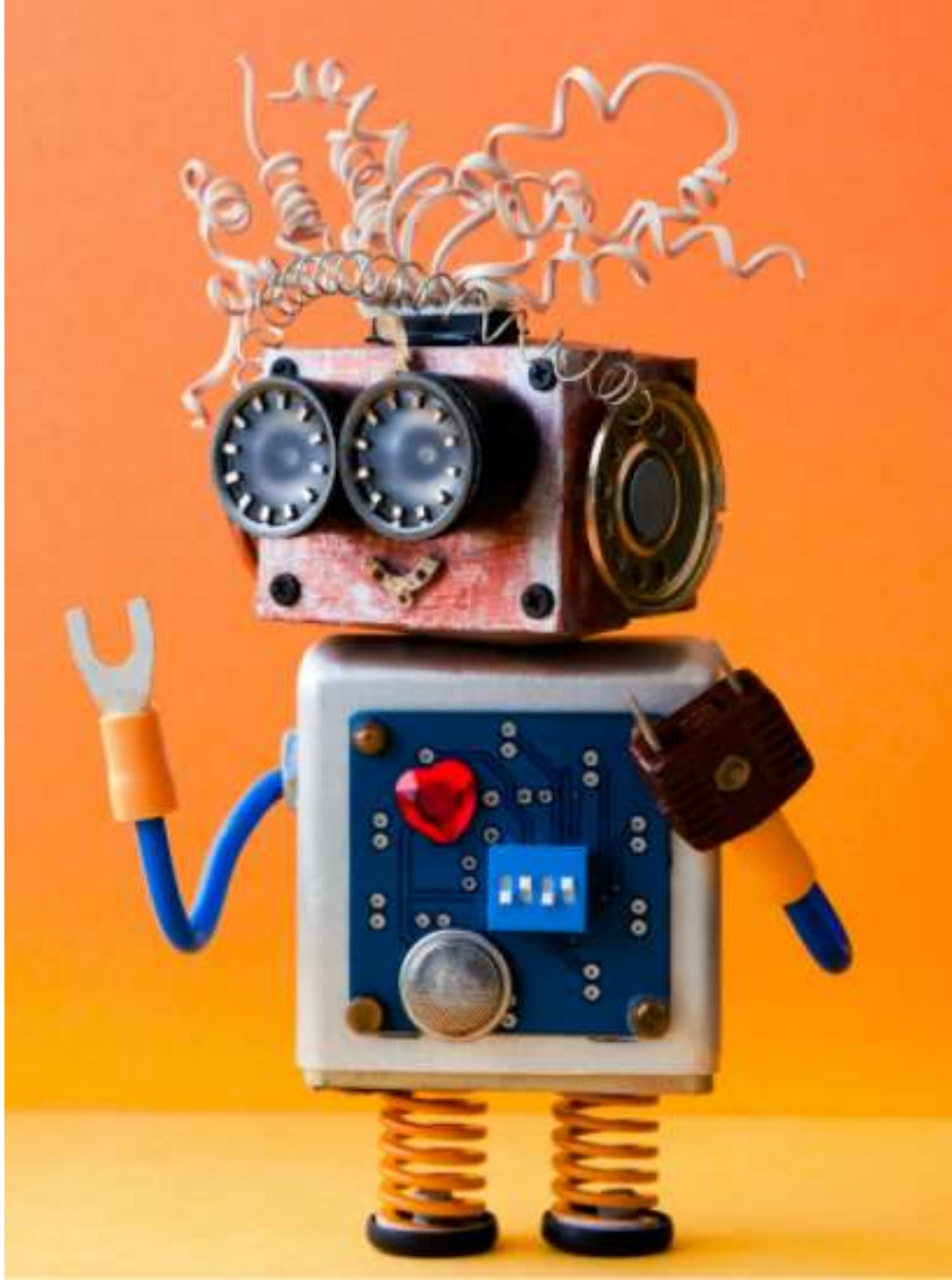
- Note that the first
- Total Cost = (50,000 x $0.08 per article) + (35,000 posts x $0.04 per post) + (85,000 human-labeled posts x $0.036 per post x 3 labelers per object) = $14,580

AWS SageMaker Pricing Examples: https://aws.amazon.com/sagemaker/data-labeling/pricing/?nc=sn&loc=3&refid=ps_a134p0000078pqxaae&trkcampaign=acq_paid_search_brand

GROUNDTRUTH PRICING PRACTICE OPPORTUNITY

EASY

ADVANCED



PRACTICE OPPORTUNITY

QUESTION

- A publishing company wants to natural language processing application to classify newspaper articles. To do this, they want to label 200,000 articles.
- Each article need to be labeled three times.
- Using SageMaker Ground Truth, humans label 40,000 articles and 160,000 are labeled automatically.
- Because the company used Mechanical Turk, the text classification workflow included an additional charge of $0.012 for each human-labeled article to pay the labeler.

AWS SageMaker Pricing Examples: https://aws.amazon.com/sagemaker/data-labeling/pricing/?nc=sn&loc=3&refid=ps_a134p0000078pqxaae&trkcampaign=acq_paid_search_brand

PRACTICE OPPORTUNITY

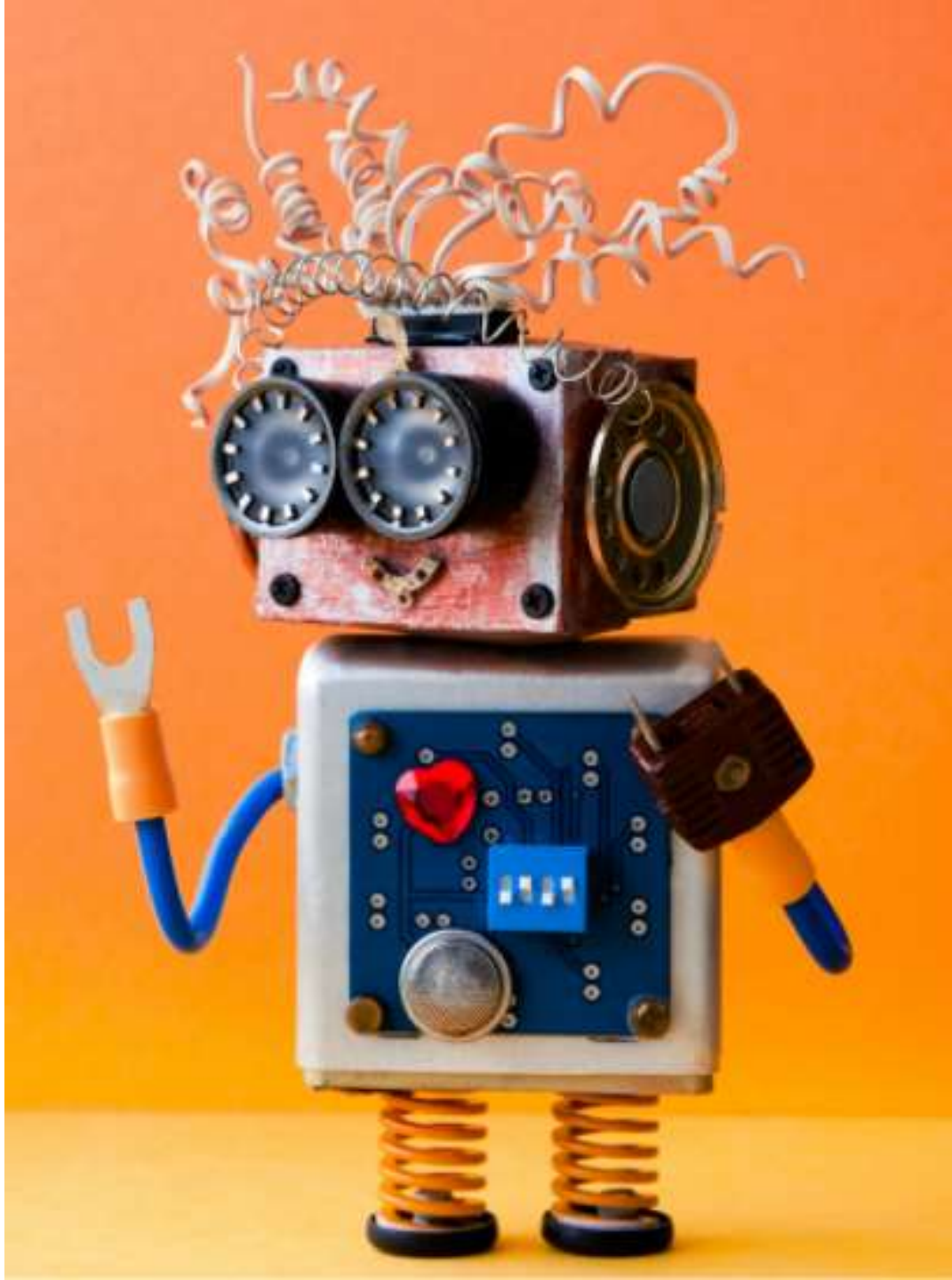
SOLUTION

- Total Cost = (50,000 x $0.08 per article) + (150,000 articles x $0.04 per article) + (40,000 human-labeled articles x $0.012 per article x 3 labelers per object) + Amazon SageMaker training & inference costs** = $11,440 + Amazon SageMaker training & inference costs**

- **These costs depend on a variety of factors, including the type of dataset being used, the type of labeling task, and the resolution of the images in your dataset.

AWS SageMaker Pricing Examples: https://aws.amazon.com/sagemaker/data-labeling/pricing/?nc=sn&loc=3&refid=ps_a134p0000078pqxaae&trkcampaign=acq_paid_search_brand

FINAL END-OF-DAY CAPSTONE PROJECT



PROJECT

1. Using the 3 kitchen images included in the course package.
2. Create a bounding box labeling job using Amazon SageMaker GroundTruth to label only the “fridge” and “oven”.
3. Review the input manifest file.
4. Perform the labeling job.
5. Review the output manifest file and ensure that the labeling job was successful.

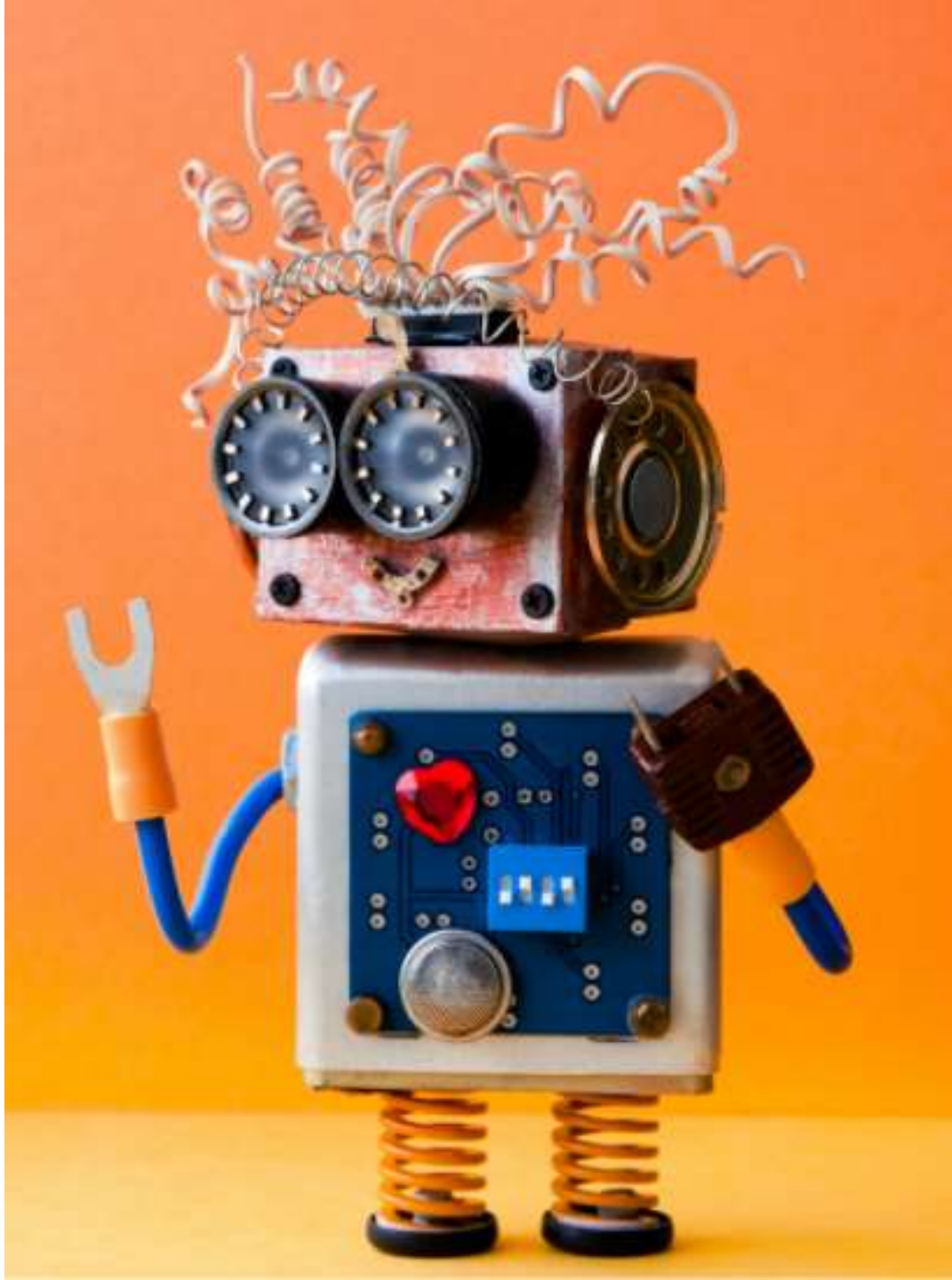
kitchen 1

kitchen 2

kitchen 3

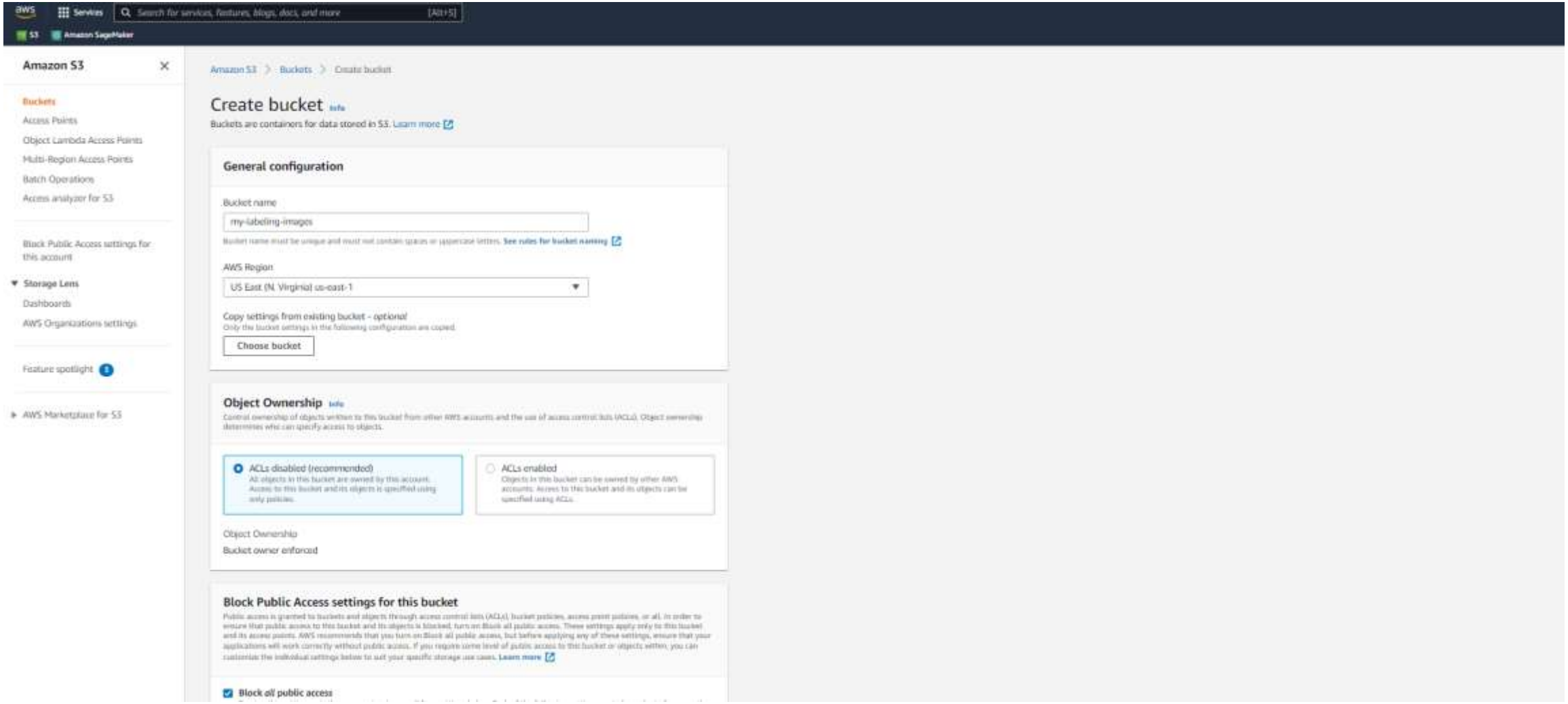
FINAL END-OF-DAY CAPSTONE PROJECT SOLUTION

EASY ADVANCED



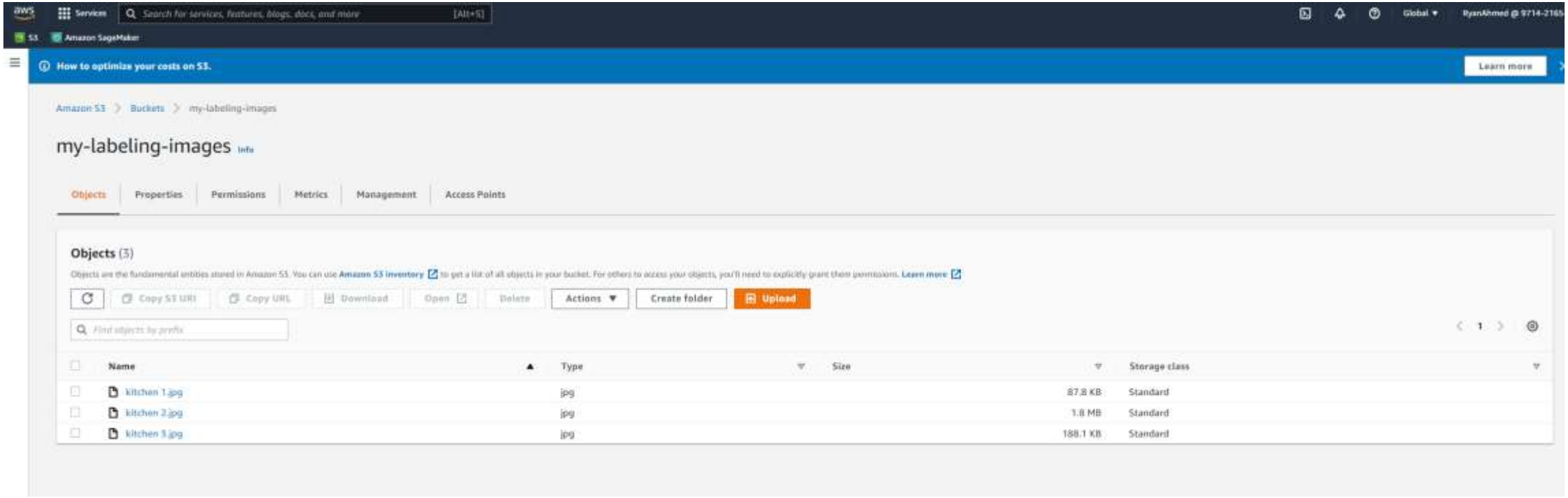
LABELING IMAGES WITH BOUNDING BOXES

UPLOAD THE IMAGES TO S3 BUCKET



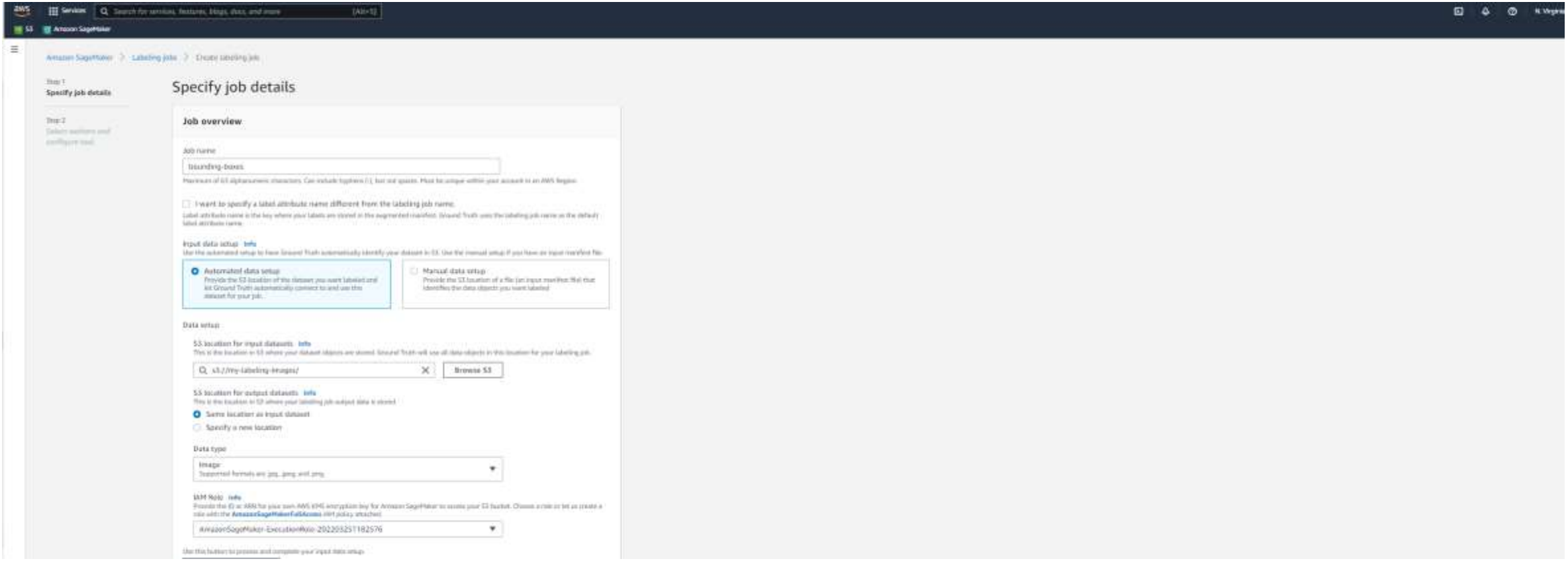
LABELING IMAGES WITH BOUNDING BOXES

UPLOAD IMAGES



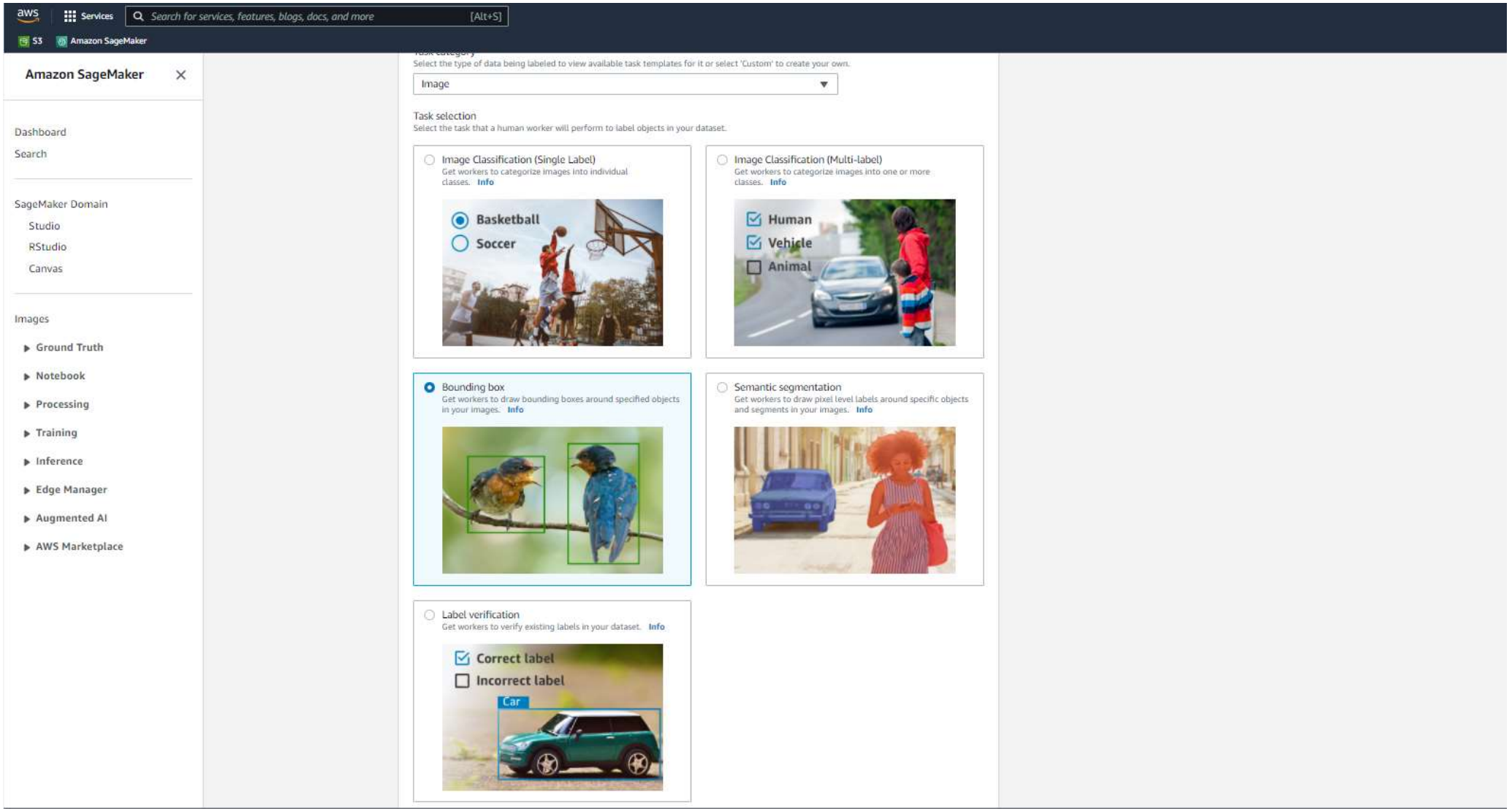
LABELING IMAGES WITH BOUNDING BOXES

CREATE A LABELING JOB. DON'T FORGET TO CLICK ON
COMPLETE DATA SETUP TO GENERATE THE MANIFEST FILE



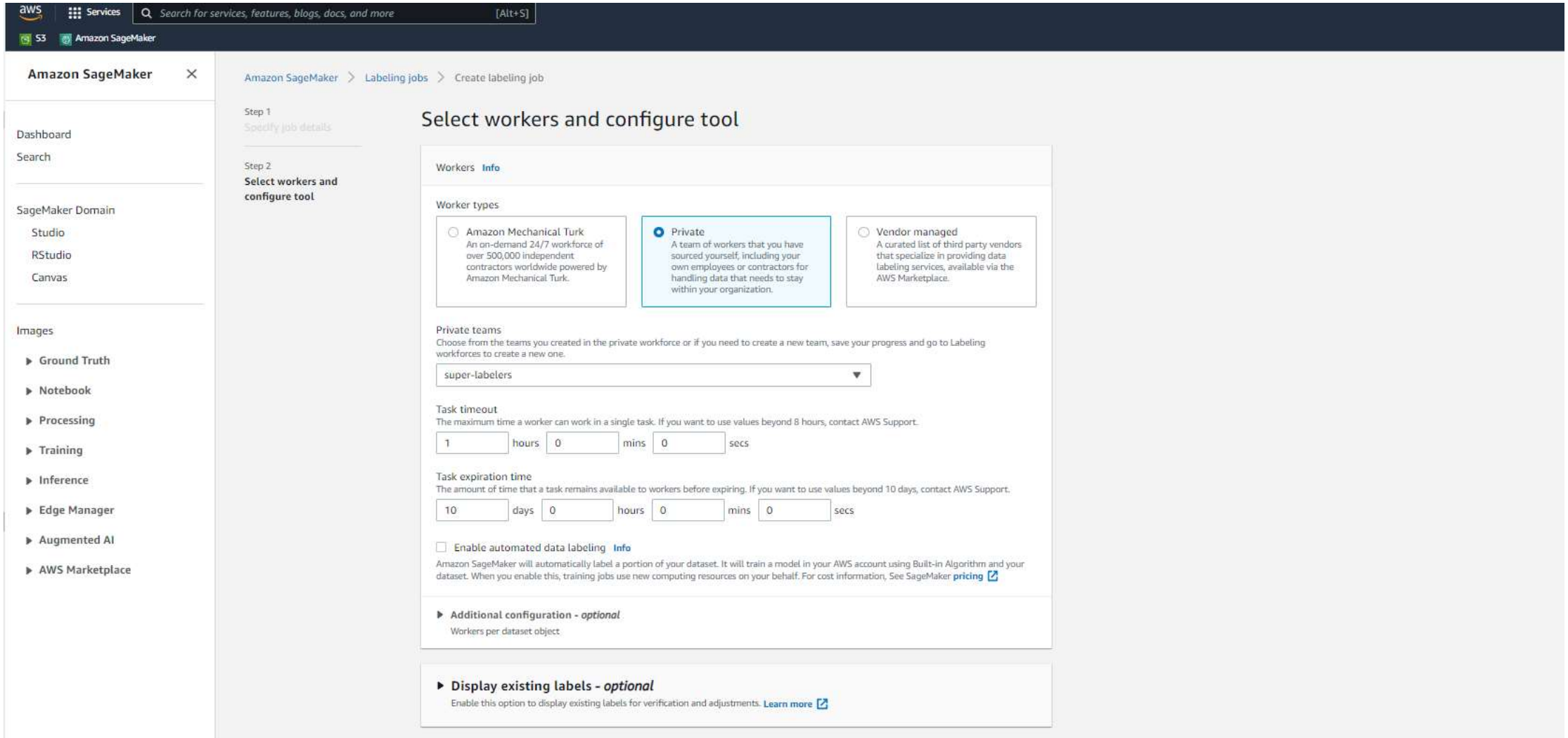
LABELING IMAGES WITH BOUNDING BOXES

SELECT THE BOUNDING BOX OPTION AND CLICK NEXT



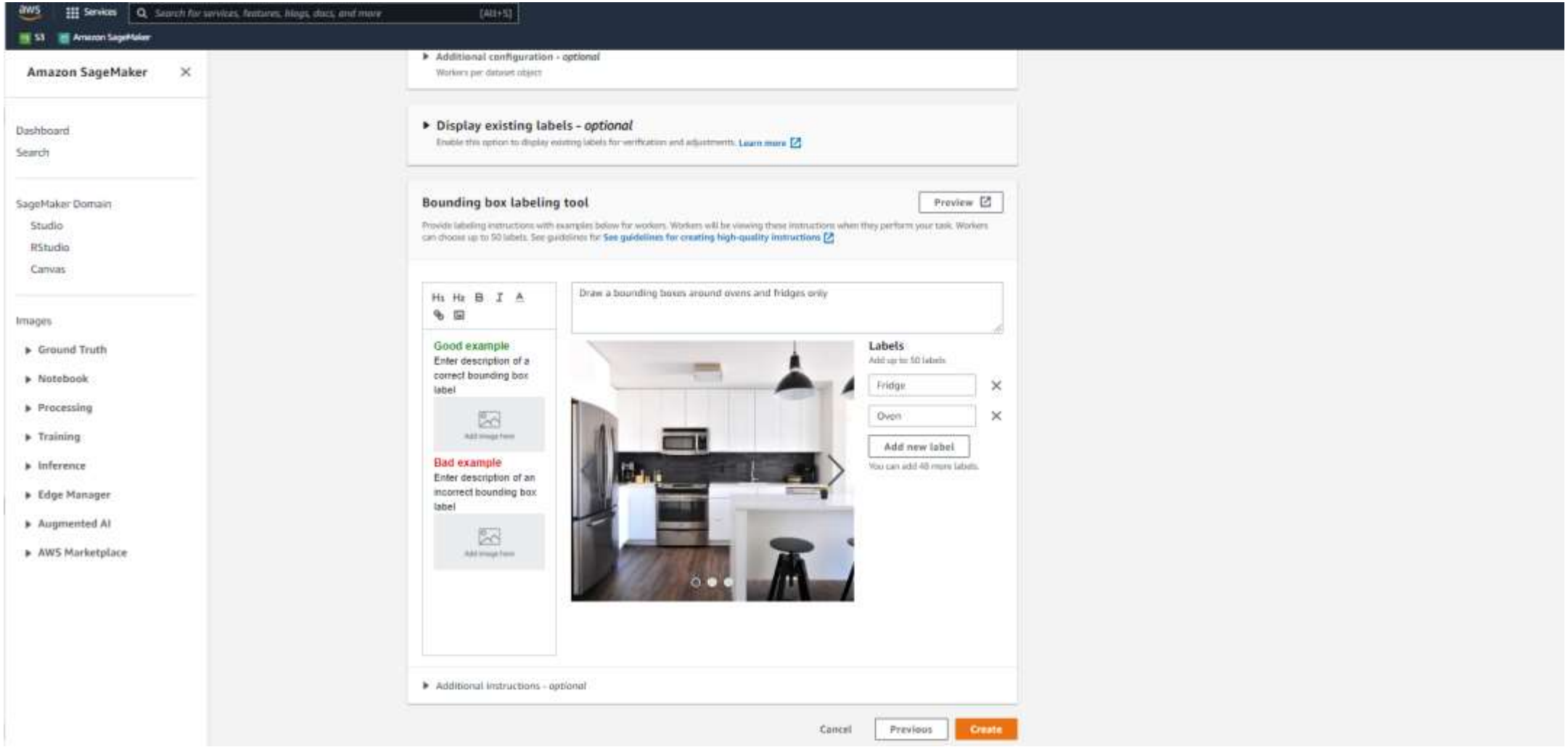
LABELING IMAGES WITH BOUNDING BOXES

SELECT PRIVATE AND CHOOSE SUPER-LABELERS



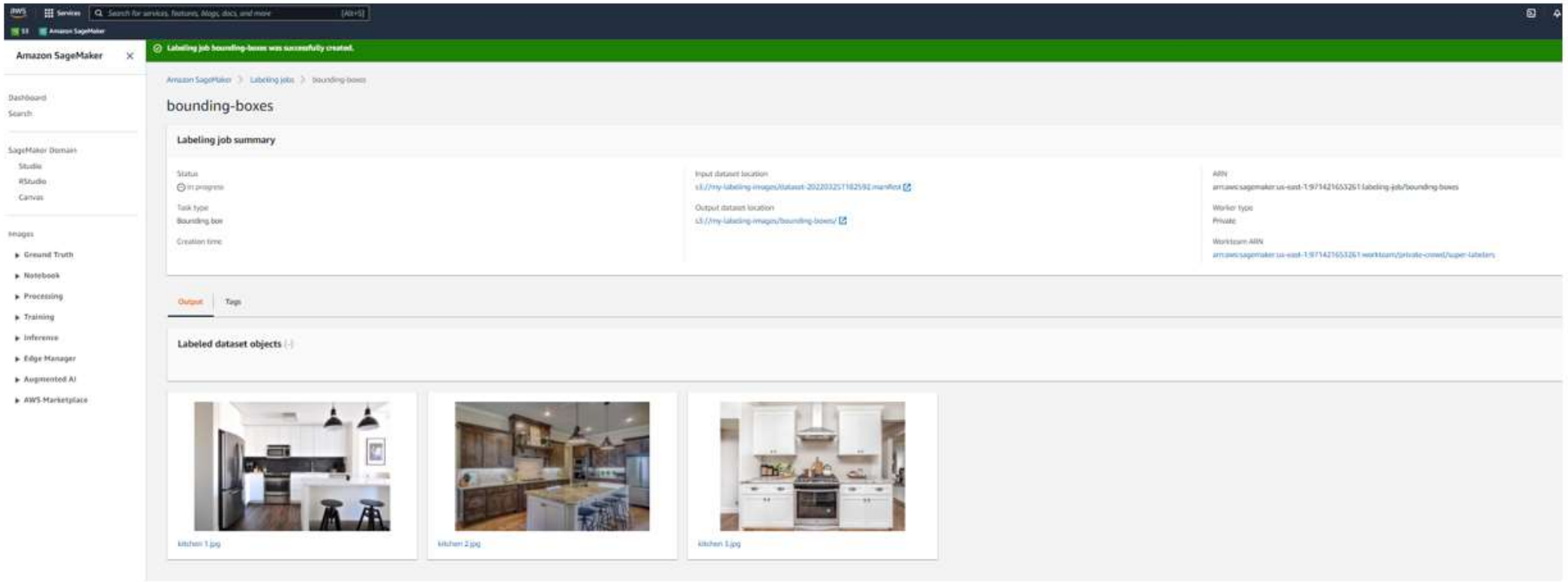
LABELING IMAGES WITH BOUNDING BOXES

PROVIDE INSTRUCTIONS AND CLICK CREATE!



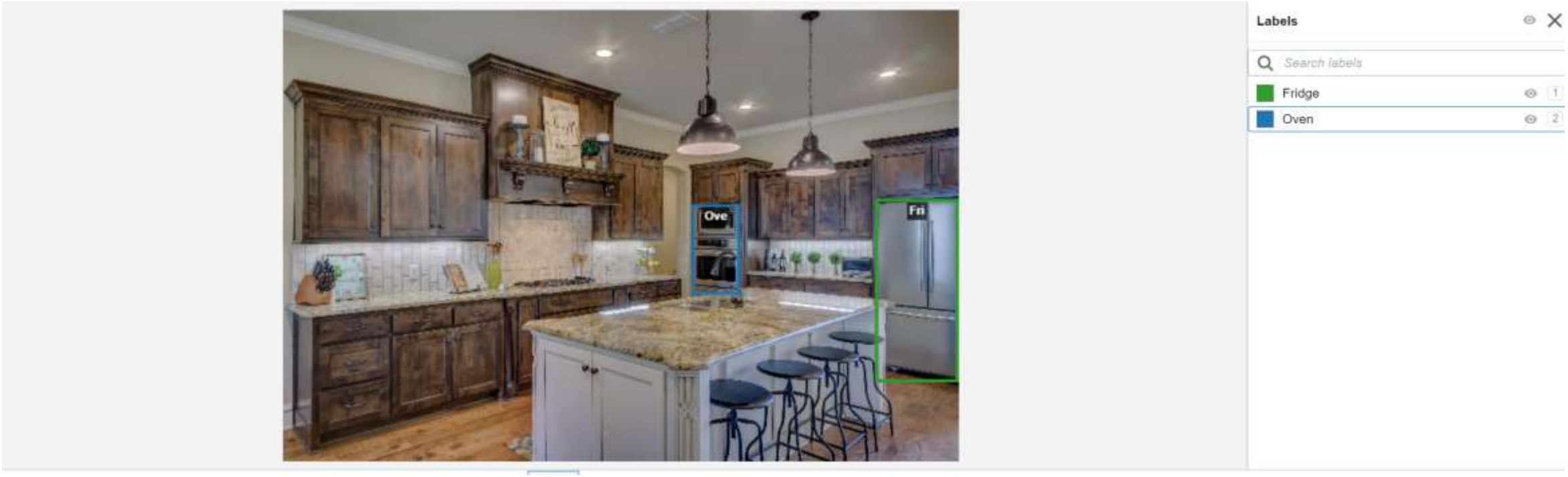
LABELING IMAGES WITH BOUNDING BOXES

REVIEW THE LABELING JOB HERE!



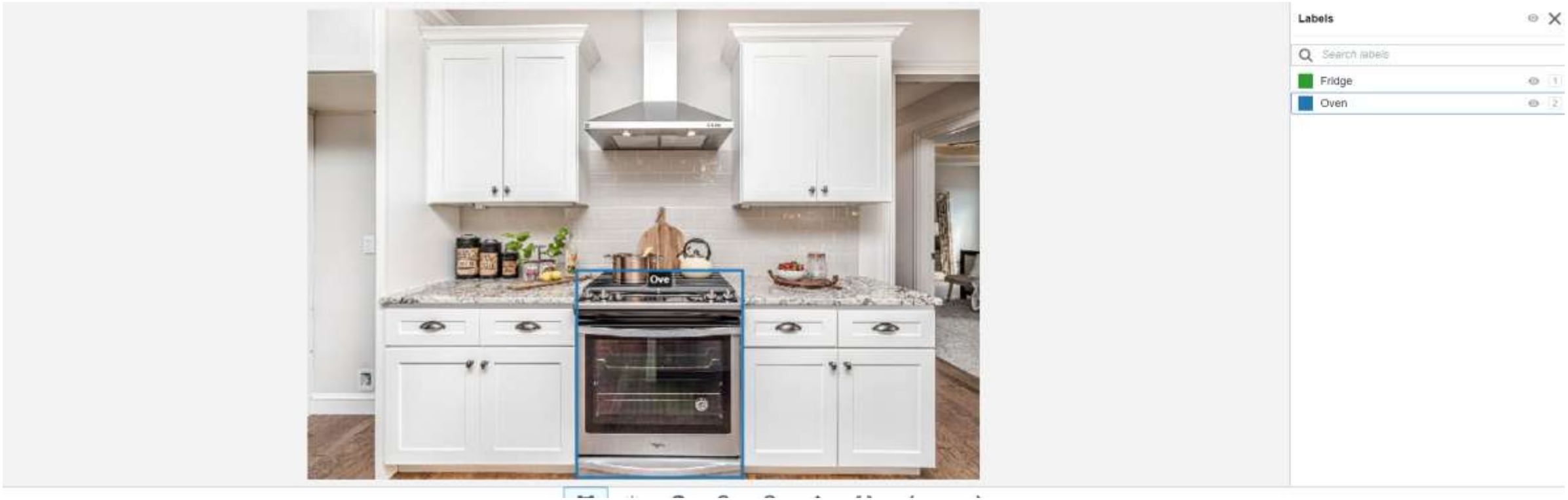
LABELING IMAGES WITH BOUNDING BOXES

LABEL THE FRIDGE AND OVEN



LABELING IMAGES WITH BOUNDING BOXES

LABEL THE OVEN IN THE IMAGE



LABELING IMAGES WITH BOUNDING BOXES

COORDINATES OF THE BOUNDING BOXES

Previewing Answers Submitted by Workers

This message is only visible to you and will not be shown to Workers.

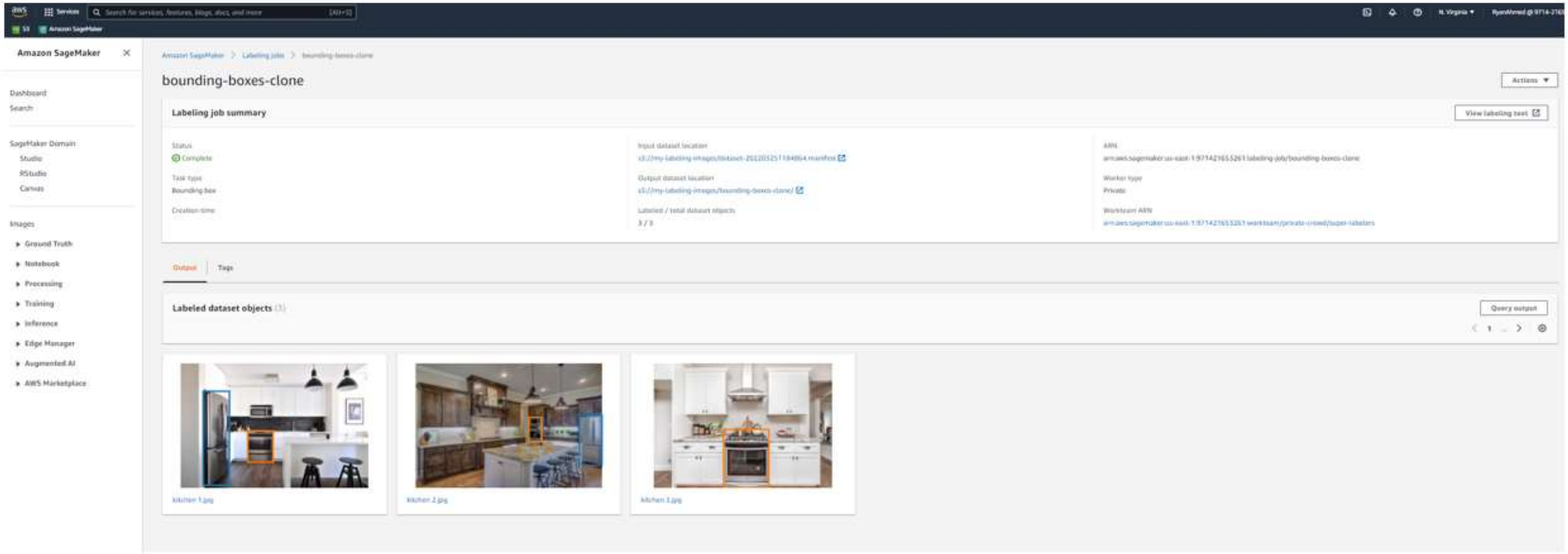
The submitted result data is shown below:

taskAnswers


```
[
  {
    "boundingBox": {
      "boundingBoxes": [
        {
          "height": 400,
          "label": "Oven",
          "left": 510,
          "top": 495,
          "width": 321
        }
      ],
      "inputImageProperties": {
        "height": 896,
        "width": 1280
      }
    }
  }
]
```

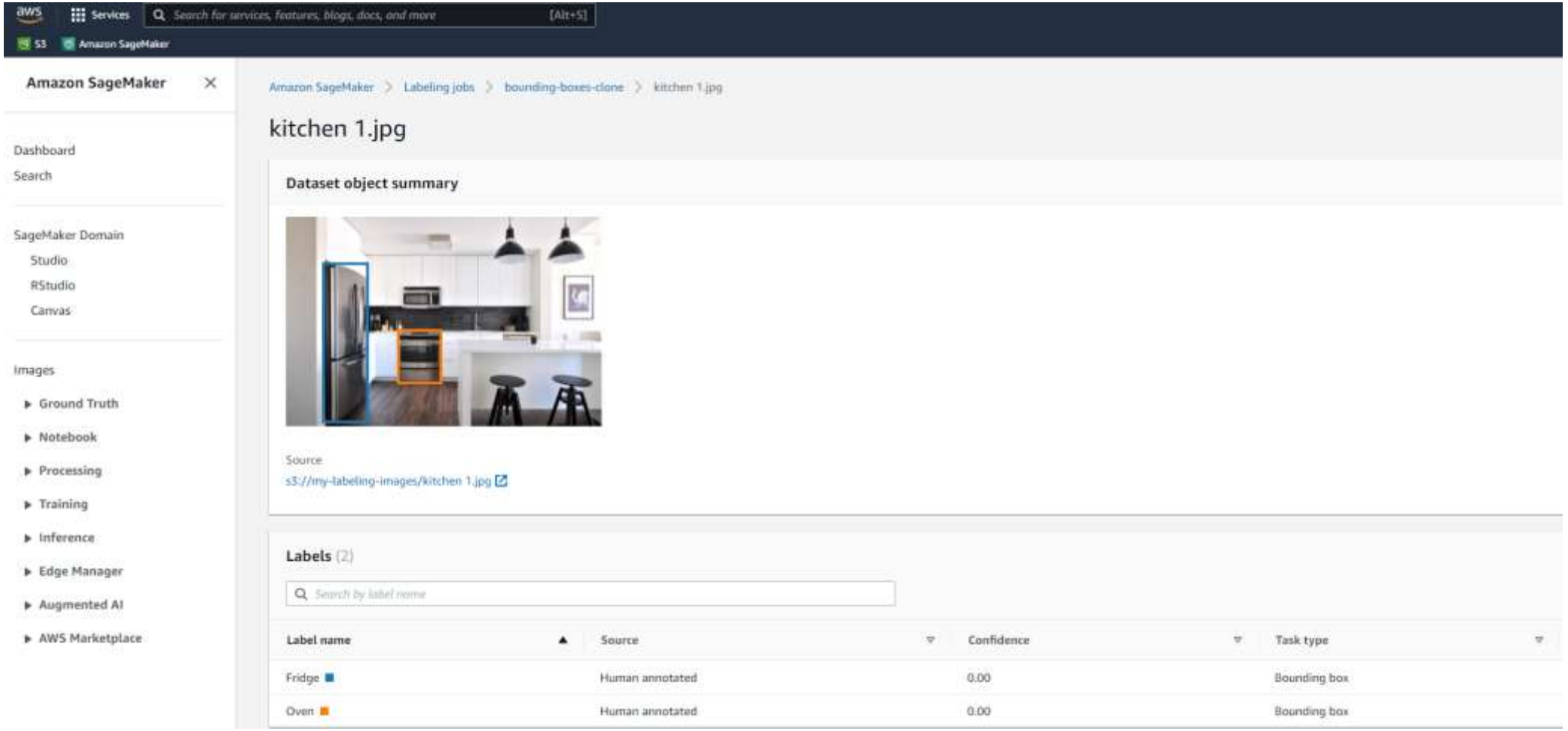
LABELING IMAGES WITH BOUNDING BOXES

BOUNDING BOXES LABELING JOB IS NOW COMPLETE



LABELING IMAGES WITH BOUNDING BOXES

SELECT A SAMPLE IMAGE TO REVIEW THE RESULTING LABEL



LABELING IMAGES WITH BOUNDING BOXES

OUTPUT MANIFEST FILE

output (3).manifest - Notepad

File Edit View


```
{"source-ref":"s3://my-labeling-images/kitchen 1.jpg","bounding-boxes-
clone":{"image_size":[{"width":800,"height":531,"depth":3}],"annotations":[{"class_id":0,"top":119,"left":95,"height":401,"wid
th":112},{"class_id":1,"top":288,"left":285,"height":133,"width":108}]},"bounding-boxes-clone-
metadata":{"objects":[{"confidence":0},{"confidence":0}],"class-map":{"0":"Fridge","1":"Oven"},"type":"groundtruth/object-
detection","human-annotated":"yes","creation-date":
clone"}}
{"source-ref":"s3://my-labeling-images/kitchen 2.jpg","bounding-boxes-
clone":{"image_size":[{"width":5998,"height":3999,"depth":3}],"annotations":[{"class_id":0,"top":1662,"left":5272,"height":160
5,"width":723},{"class_id":1,"top":1734,"left":3639,"height":774,"width":416}]},"bounding-boxes-clone-
metadata":{"objects":[{"confidence":0},{"confidence":0}],"class-map":{"0":"Fridge","1":"Oven"},"type":"groundtruth/object-
detection","human-annotated":"yes"                                                                                    -
clone"}}
{"source-ref":"s3://my-labeling-images/kitchen 3.jpg","bounding-boxes-
clone":{"image_size":[{"width":1280,"height":896,"depth":3}],"annotations":[{"class_id":1,"top":476,"left":501,"height":420,"w
idth":325}]},"bounding-boxes-clone-metadata":{"objects":[{"confidence":0}],"class-
map":{"1":"Oven"},"type":"groundtruth/object-detection","human-annotated":"yes",
                                "job-name":"labeling-job/bounding-boxes-clone"}}
```