AWS AI services

AWS AI services

- Amazon Transcribe
- Amazon Translate
- Amazon Comprehend
- Amazon Polly
- Amazon Lex
- Amazon Rekognition

Amazon Transcribe

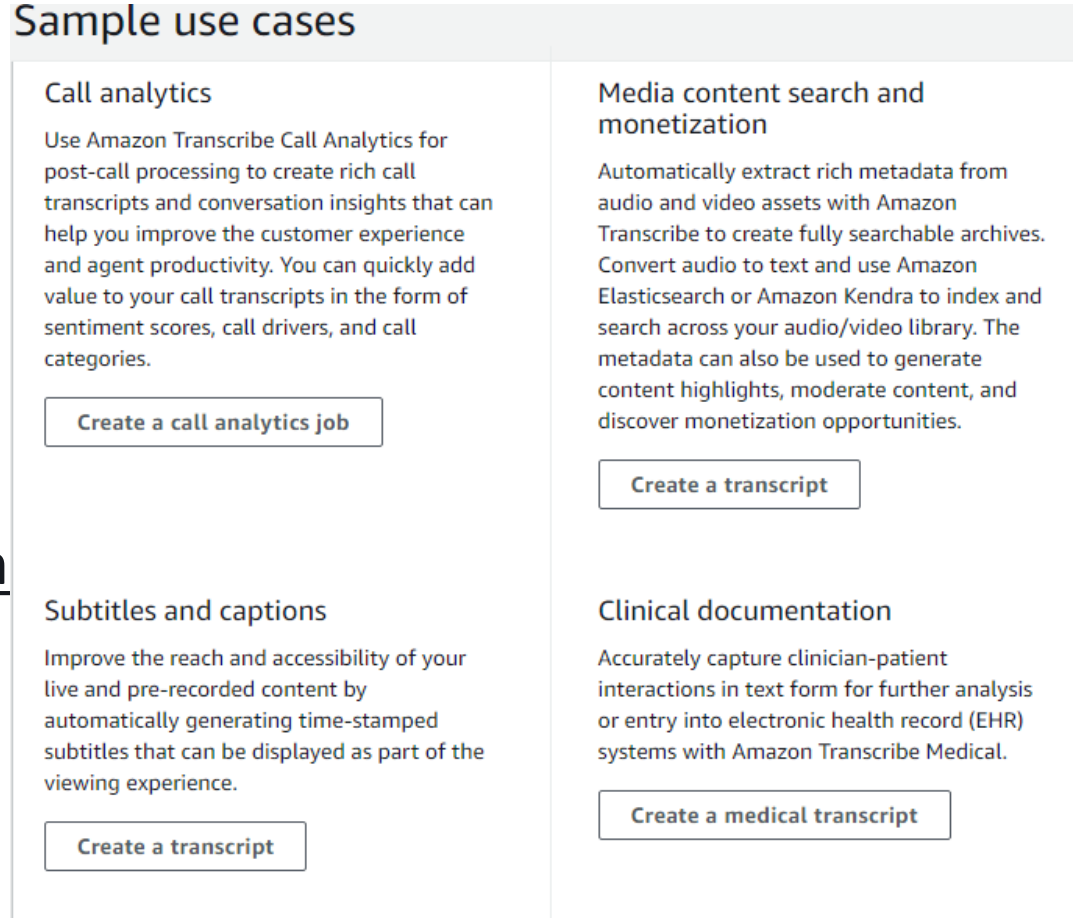
Amazon Transcribe

- **Why**: It is not easy (if possible) to search inside audio data. So, it is better we convert **speech/audio to text**
- Amazon Transcribe is an automatic speech recognition service that uses machine learning models to convert audio to text.
- You can use Amazon Transcribe as a standalone transcription service or to add speech-to-text capabilities to any application.
- Watch this video:

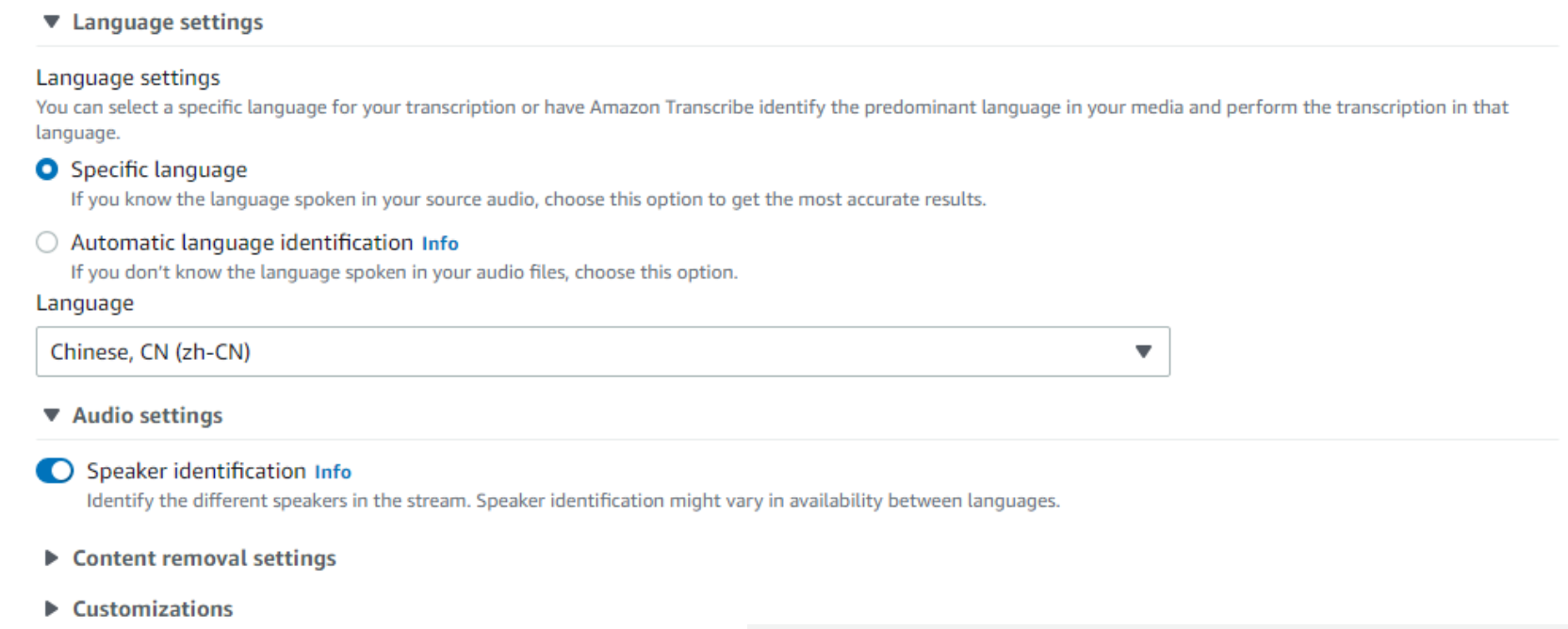
https://youtu.be/zD8NMw4T1TI

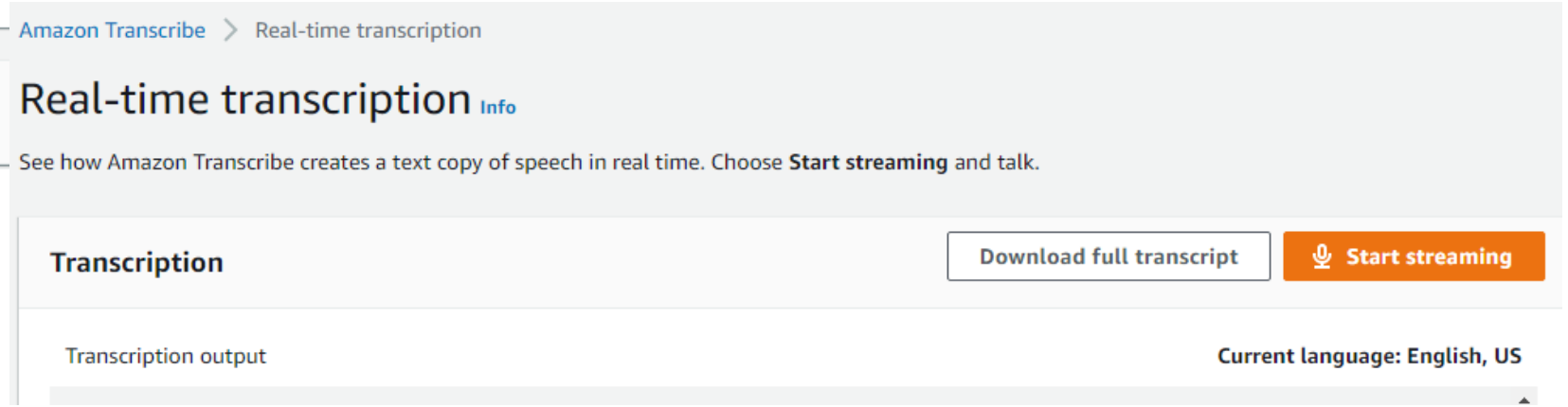
Use cases

- When you have the text, you can use it:
  - For call center post-call processing
  - Extracting rich meta data from audio and video assets
  - For closed caption
  - To accurately capture clinician-patient interaction in text form for further analysis
  - For translation to another language
  - To comprehend the sentiment
  - For searching inside the text
  - …



Demo: Realtime transcription





Demo: let’s try a domain specific words

- Read the first paragraph in this link for transcribe:

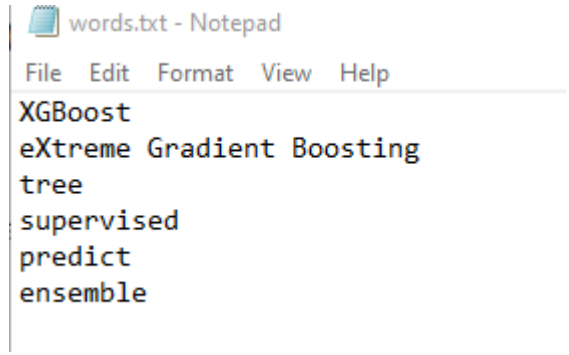
https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost.html

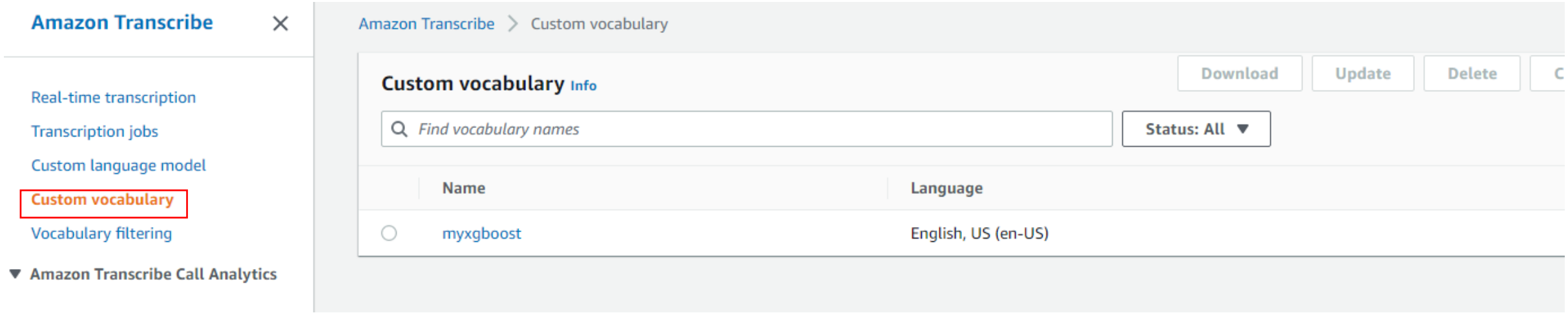
You probably get something like this:

- The exchange boost. Extremely radiant boosting is a popular and efficient open source. Implementation of the gradient boosted three algorithm.
- Gradient boosting is a super voice learning algorithm that attempts to accurately pretty target variable by combining and in some bubble of estimates from a set of simpler and weaker models.

Improving the performance of the Transcribe

- Upload the **words.txt** to Transcribe to custom vocab.





Re-read that XGBoost for Transcribe

- Here is the second-round sample:

The XG boost. Extreme Caribbean boosting is a popular and efficient open source. Implementation of the gradient boosted the algorithm.

Great posting is a supervised learning algorithm that attempts to accurately predicted target valuable by combining and in some below estimates from a set of simpler and bigger models.

Maybe we should add Gradient to the words list, do you agree?

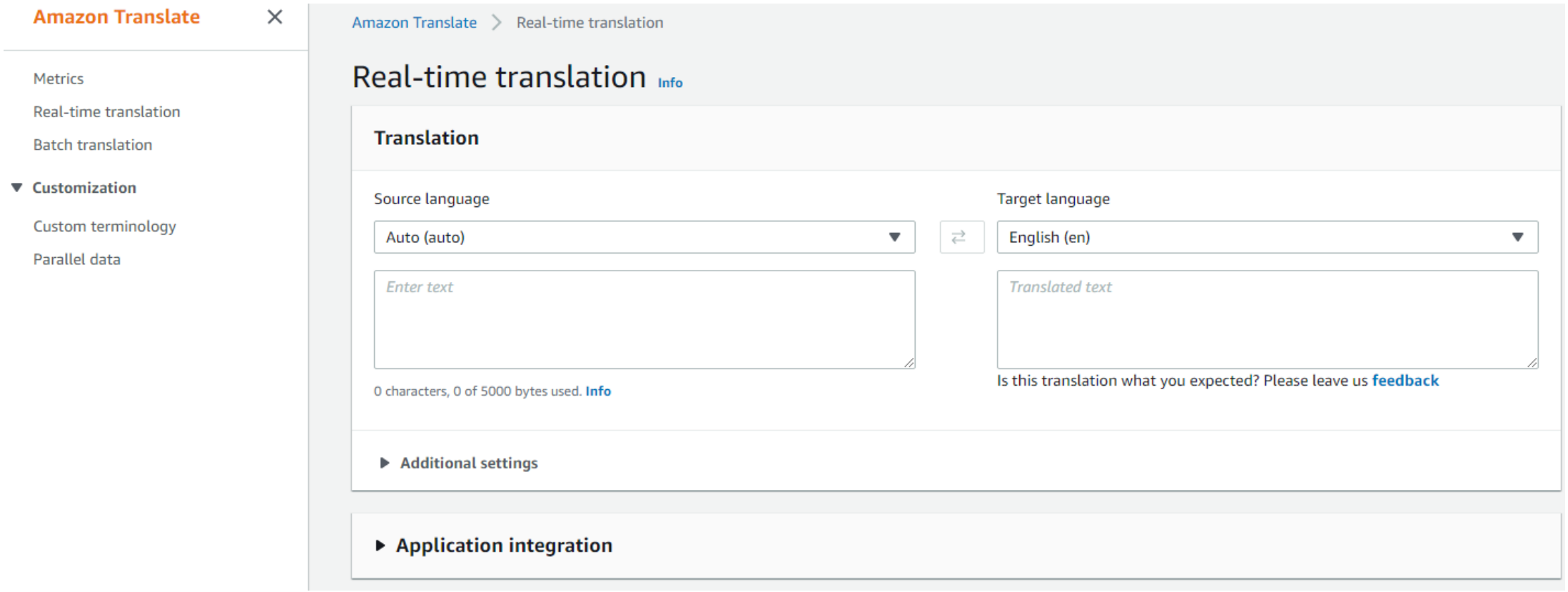
- Try it and show me

Lesson learned

- We still need to work on the output of the Transcribe to make it better
- This is a harder problem and that is the reason we have Deep learning with many complex networks to solve these kinds of problems
- Different accent and people voices makes this service generate wrong content and as a result we need a human supervision to improve the performance of the model

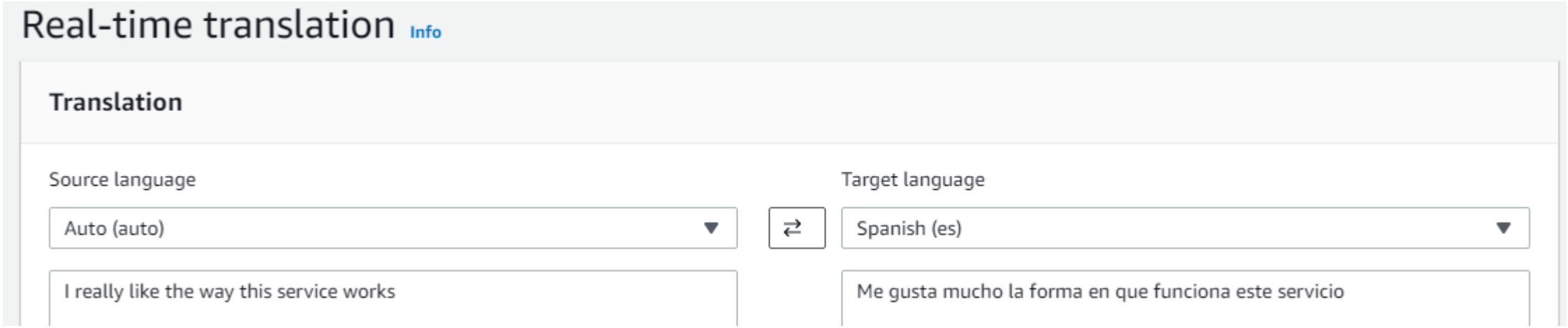
Amazon Translate

Amazon Translate



Demo: Amazon Translate

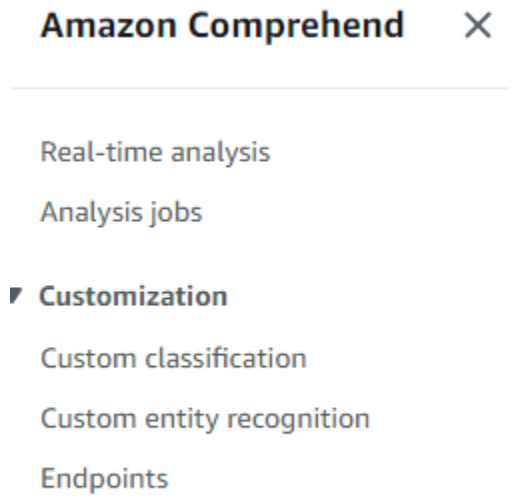
- Let’s try it



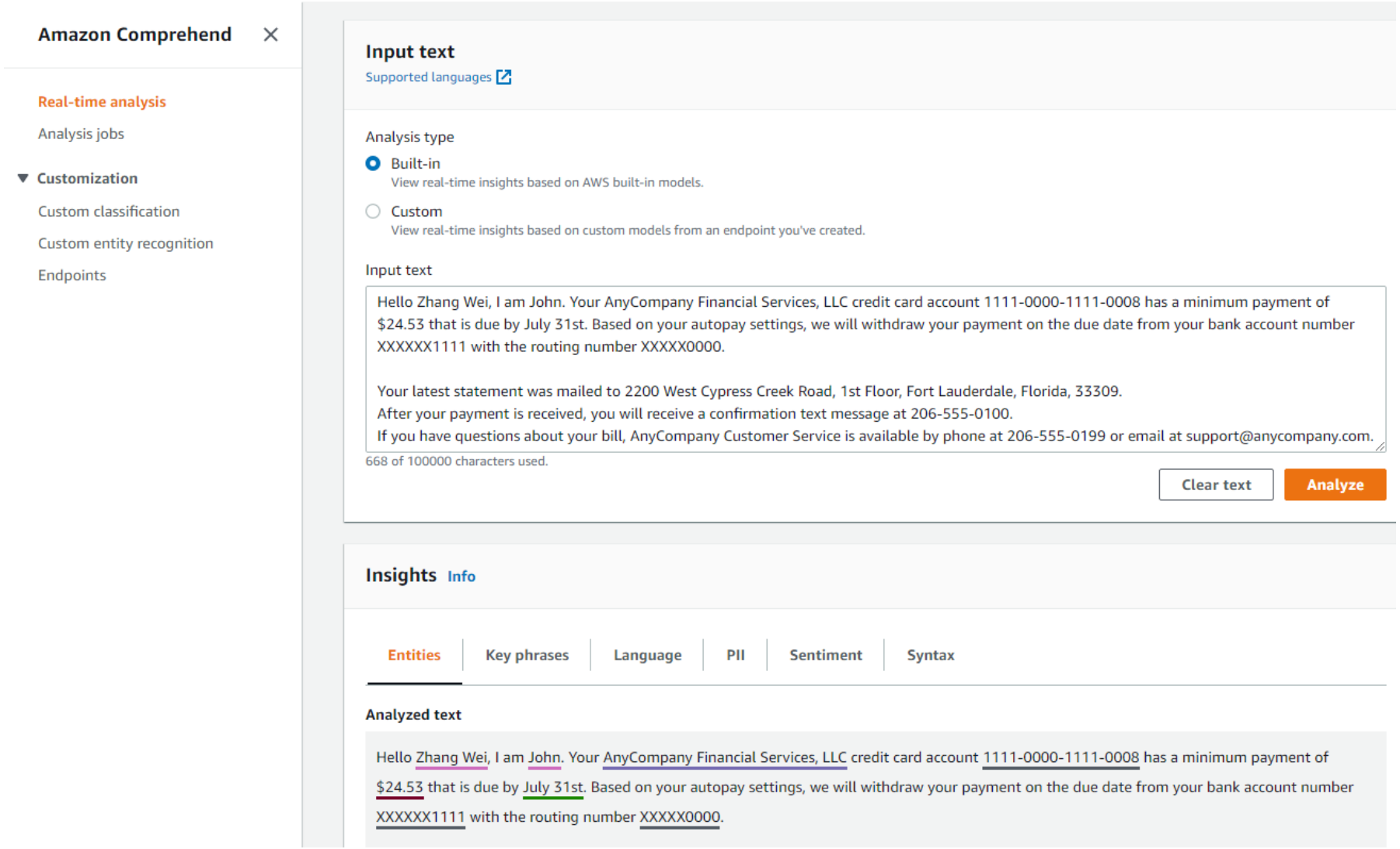
Amazon Comprehend

Amazon Comprehend

- Sentiment Analysis
- Text Classification
- Insights: Entities, Key Phrases, Language, PII, etc

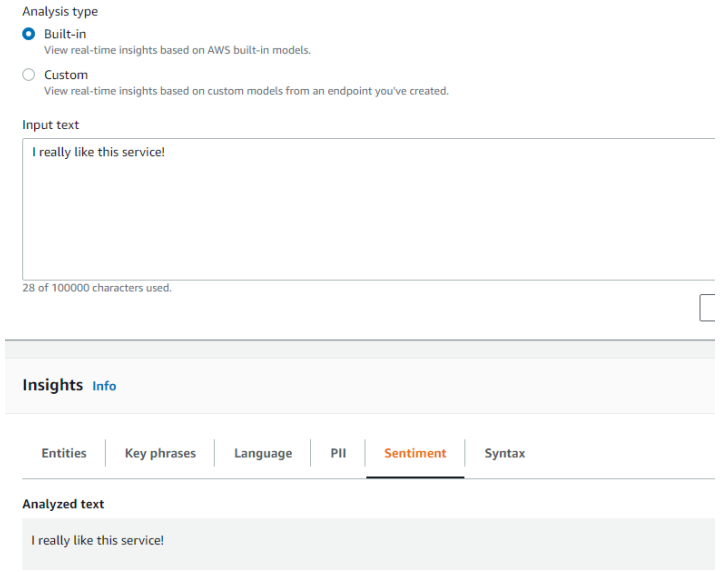


Amazon Comprehend Demo: : Real-Time Analysis



Amazon Comprehend Demo : Sentiment Analysis

- *I really like my cell phone, but I do not want to recommend it!*
- *Nothing, but in general is it fine*
- *What should I say*
- *To my surprise, the feedback was great but awful!*



Amazon Comprehend use cases

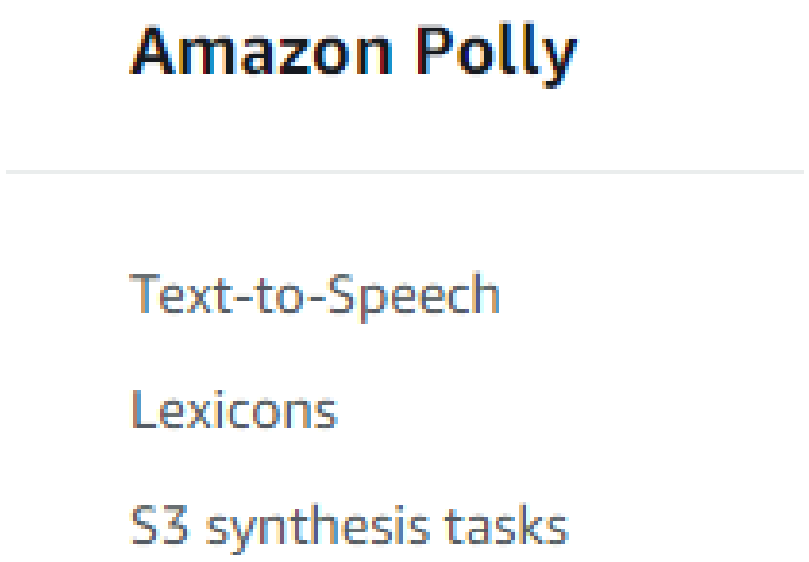
- Here is a list of example to show case other scenarios:

https://aws.amazon.com/comprehend/features/

Based on the above examples, can you describe what Topic Modeling is?

Can you describe the difference between Sentiment Analysis and Targeted Sentiment?

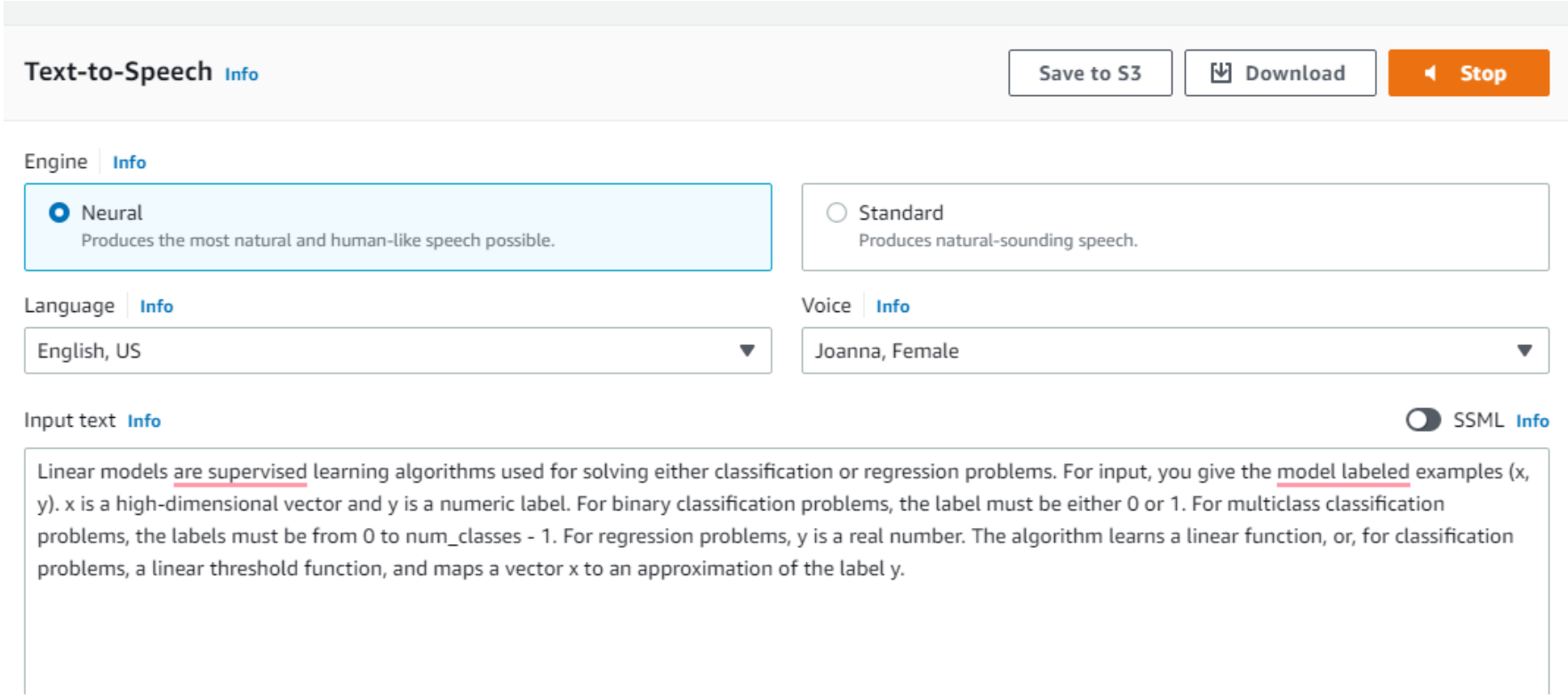
Amazon Polly (Text to Speech)



Amazon Polly Demo

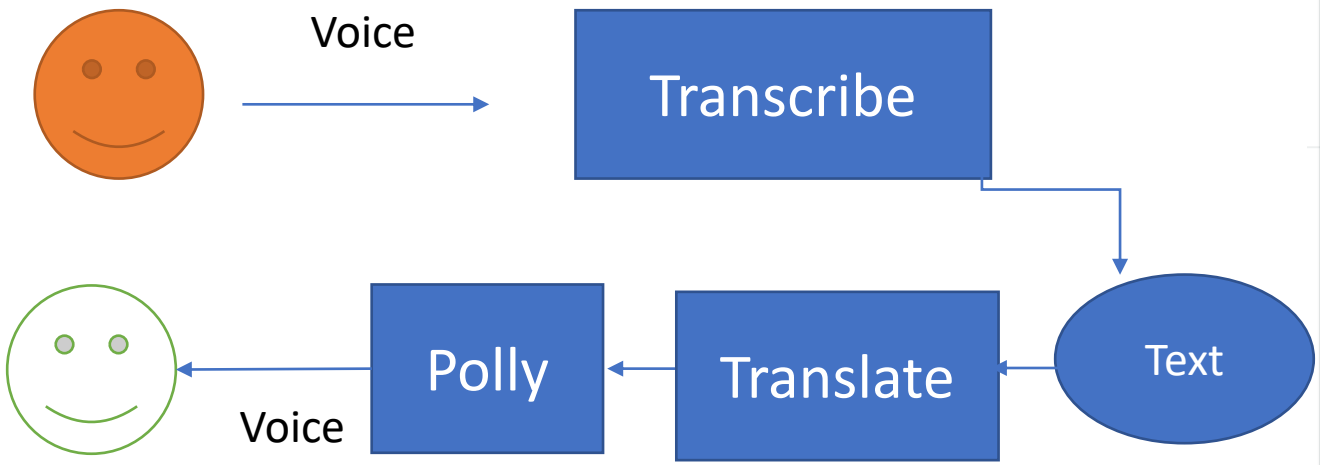
- Let's use Linear Learner text:

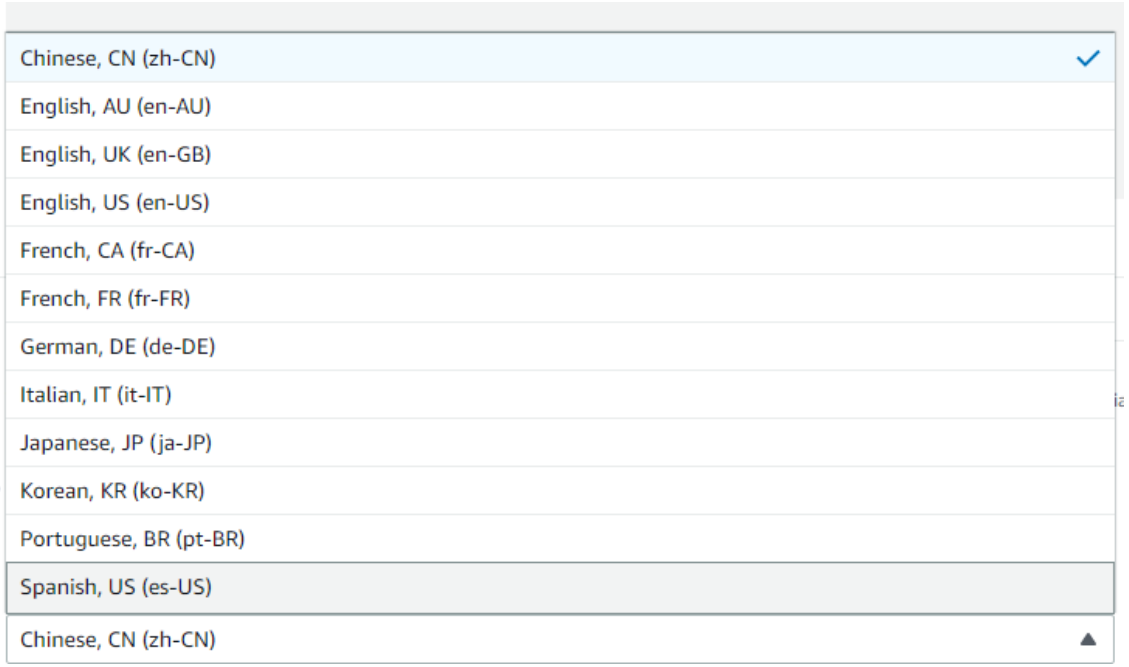
https://docs.aws.amazon.com/sagemaker/latest/dg/linear-learner.html



Show me in class

- If you know any of languages than English that **Transcribe** supports (see below picture), say something in the language. Use the generated text and translate that to English with the Amazon **Translate** and then use the **Polly** to read that for you



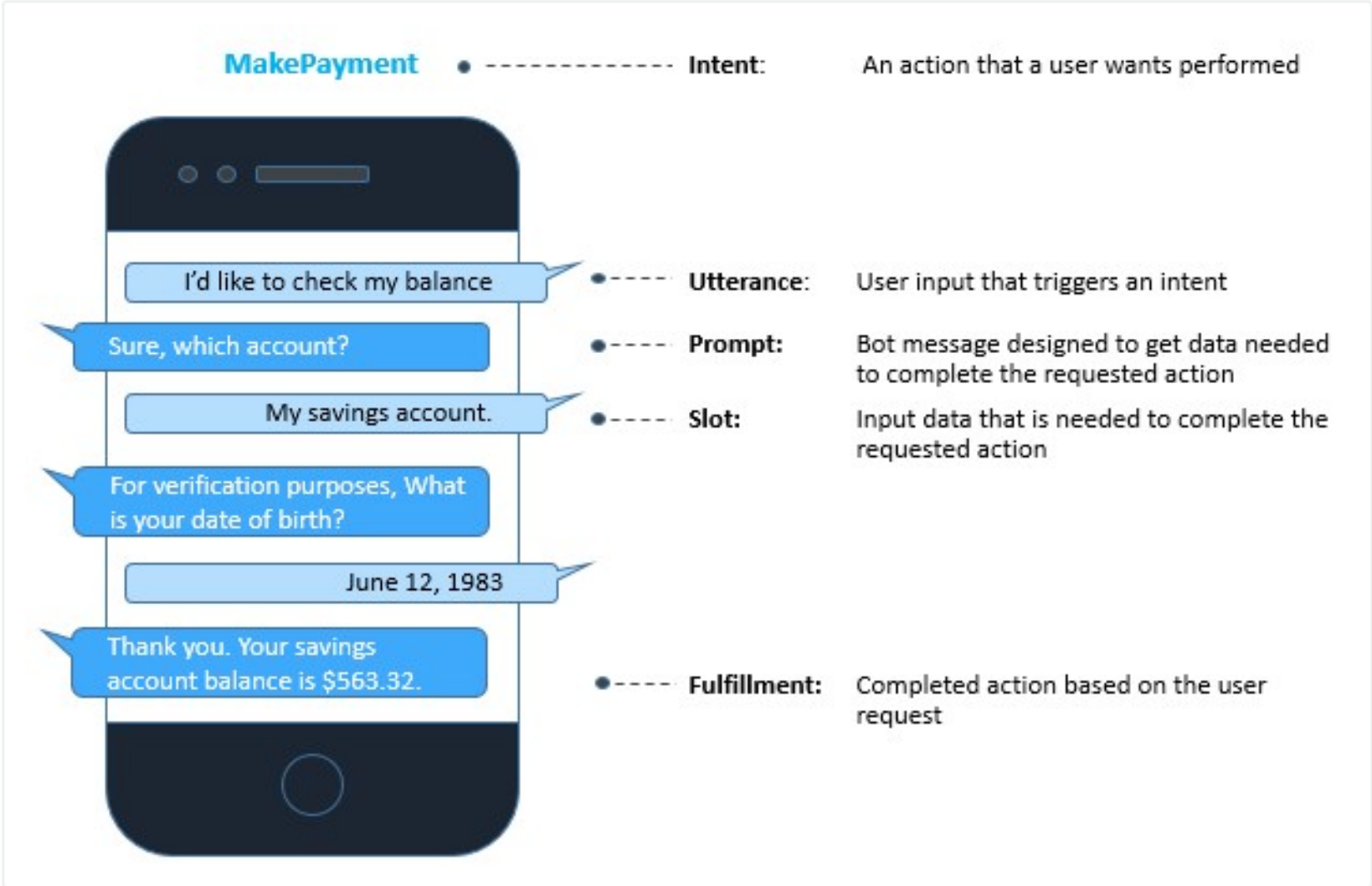


Amazon Lex

Amazon Lex

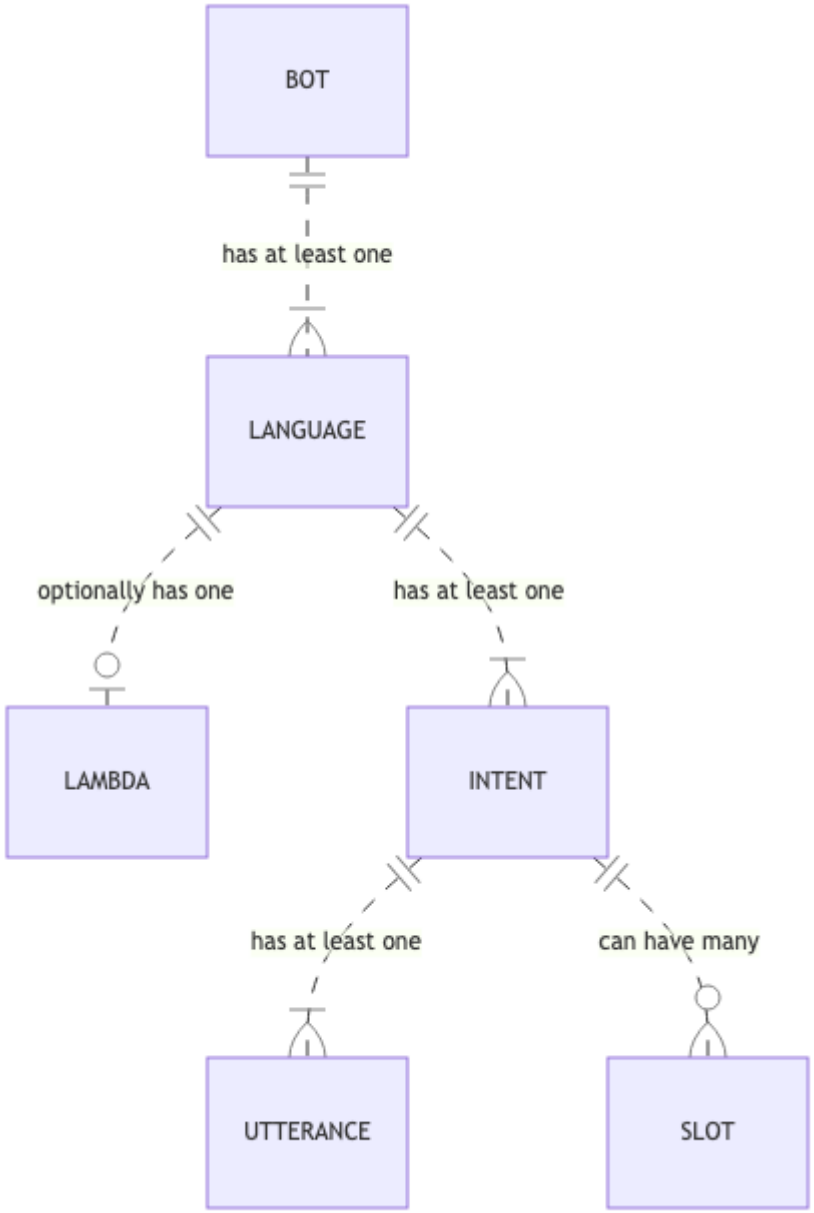
- It is a service that does more than just one of the mentioned AI services, it is a combination of those
- Create a conversational interface using voice and text
- Amazon Alexa
- Lex services:
  - Converting Speech to text
  - NLP to comprehend the intend
  - Running a business logic by Lambda to do further action
  - Return the result in the form of voice or text

Amazon Lex Terminology



Another view:

Intent: is about the reason user is calling this bot



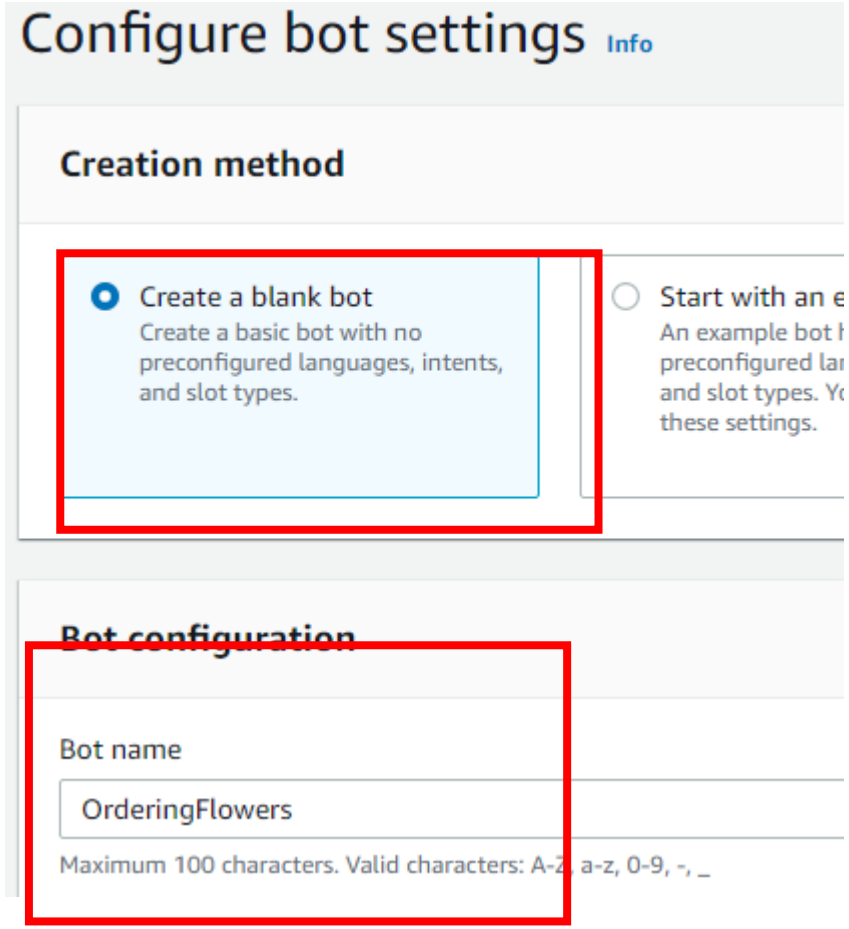
Let’s create a flower order bot

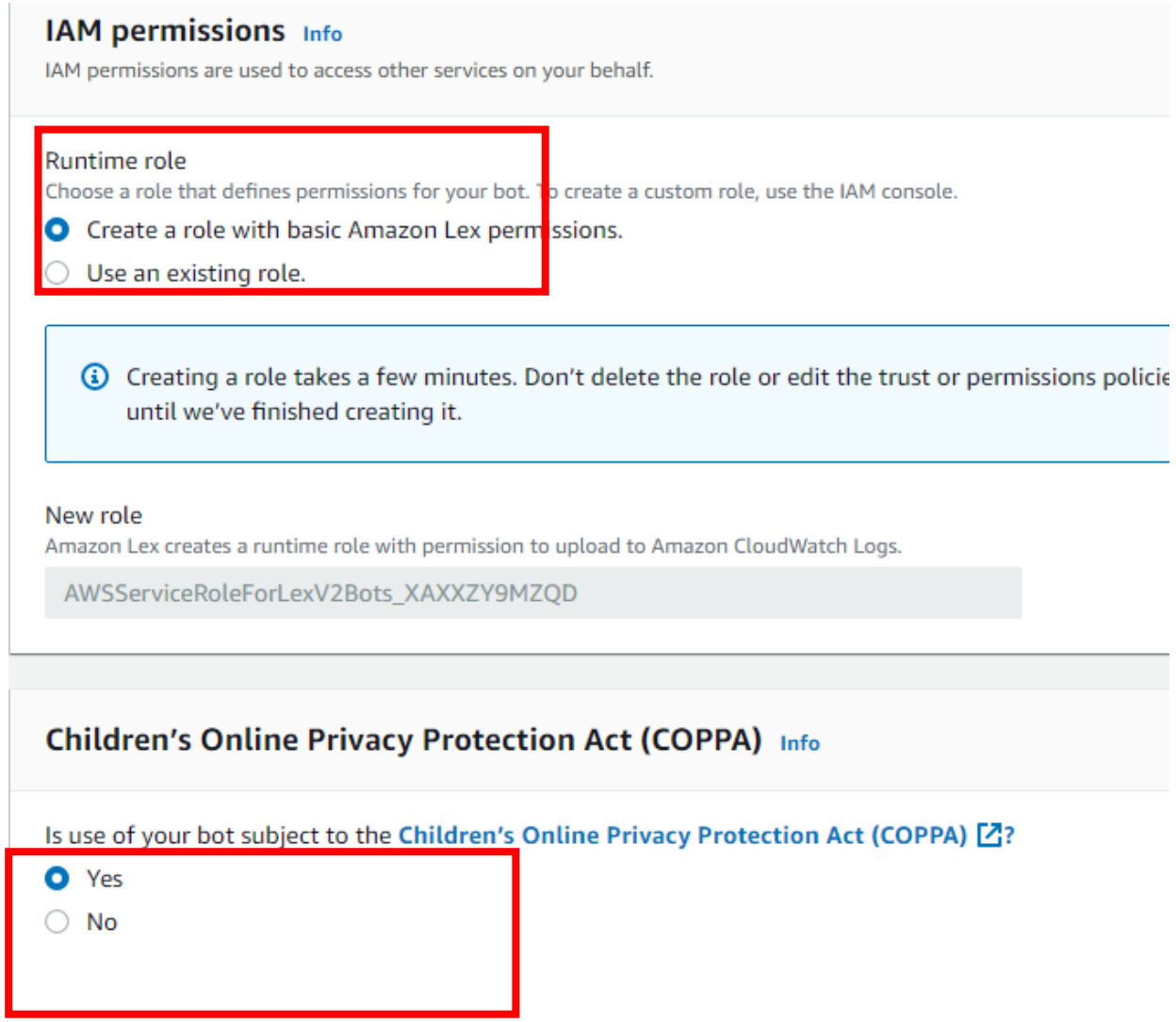
- We follow the instructions in this link: **Create a bot from an example**

https://docs.aws.amazon.com/lexv2/latest/dg/exercise-1.html

NOTE: We use the text messages in the above link but we use a bit different approach in creating the bot, so please follow through next slides and use the above link just to paste the messages

Create a new bot

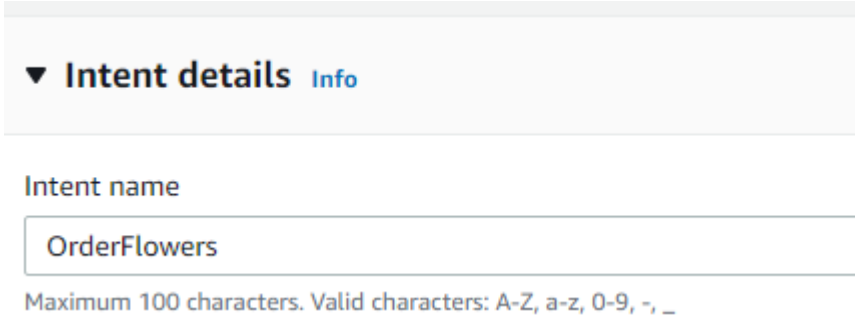


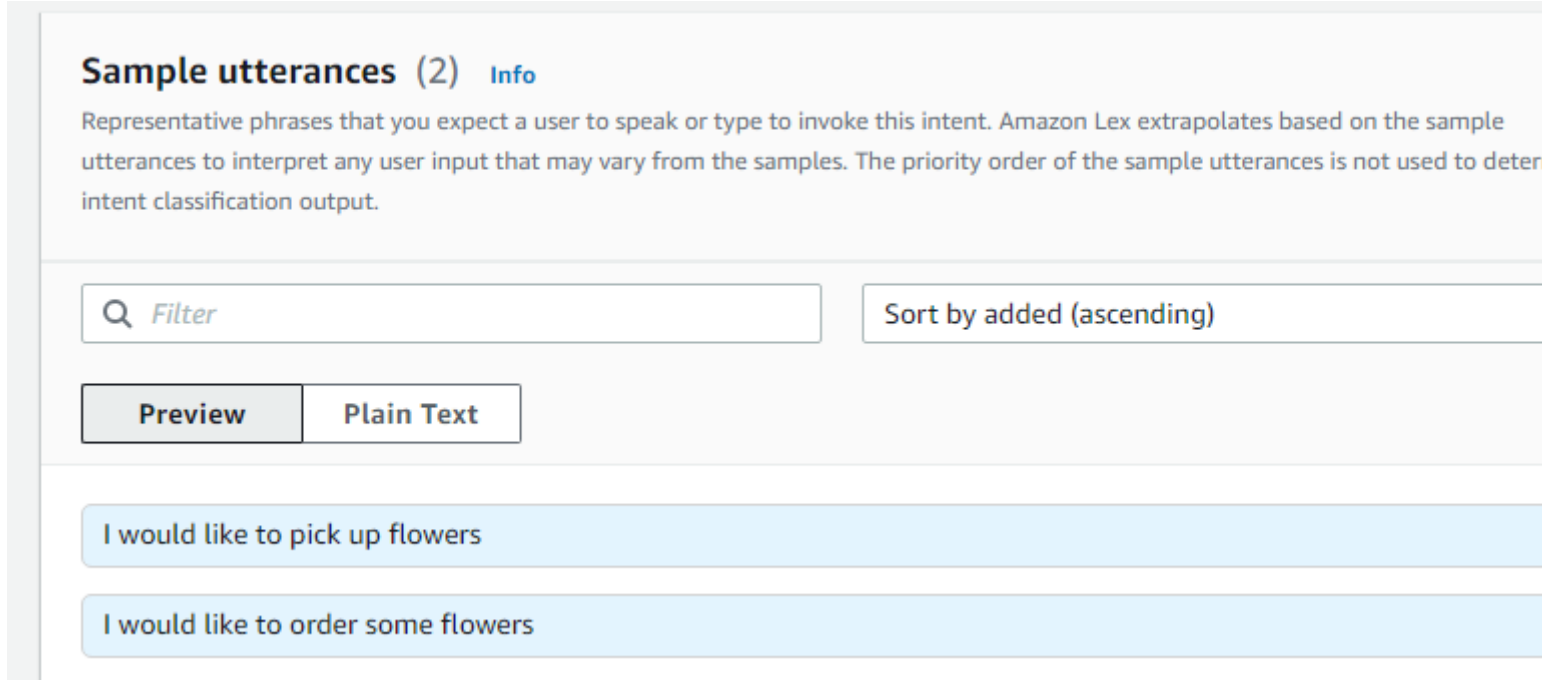


Then click Next and Done

Create the intend and sample utterances

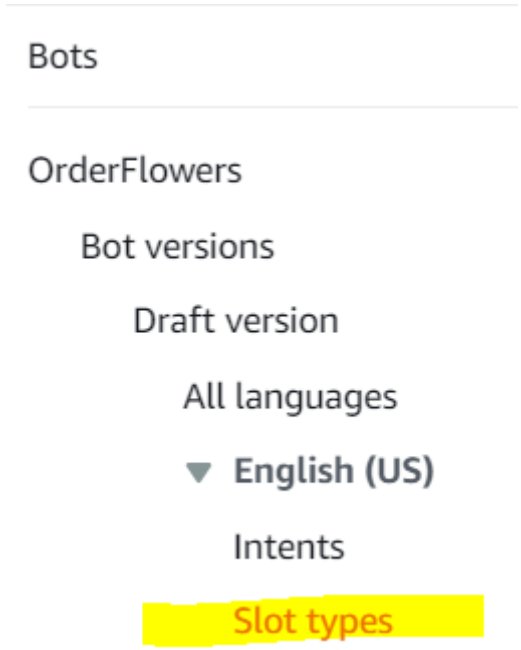
- Here we create the first Intent
- Add the Intent name
- In Sample utterances, add a few utterances
- Click Save Intent



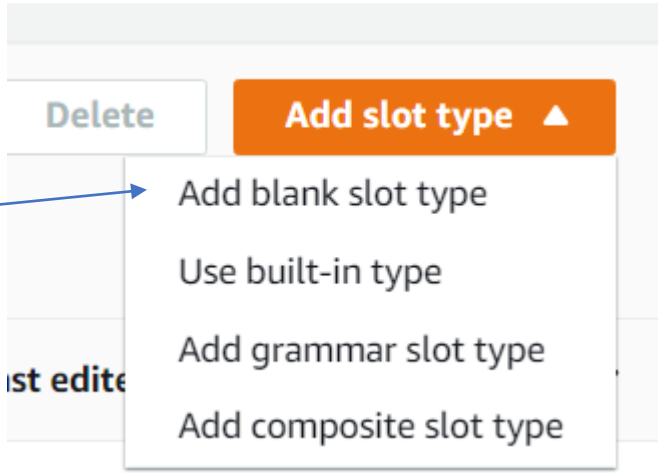


Create a custom slot

- Select Slot Type from the left menu

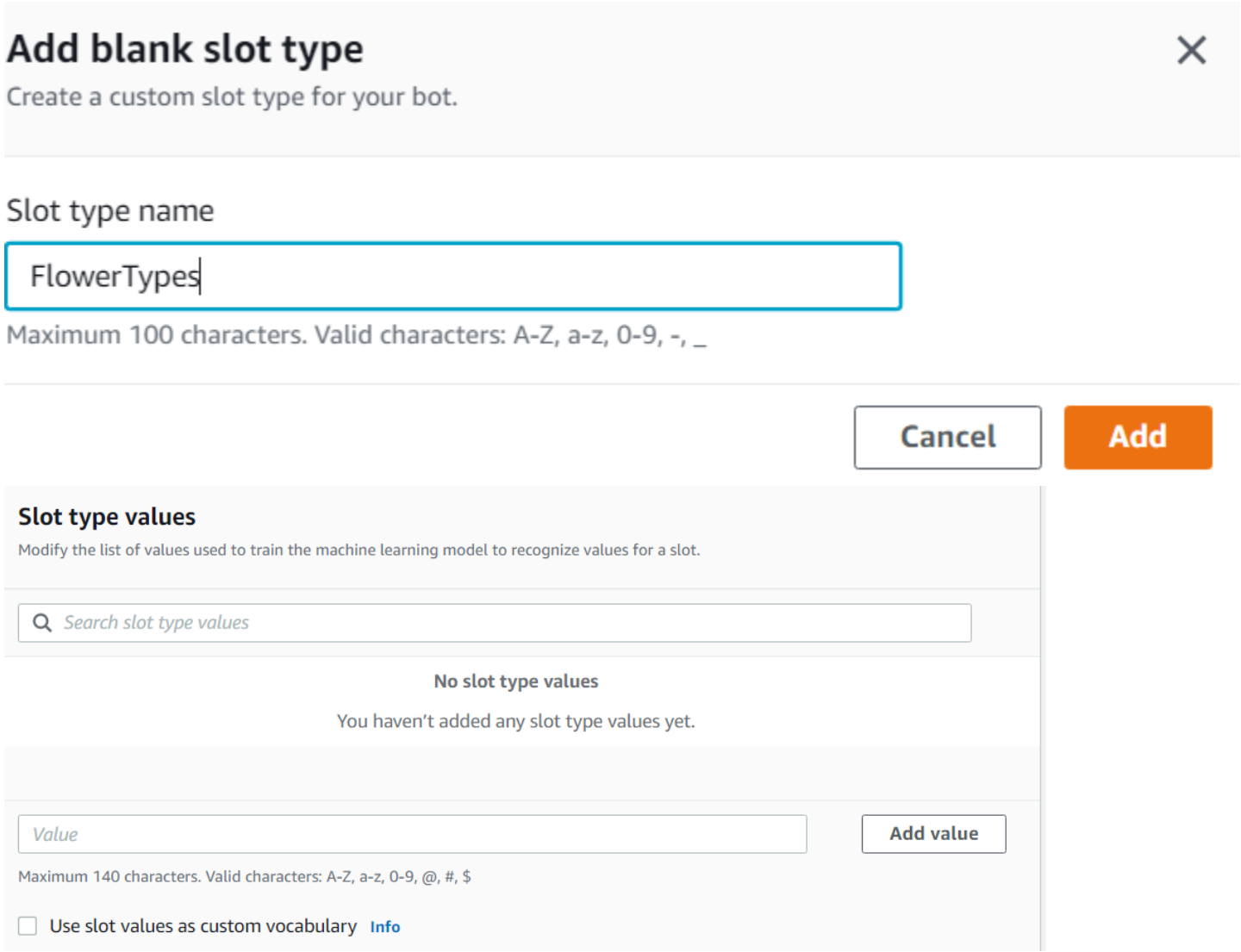


Select **Add blank slot type**

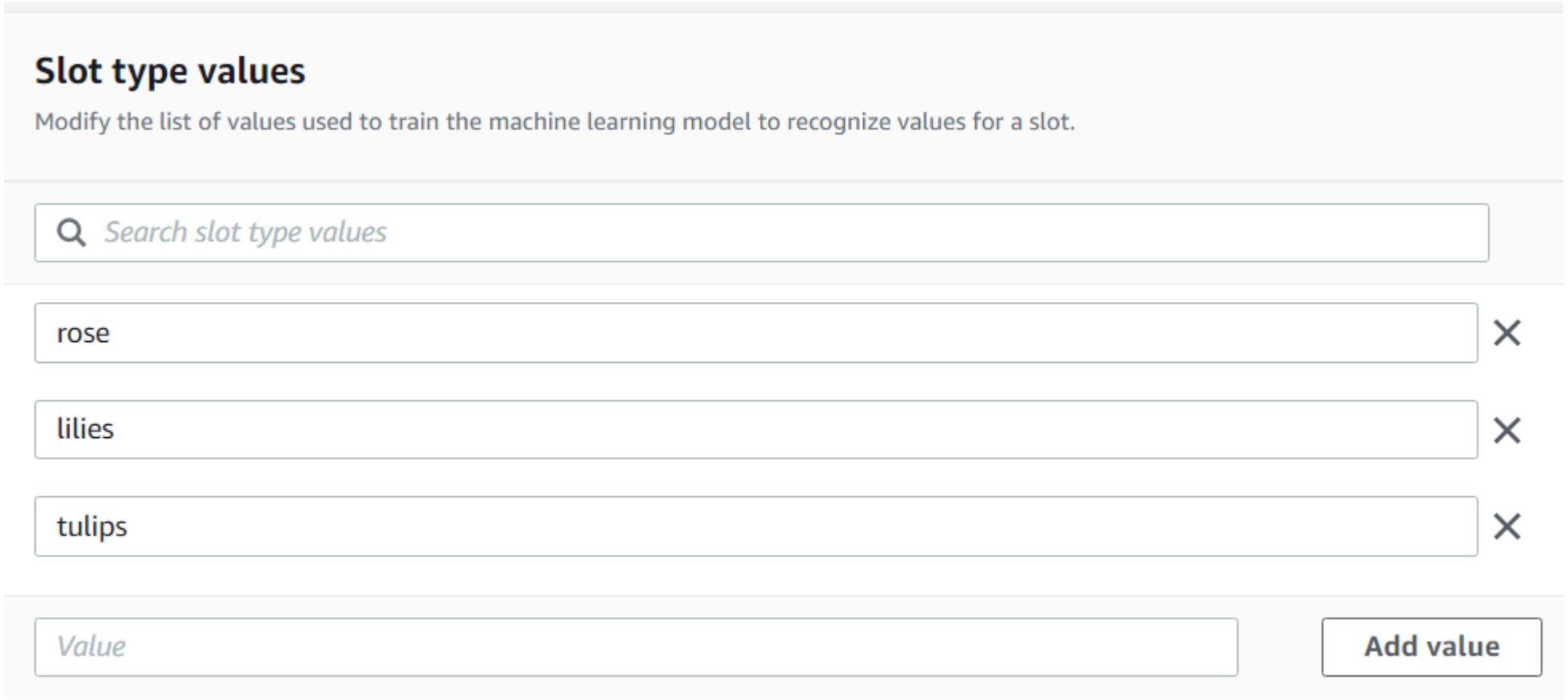


Add a new slot and its values

- Slot Type Name
- Add Values

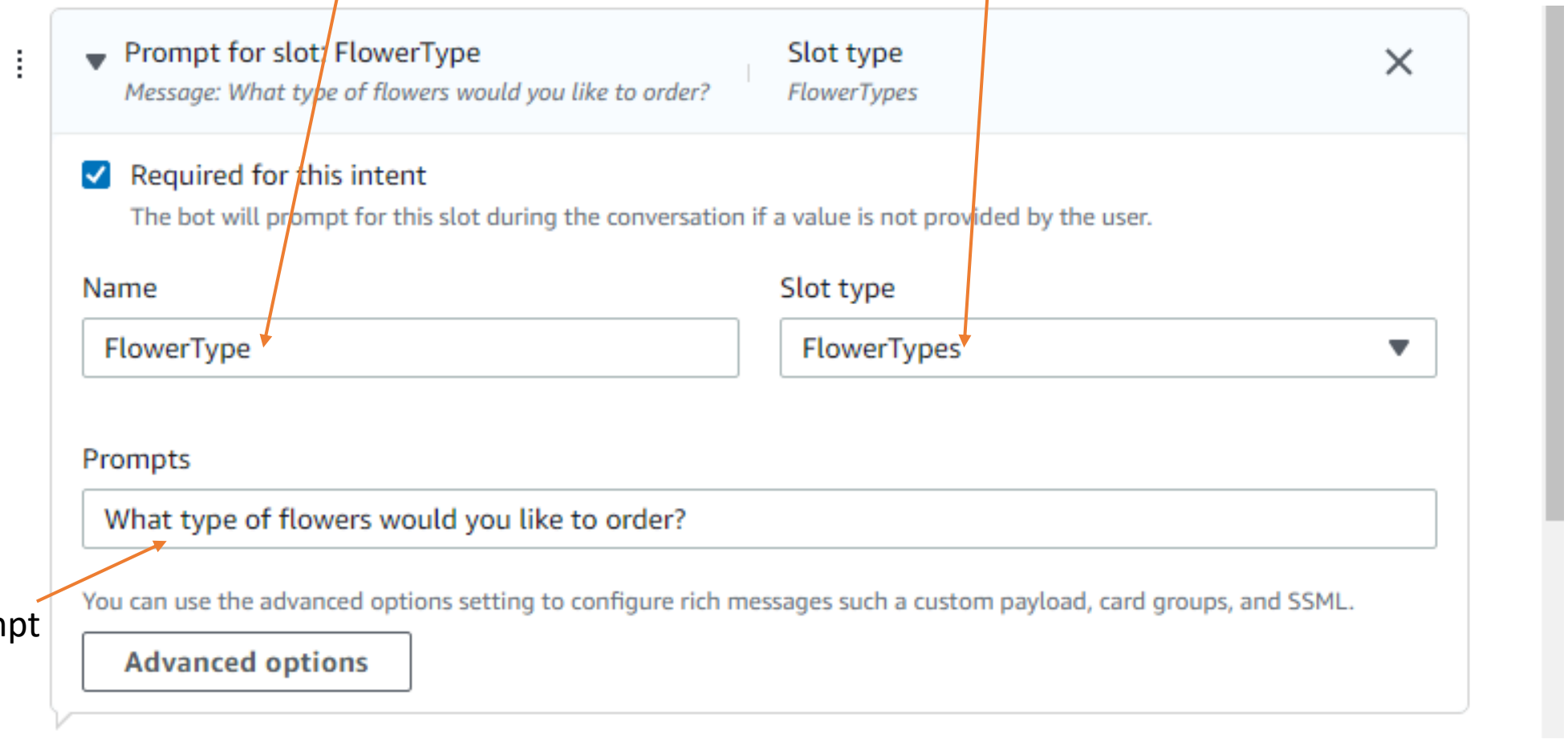


Save the Slot Type after adding the values



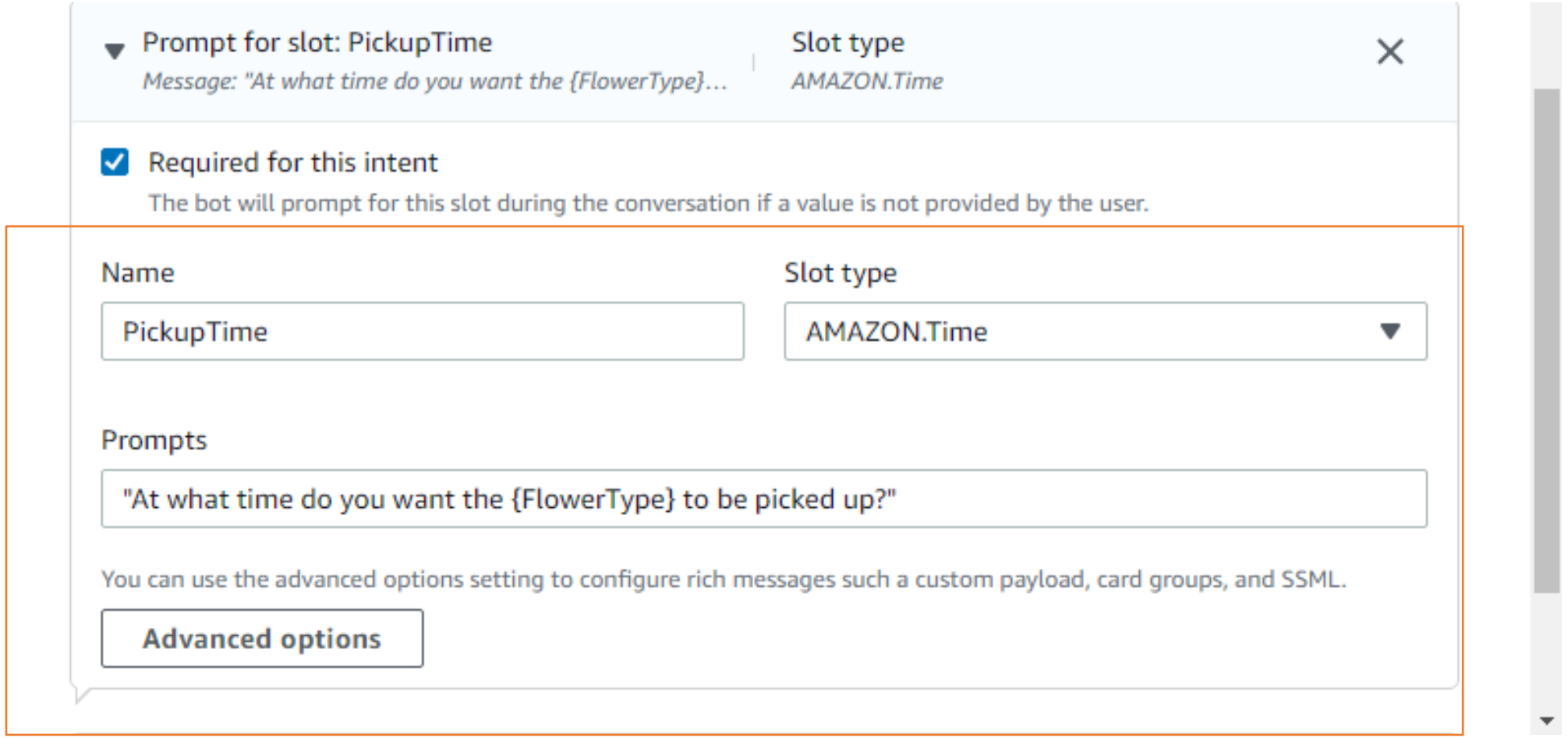
Create the slot FlowerType for the intend

Go back to Intent, select OrderFlowers, Add slot, name it and select the flower types we created in the previous step

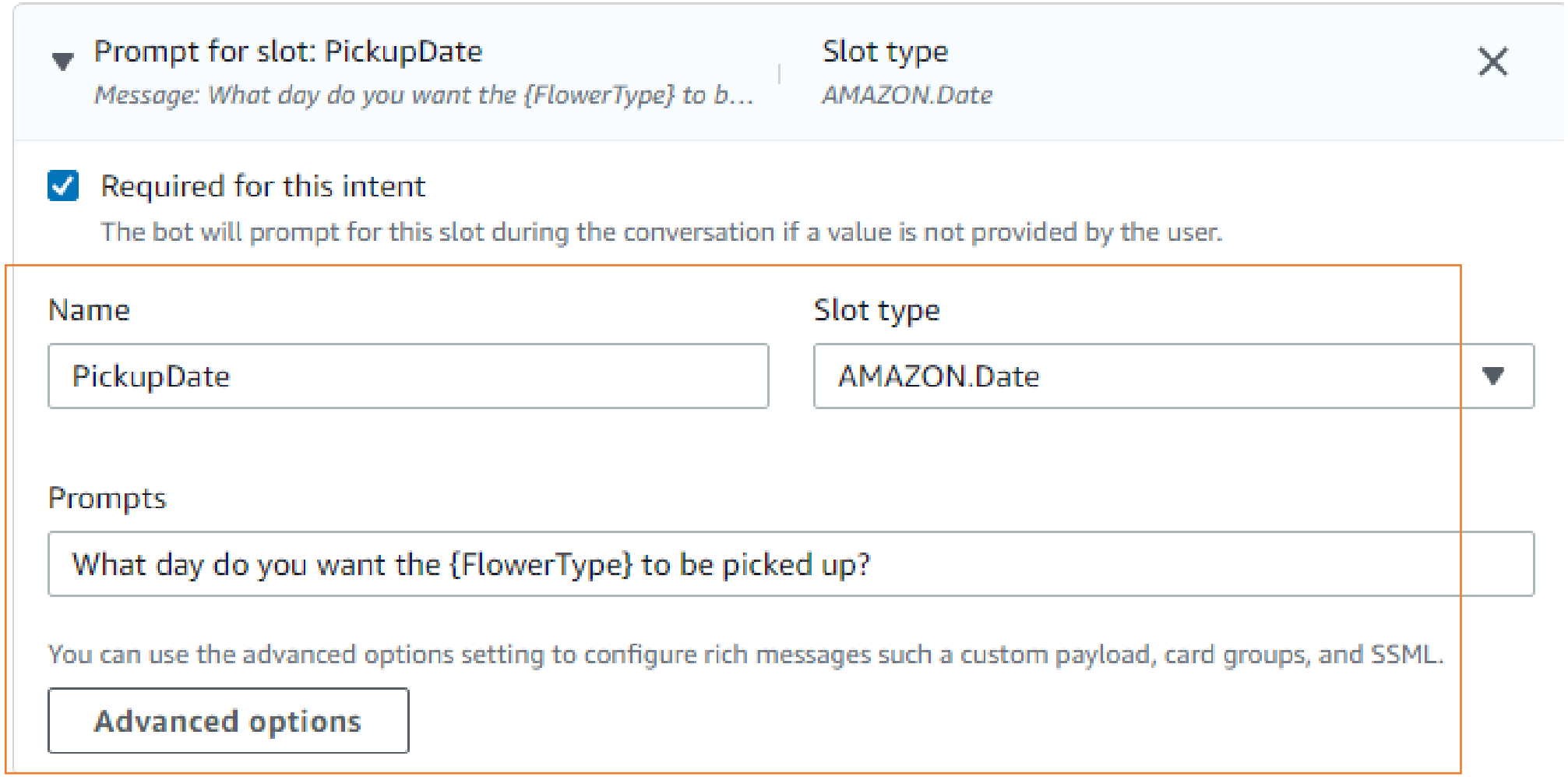


Add the prompt

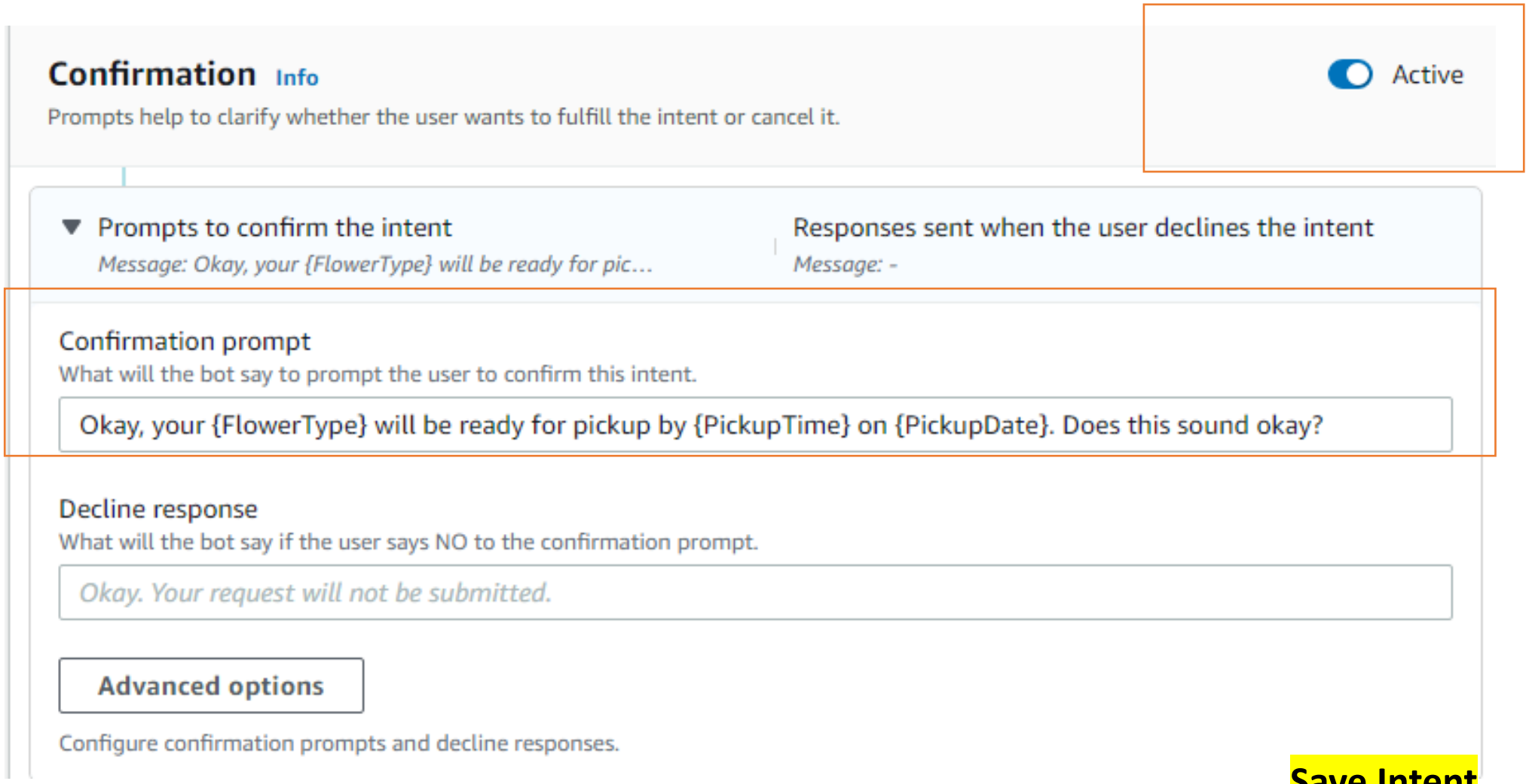
Create a new slot for the intend



Create the slot PickupDate for the intend

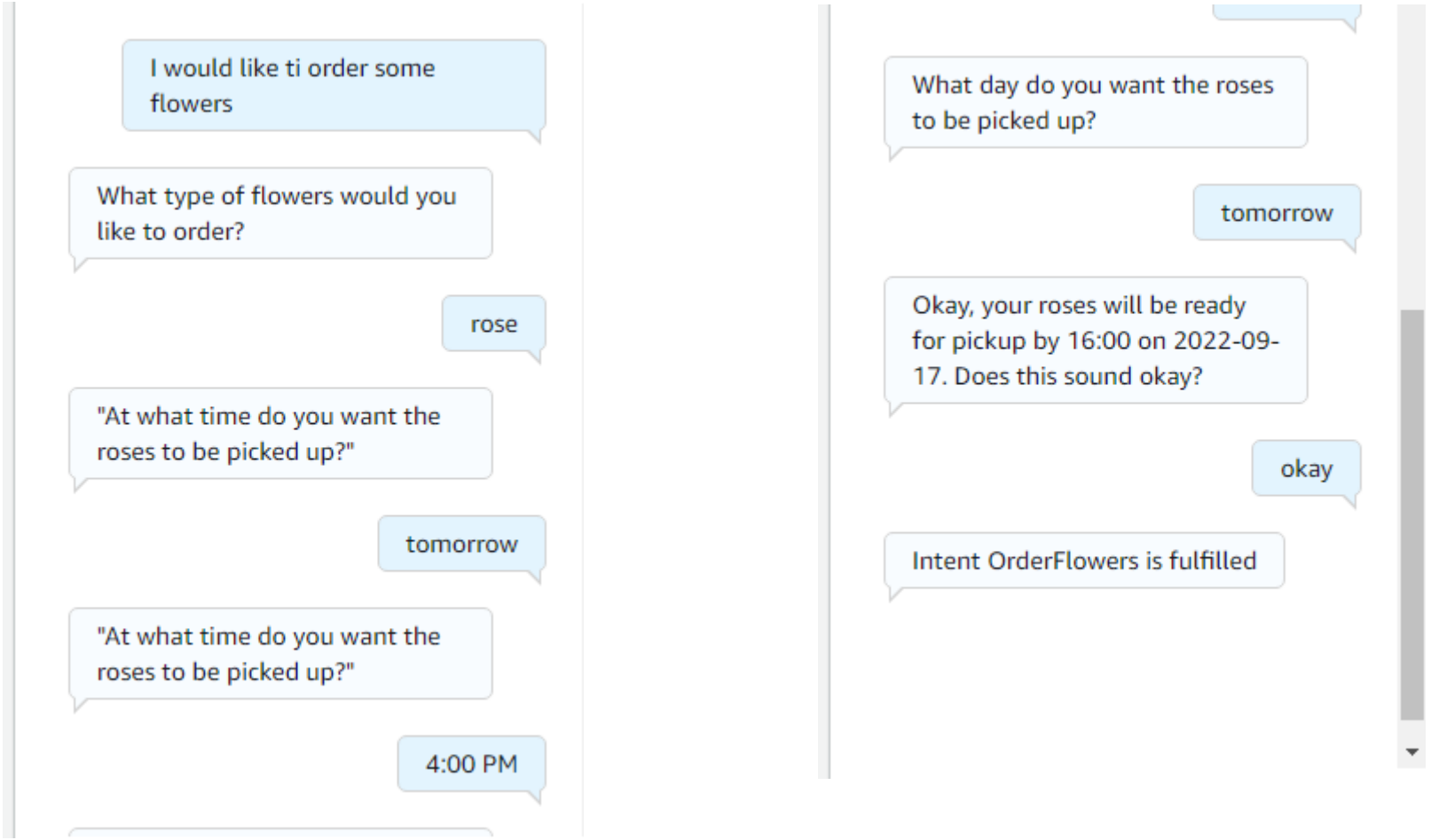


Create Confirmation



Save Intent

Build and Test



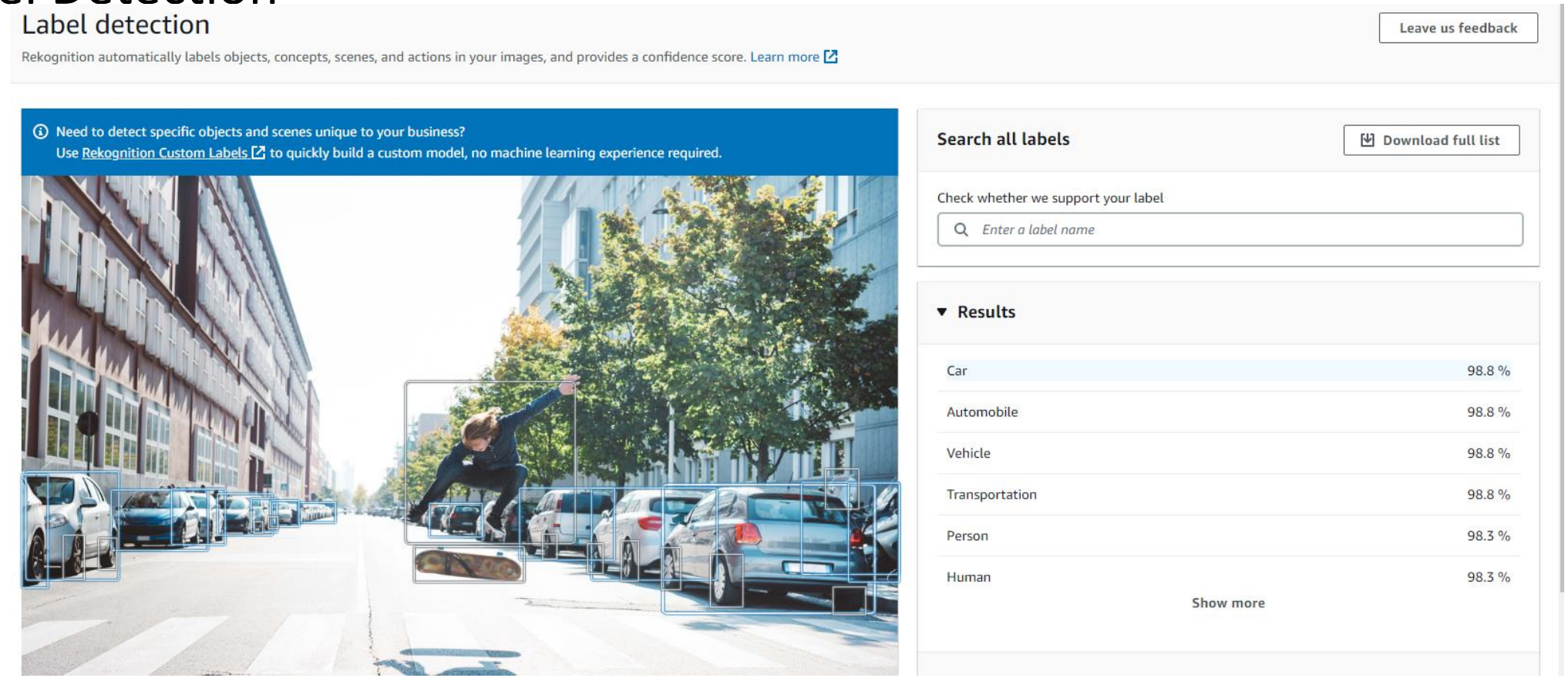
Amazon Rekognition

Amazon Rekognition

- Object (label) detection
- Activity detection
- Scene detection
- Facial recognition
- Facial Analysis
- Pathing (tracking movements)
- Unsafe image detection
- Celebrity recognition
- Recognizing text in image

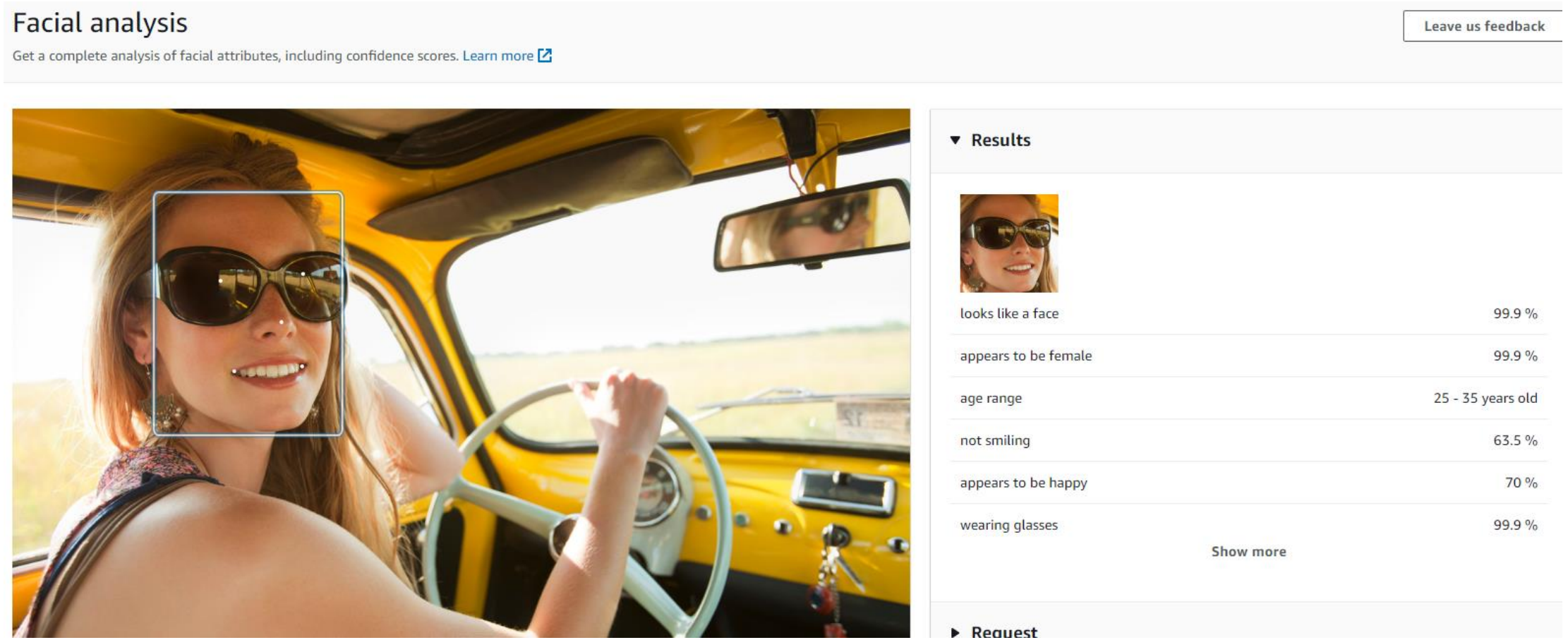
Amazon Rekognition Demo

- Label Detection



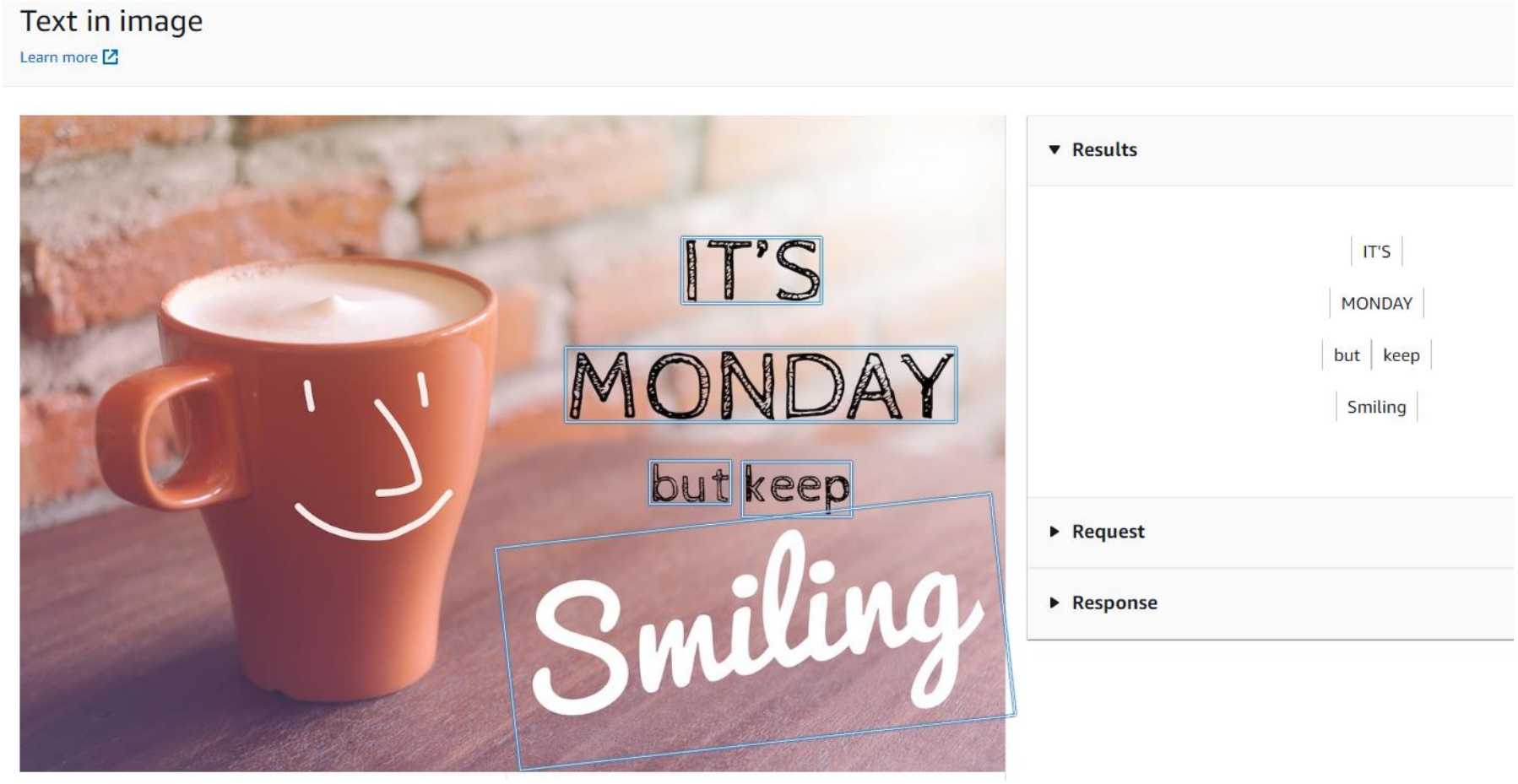
Amazon Rekognition Demo

- Facial analysis



Amazon Rekognition Demo

- Text in image



Assignment

- Use one of the services in this module like translate, transcribe, Rekognition, etc. and upload an appropriate file in S3 (depending on the target AI service, you need to upload different file formats). That upload, triggers a lambda function code and you need to use that event to learn about the location of object in the S3 bucket. After that, you pass that object to the AI service to do the job (translation, or detecting some sentiment or detecting text inside image, whatever you have chosen). The result of that AI service should be either printed in Lambda function or show in cloudwatch or save in S3.