

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
 - ...

Sample use cases

Call analytics

Use Amazon Transcribe Call Analytics for post-call processing to create rich call transcripts and conversation insights that can help you improve the customer experience and agent productivity. You can quickly add value to your call transcripts in the form of sentiment scores, call drivers, and call categories.

Create a call analytics job

Subtitles and captions

Improve the reach and accessibility of your live and pre-recorded content by automatically generating time-stamped subtitles that can be displayed as part of the viewing experience.

Create a transcript

Media content search and monetization

Automatically extract rich metadata from audio and video assets with Amazon Transcribe to create fully searchable archives. Convert audio to text and use Amazon Elasticsearch or Amazon Kendra to index and search across your audio/video library. The metadata can also be used to generate content highlights, moderate content, and discover monetization opportunities.

Create a transcript

Clinical documentation

Accurately capture clinician-patient interactions in text form for further analysis or entry into electronic health record (EHR) systems with Amazon Transcribe Medical.

Create a medical transcript

Demo: Realtime transcription

▼ Language settings

Language settings

You can select a specific language for your transcription or have Amazon Transcribe identify the predominant language in your media and perform the transcription in that language.

☒ Specific language

If you know the language spoken in your source audio, choose this option to get the most accurate results.

☐ Automatic language identification [Info](#)

If you don't know the language spoken in your audio files, choose this option.

Language

Chinese, CN (zh-CN)

▼ Audio settings

☒ Speaker identification [Info](#)

Identify the different speakers in the stream. Speaker identification might vary in availability between languages.

► Content removal settings

► Customizations

[Amazon Transcribe](#) > Real-time transcription

Real-time transcription [Info](#)

See how Amazon Transcribe creates a text copy of speech in real time. Choose **Start streaming** and talk.

Transcription

[Download full transcript](#)

 [Start streaming](#)

Transcription output

Current language: English, US

► Application integration

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**.

words.txt - Notepad

File Edit Format View Help

```
XGBoost
eXtreme Gradient Boosting
tree
supervised
predict
ensemble
```

Amazon Transcribe X

[Real-time transcription](#)
[Transcription jobs](#)
[Custom language model](#)
[Custom vocabulary](#)
[Vocabulary filtering](#)

[Amazon Transcribe Call Analytics](#)

Amazon Transcribe > Custom vocabulary

Custom vocabulary [Info](#)

Download Update Delete C

Status: All ▼

	Name	Language
<input type="radio"/>	myxgboost	English, US (en-US)

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

Amazon Translate ×

[Metrics](#)
[Real-time translation](#)
[Batch translation](#)
▼ Customization
[Custom terminology](#)
[Parallel data](#)

[Amazon Translate](#) > [Real-time translation](#)

Real-time translation [Info](#)

Translation

Source language

Auto (auto) ▼

Target language

English (en) ▼

↔

Enter text

Translated text

0 characters, 0 of 5000 bytes used. [Info](#)

Is this translation what you expected? Please leave us [feedback](#)

▶ Additional settings

▶ Application integration

Demo: Amazon Translate

- Let's try it

Real-time translation [Info](#)

Translation

Source language

Auto (auto) ▼

Target language

Spanish (es) ▼

I really like the way this service works

Me gusta mucho la forma en que funciona este servicio

Amazon Comprehend

Amazon Comprehend

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

Amazon Comprehend ×

Real-time analysis

Analysis jobs

▾ Customization

Custom classification

Custom entity recognition

Endpoints

Amazon Comprehend Demo: : Real-Time Analysis

Amazon Comprehend

Real-time analysis

Analysis jobs

Customization

Custom classification

Custom entity recognition

Endpoints

Input text

Supported languages

Analysis type

Built-in

View real-time insights based on AWS built-in models.

Custom

View real-time insights based on custom models from an endpoint you've created.

Input text

Hello Zhang Wei, I am John. Your AnyCompany Financial Services, LLC credit card account 1111-0000-1111-0008 has a minimum payment of \$24.53 that is due by July 31st. Based on your autopay settings, we will withdraw your payment on the due date from your bank account number XXXXXX1111 with the routing number XXXXX0000.

Your latest statement was mailed to 2200 West Cypress Creek Road, 1st Floor, Fort Lauderdale, Florida, 33309. After your payment is received, you will receive a confirmation text message at 206-555-0100. If you have questions about your bill, AnyCompany Customer Service is available by phone at 206-555-0199 or email at support@anycompany.com.

668 of 100000 characters used.

Clear text

Analyze

Insights

Info

Entities

Key phrases

Language

PII

Sentiment

Syntax

Analyzed text

Hello Zhang Wei, I am John. Your AnyCompany Financial Services, LLC credit card account 1111-0000-1111-0008 has a minimum payment of \$24.53 that is due by July 31st. Based on your autopay settings, we will withdraw your payment on the due date from your bank account number XXXXXX1111 with the routing number XXXXX0000.

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!*

Analysis type

☒ Built-in
View real-time insights based on AWS built-in models.

☐ Custom
View real-time insights based on custom models from an endpoint you've created.

Input text

I really like this service!

28 of 100000 characters used.

Insights [Info](#)

Entities | Key phrases | Language | PII | **Sentiment** | Syntax

Analyzed text

I really like this service!

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

Text-to-Speech

Lexicons

S3 synthesis tasks

Amazon Polly Demo

- Let's use Linear Learner text:

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

Text-to-Speech [Info](#)

Save to S3

Download

Stop

Engine [Info](#)

☒ Neural
Produces the most natural and human-like speech possible.

☐ Standard
Produces natural-sounding speech.

Language [Info](#)

English, US ▼

Voice [Info](#)

Joanna, Female ▼

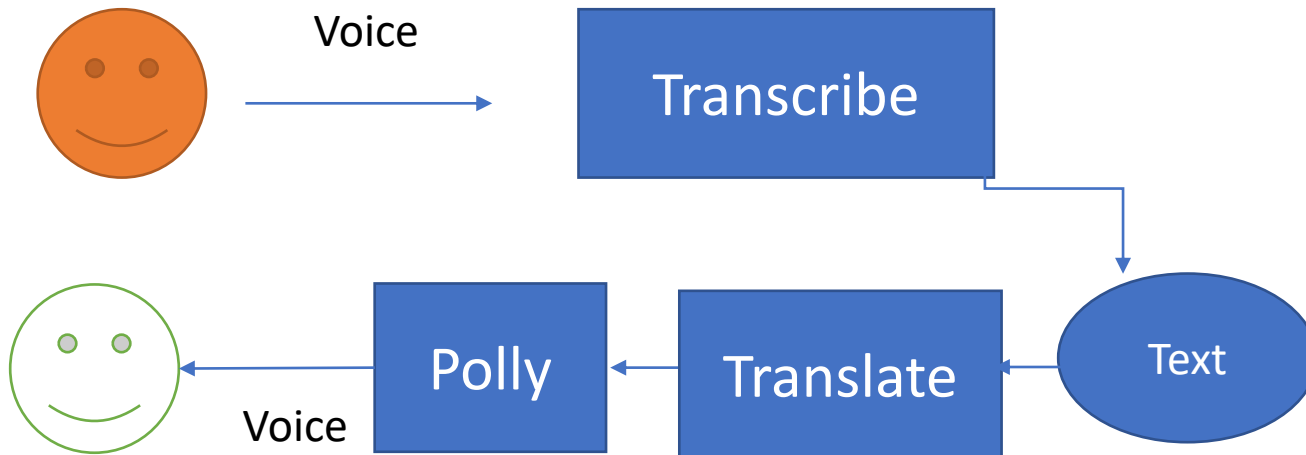
Input text [Info](#)

☐ SSML [Info](#)

Linear models are supervised learning algorithms used for solving either classification or regression problems. For input, you give the model labeled examples (x, y). x is a high-dimensional vector and y is a numeric label. For binary classification problems, the label must be either 0 or 1. For multiclass classification problems, the labels must be from 0 to num_classes - 1. For regression problems, y is a real number. The algorithm learns a linear function, or, for classification problems, a linear threshold function, and maps a vector x to an approximation of the label y.

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



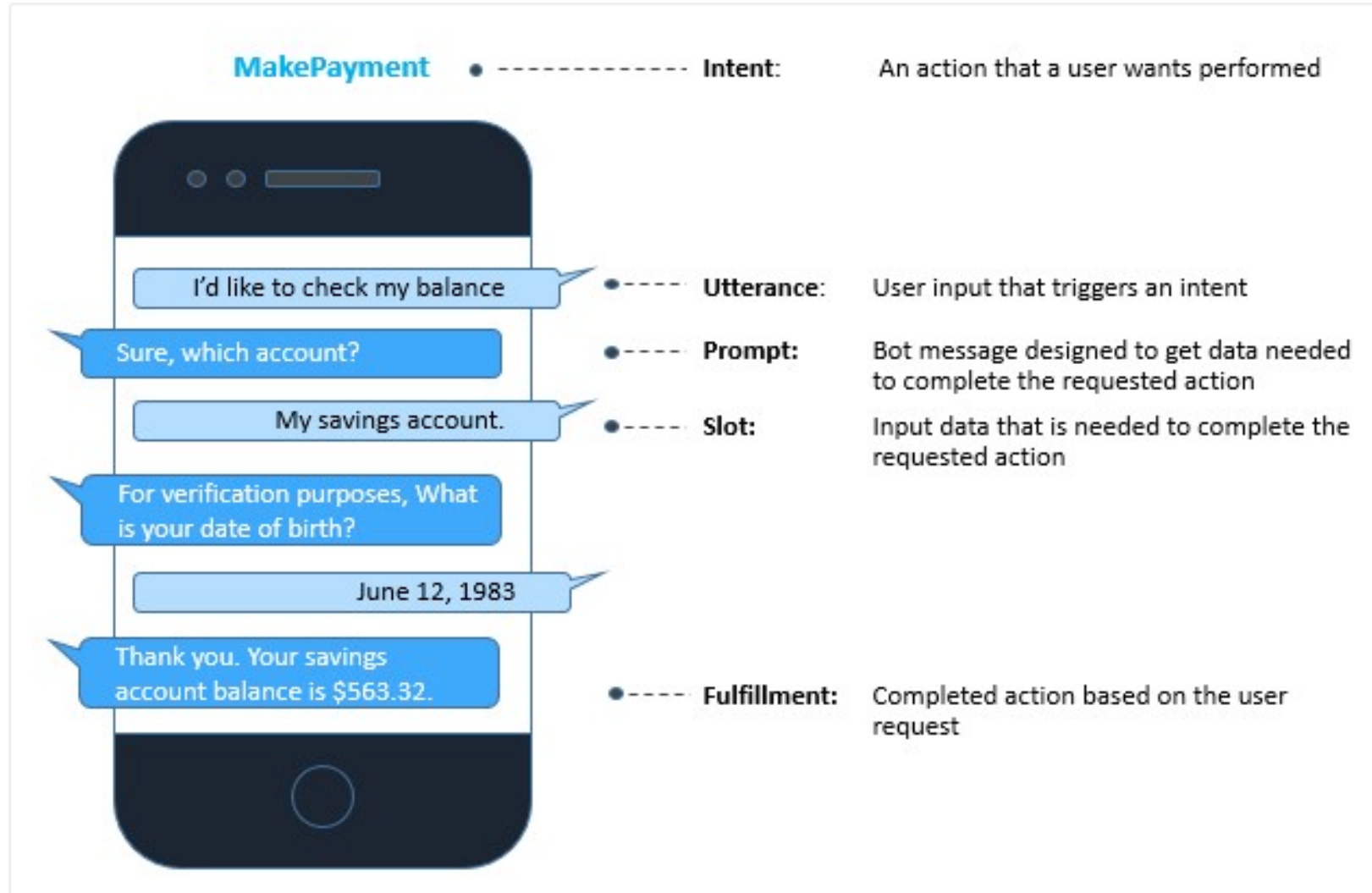
Chinese, CN (zh-CN)	✓
English, AU (en-AU)	
English, UK (en-GB)	
English, US (en-US)	
French, CA (fr-CA)	
French, FR (fr-FR)	
German, DE (de-DE)	
Italian, IT (it-IT)	
Japanese, JP (ja-JP)	
Korean, KR (ko-KR)	
Portuguese, BR (pt-BR)	
Spanish, US (es-US)	
Chinese, CN (zh-CN)	▲

Amazon Lex

Amazon Lex

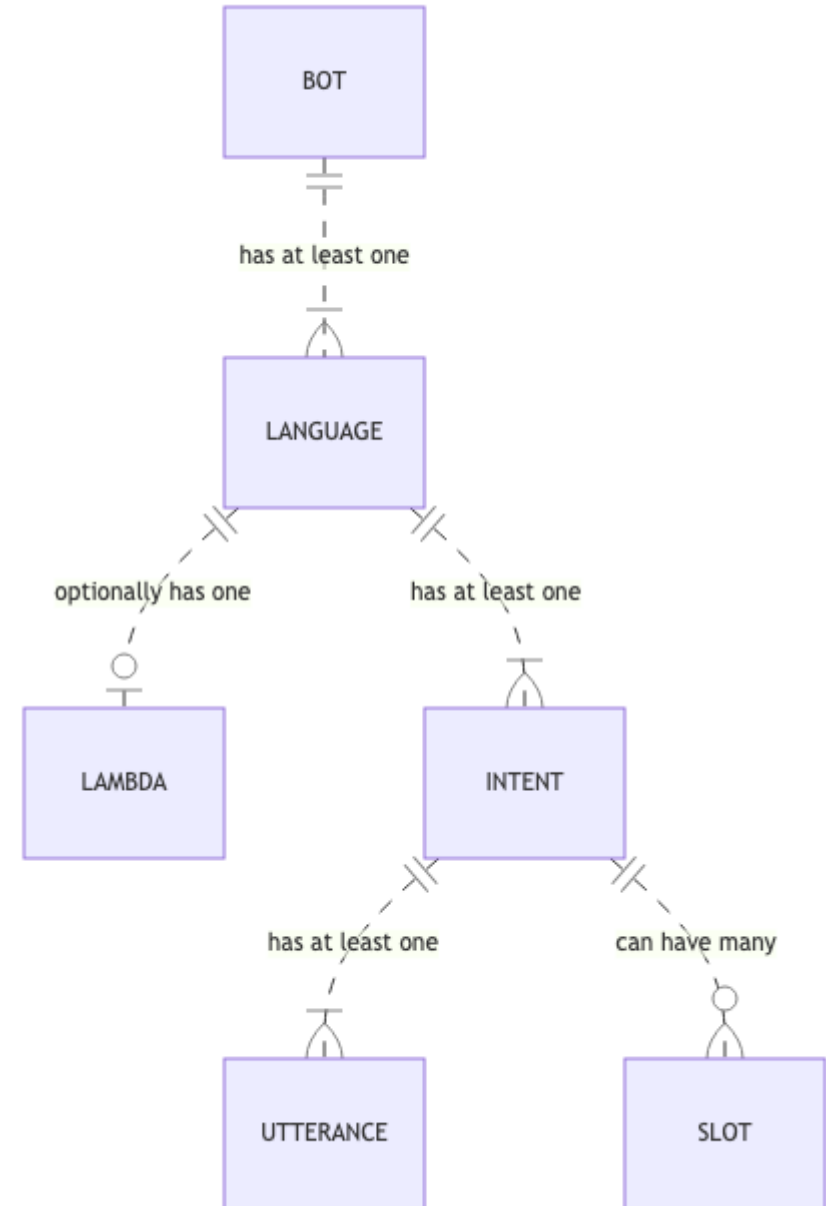
- It is a service that does more than just one of the mentioned AI services, it is a combination of those
- Create a conversational interface using voice and text
- Amazon Alexa
- Lex services:
 - Converting Speech to text
 - NLP to comprehend the intend
 - Running a business logic by Lambda to do further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

Configure bot settings [Info](#)

Creation method

☒ **Create a blank bot**
Create a basic bot with no preconfigured languages, intents, and slot types.

☐ **Start with an example bot**
An example bot with preconfigured languages and slot types. You can modify these settings.

Bot configuration

Bot name

OrderingFlowers

Maximum 100 characters. Valid characters: A-Z, a-z, 0-9, -, _

IAM permissions [Info](#)

IAM permissions are used to access other services on your behalf.

☒ **Runtime role**
Choose a role that defines permissions for your bot. To create a custom role, use the IAM console.

☐ Use an existing role.

Info Creating a role takes a few minutes. Don't delete the role or edit the trust or permissions policies until we've finished creating it.

New role
Amazon Lex creates a runtime role with permission to upload to Amazon CloudWatch Logs.

AWSServiceRoleForLexV2Bots_XAXXZY9MZQD

Children's Online Privacy Protection Act (COPPA) [Info](#)

Is use of your bot subject to the [Children's Online Privacy Protection Act \(COPPA\)](#) [?](#)

☒ Yes

☐ No

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**

▼ Intent details [Info](#)

Intent name

OrderFlowers

Maximum 100 characters. Valid characters: A-Z, a-z, 0-9, -, _

Sample utterances (2) [Info](#)

Representative phrases that you expect a user to speak or type to invoke this intent. Amazon Lex extrapolates based on the sample utterances to interpret any user input that may vary from the samples. The priority order of the sample utterances is not used to determine intent classification output.

🔍 *Filter*

Sort by added (ascending)

Preview

Plain Text

I would like to pick up flowers

I would like to order some flowers

Create a custom slot

- Select Slot Type from the left menu

Bots

OrderFlowers

Bot versions

Draft version

All languages

▼ English (US)

Intents

Slot types

Select **Add blank slot type**

Delete

Add slot type ▲

Add blank slot type

Use built-in type

Add grammar slot type

Add composite slot type

Add a new slot and its values

- Slot Type Name

Add blank slot type

Create a custom slot type for your bot.



Slot type name

FlowerTypes

Maximum 100 characters. Valid characters: A-Z, a-z, 0-9, -, _

Cancel

Add

Slot type values

Modify the list of values used to train the machine learning model to recognize values for a slot.

Search slot type values

No slot type values

You haven't added any slot type values yet.

Value

Add value

Maximum 140 characters. Valid characters: A-Z, a-z, 0-9, @, #, \$

☐ Use slot values as custom vocabulary [Info](#)

- Add Values

Save the Slot Type after adding the values

Slot type values

Modify the list of values used to train the machine learning model to recognize values for a slot.

rose

×

lilies

×

tulips

×

Add value

Create the slot FlowerType for the intent

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

The screenshot shows the 'Add slot' configuration window in the AWS Lex console. At the top, a header bar contains a dropdown menu labeled 'Prompt for slot: FlowerType' with the message 'What type of flowers would you like to order?' and a 'Slot type' dropdown set to 'FlowerTypes'. Below this, a checkbox labeled 'Required for this intent' is checked, with a subtext: 'The bot will prompt for this slot during the conversation if a value is not provided by the user.' The 'Name' field is set to 'FlowerType' and the 'Slot type' dropdown is set to 'FlowerTypes'. Under the 'Prompts' section, a text box contains the prompt 'What type of flowers would you like to order?'. At the bottom, there is a button labeled 'Advanced options'.

⋮

▼ Prompt for slot: FlowerType
Message: What type of flowers would you like to order?

Slot type
FlowerTypes

☒ Required for this intent
The bot will prompt for this slot during the conversation if a value is not provided by the user.

Name
FlowerType

Slot type
FlowerTypes

Prompts
What type of flowers would you like to order?

You can use the advanced options setting to configure rich messages such a custom payload, card groups, and SSML.

Advanced options

Add the prompt

Create a new slot for the intend

▼ Prompt for slot: PickupTime
Message: "At what time do you want the {FlowerType}..."

Slot type
AMAZON.Time

✕

☒ Required for this intent
The bot will prompt for this slot during the conversation if a value is not provided by the user.

Name
PickupTime

Slot type
AMAZON.Time ▼

Prompts
"At what time do you want the {FlowerType} to be picked up?"

You can use the advanced options setting to configure rich messages such a custom payload, card groups, and SSML.

Advanced options

Create the **slot** PickupDate for the intend

▼ Prompt for slot: PickupDate

Message: What day do you want the {FlowerType} to b...

Slot type
AMAZON.Date

✕

☒ Required for this intent

The bot will prompt for this slot during the conversation if a value is not provided by the user.

Name

PickupDate

Slot type

AMAZON.Date

▼

Prompts

What day do you want the {FlowerType} to be picked up?

You can use the advanced options setting to configure rich messages such a custom payload, card groups, and SSML.

Advanced options

Create Confirmation

Confirmation [Info](#)

Prompts help to clarify whether the user wants to fulfill the intent or cancel it.

▼ Prompts to confirm the intent

Message: *Okay, your {FlowerType} will be ready for pic...*

Responses sent when the user declines the intent

Message: -

Confirmation prompt

What will the bot say to prompt the user to confirm this intent.

Okay, your {FlowerType} will be ready for pickup by {PickupTime} on {PickupDate}. Does this sound okay?

Decline response

What will the bot say if the user says NO to the confirmation prompt.

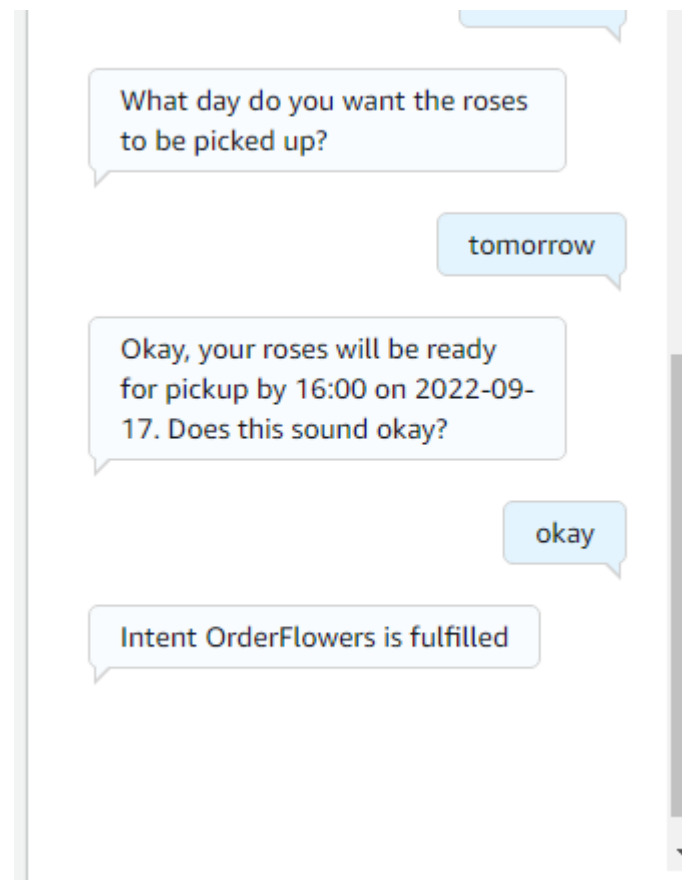
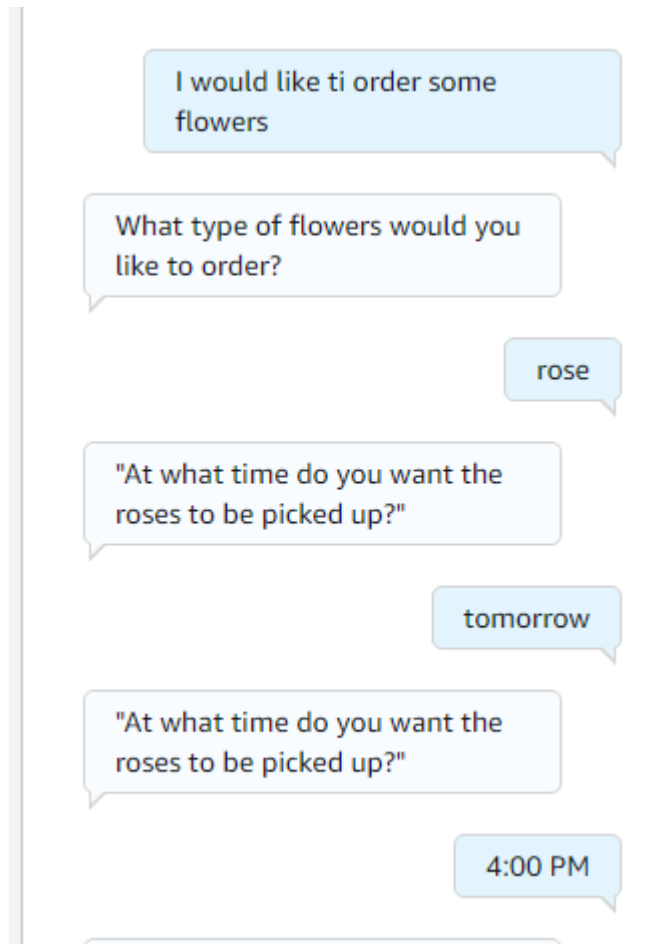
Okay. Your request will not be submitted.

Advanced options

Configure confirmation prompts and decline responses.

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

Label detection

Rekognition automatically labels objects, concepts, scenes, and actions in your images, and provides a confidence score. [Learn more](#)

[Leave us feedback](#)

Need to detect specific objects and scenes unique to your business?
Use [Rekognition Custom Labels](#) to quickly build a custom model, no machine learning experience required.



Search all labels

[Download full list](#)

Check whether we support your label

▼ Results

Car	98.8 %
Automobile	98.8 %
Vehicle	98.8 %
Transportation	98.8 %
Person	98.3 %
Human	98.3 %

[Show more](#)

Amazon Rekognition Demo

- Facial analysis

Facial analysis

Get a complete analysis of facial attributes, including confidence scores. [Learn more](#)

[Leave us feedback](#)



▼ Results



looks like a face	99.9 %
appears to be female	99.9 %
age range	25 - 35 years old
not smiling	63.5 %
appears to be happy	70 %
wearing glasses	99.9 %

[Show more](#)

► Request

Amazon Rekognition Demo

- Text in image

Text in image

[Learn more](#)



▼ Results

IT'S

MONDAY

but keep

Smiling

► Request

► Response

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.