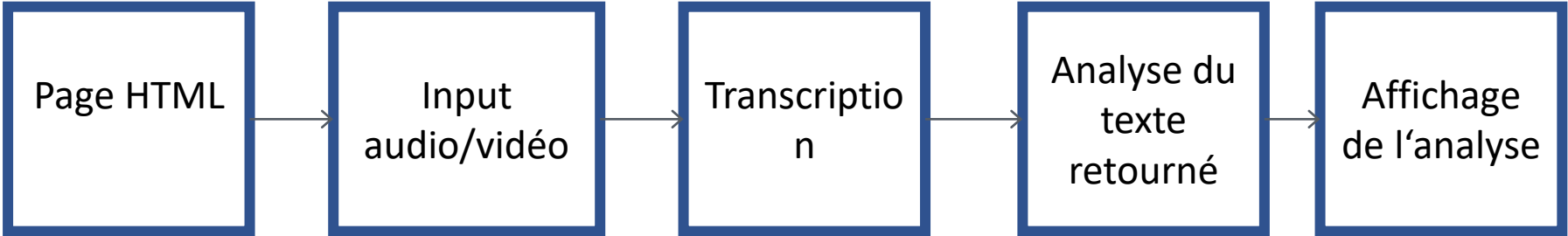


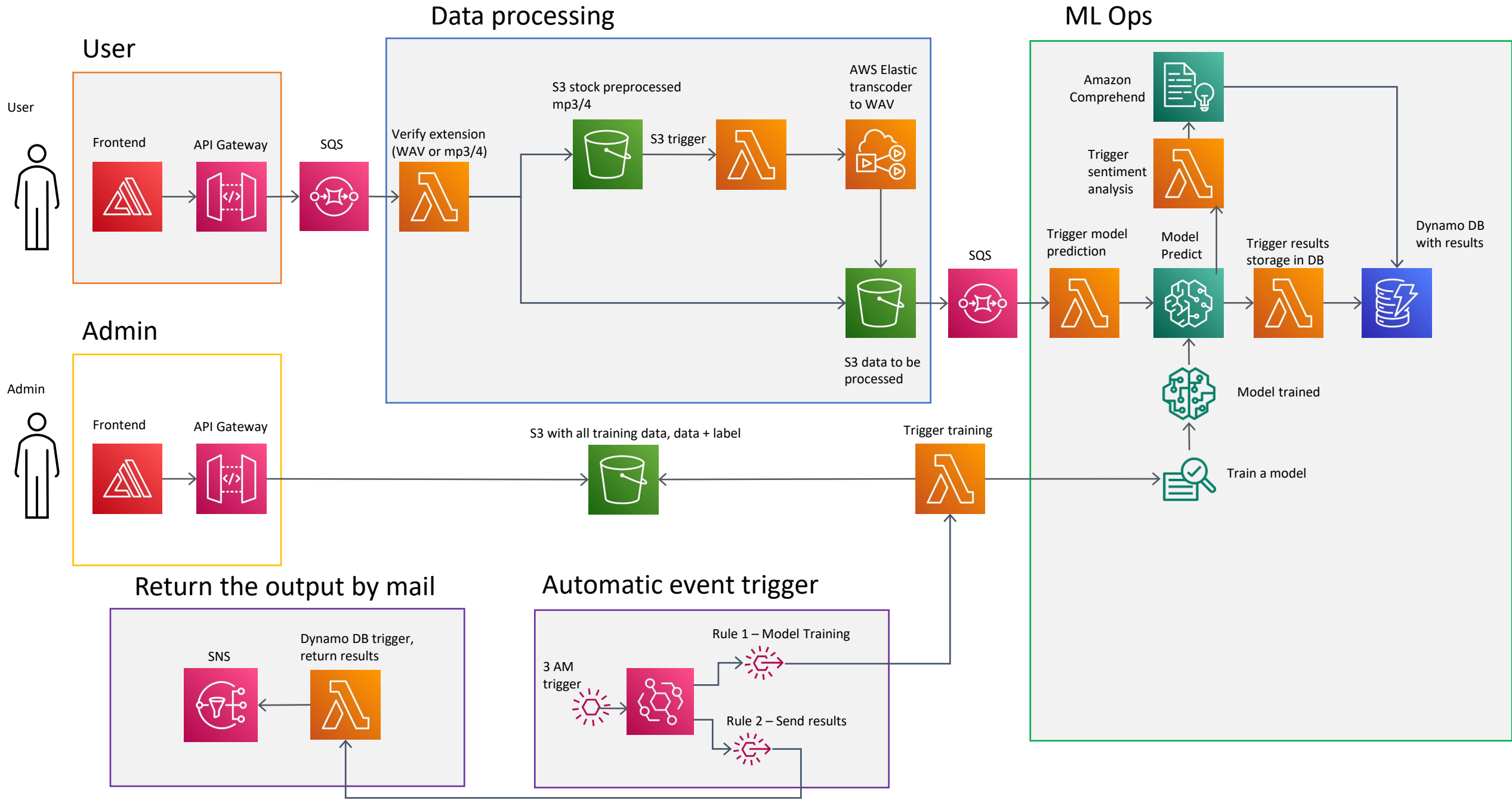
Page HTML  
Politic debate summary

Déposez ici votre  
fichier de débat:  
MP4 ou WAV

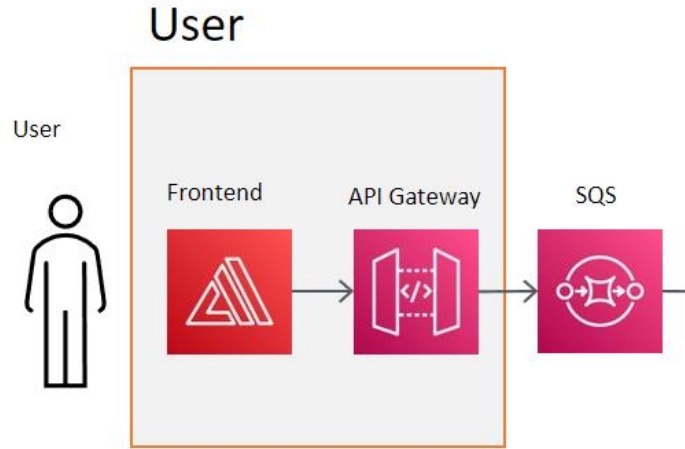
Valider

Résultat par  
mail le jour  
suivant





# Description

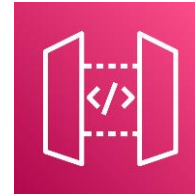


## AWS Amplify



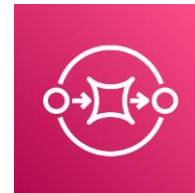
Its purpose is to create a front-end interface for the user to upload its file.

## API Gateway



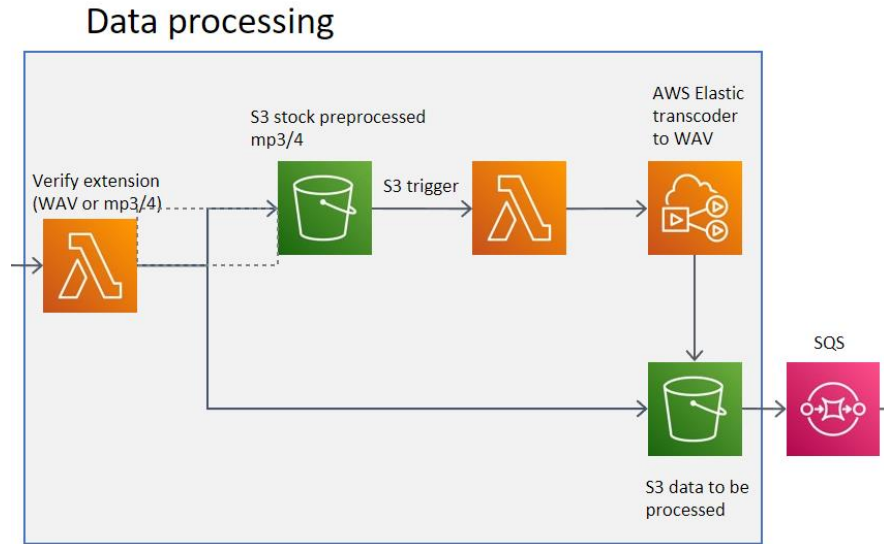
Its purpose is to make API calls to our pipeline and maintaining/securing the incoming flux.

## SQS



Its purpose is to make sure the incoming files are sent to our pipeline without any loss.

# Description



## AWS Lambda



We have two lambdas in this part :

- One is to ensure that the files have an accepted extension (either wav/mp3/mp4)
- The other is triggered when a file is put inside the S3 bucket to launch the transcription

## S3 Bucket



Used to store data.

We use this over a database because we didn't require to access the data directly, S3 are cheaper, and useful to transfer data.

## AWS Elastic Transcoder



Is only used when the incoming file is an mp3 or mp4. It will convert it to a WAV file that can be used by our model.

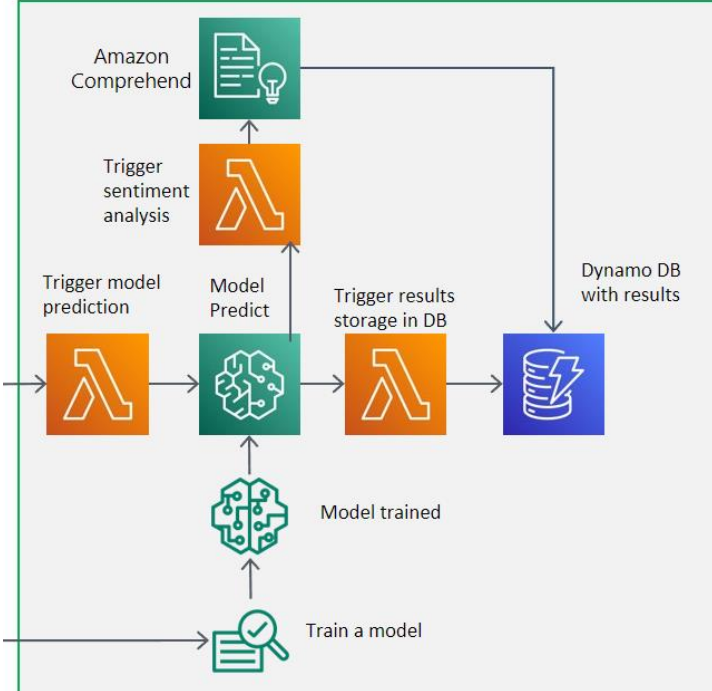
## SQS



Its purpose is to distribute all the processed data to our model.

# Description

## ML Ops



## AWS Lambda



- We have three lambdas in this part :
- The first one is going to trigger the launch of our model
  - The two others are triggered at the end of the prediction. One will trigger the analysis of the text and the other one will store the data in a database.

## Amazon SageMaker



It allows us to use our model to transcribe WAV files to text.

## Amazon Comprehend



Is used to analyze the sentiment of the transcribed text.

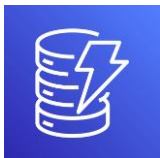


Once the model is trained, it will replace the current model



Train a model with data stored in a S3 bucket

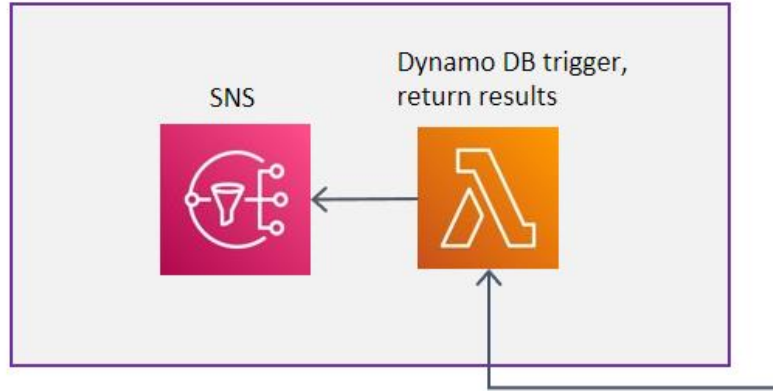
## Dynamo DB



It allows us to store a transcription and its sentiment analysis. This time we choose a database over a bucket because we need to be able to pair a transcription and an analysis. It's also an easier way to access the data to send it back to the user.

# Description

## Return the output by mail



## AWS Lambda



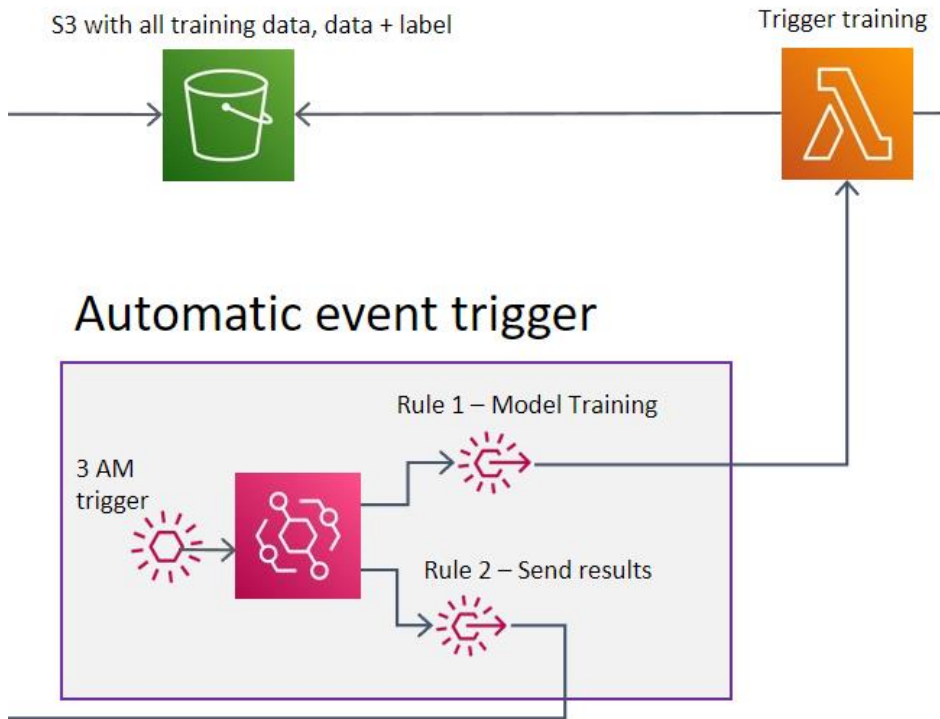
The lambda is triggered at 3 AM every day by an Event bridge. It will go fetch the data in the database and return it.

## SNS



Its purpose is to get the results of the transcription back to the user by mail at 3 AM.

# Description



## AWS Lambda



The lambda is triggered every day at 3 AM. It will go fetch new data in the Admin bucket and will feed it to the model to retrain it.

## AWS Event bridge



Its purpose is to manage multiple events from a trigger

### Event



Here this event is active once when it's 3 AM

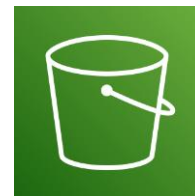
### Rule



We have 2 rules :

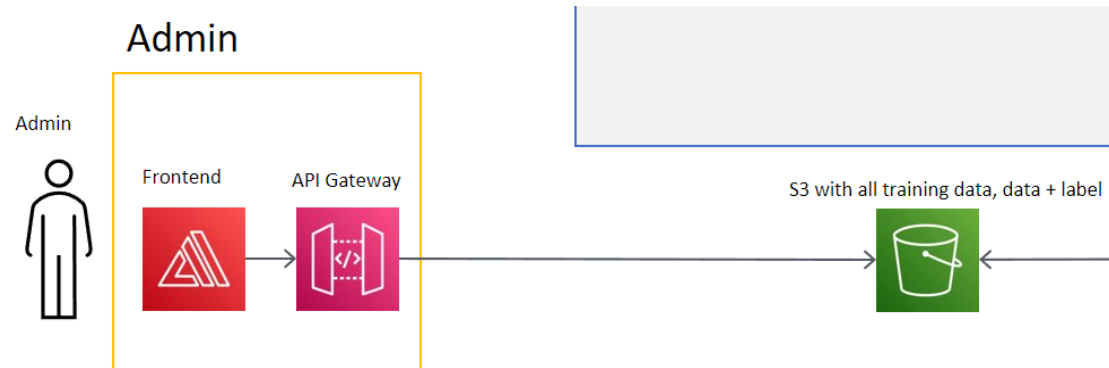
- one for triggering the model training,
- the other to trigger mails to send results back to the users

## S3 bucket



This S3 bucket stores all of the training data

# Description

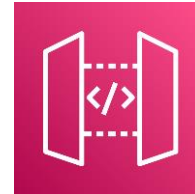


## AWS Amplify



Its purpose is to create a front-end interface for the admin to upload new data to a S3 bucket.

## API Gateway



Its purpose is to make API calls to the S3 bucket and maintaining/securing the incoming flux.

## S3 bucket



This S3 stores all of the training data given by an admin.



## Pro

- Does the job
- Low cost architecture
- Easy use/access for User and Admin
- RGPD proof
- Safe

## Cons

- Delayed results
- No training with the incoming data from user
- No user page to show and correct bad predictions from model