

Automatic Translations

Skillshare


Zombieland CLI Team (cohort 10)

Carlos Andres Garcia (agzsoftsi)

Leonardo Calderon (leocjj)

Ivan Dario Lasso (ilasso)

June 2020





Overview

0. Team members/roles
1. Company
2. The inspiration for the project
3. Technologies & Architecture
4. Core algorithms/code snippet
5. Process, collaboration, and timeline
6. Main challenge overcome
7. Live Demo (front-end, back-end)
8. Learning lessons
9. Questions session



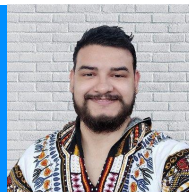
Team and roles

Carlos A. García M., SWE

@karlgarmor

agzsoftsi

FrontEnd and Design:
HTML, CSS, Javascript,
AWS CloudFront,
AWS S3



Leonardo Calderon J., SWE

@leocj

leocjj

BackEnd:
Python, AWS Lambda,
AWS API Gateway,
AWS S3
AWS Translate



Ivan D. Lasso G., SWE

@ilasso

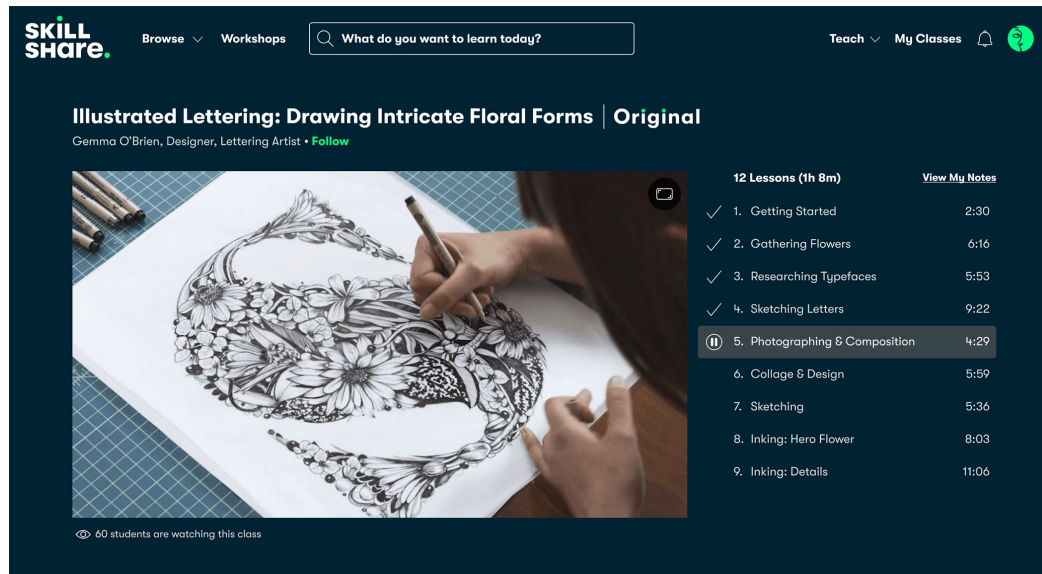
ilasso

FrontEnd and Testing:
HTML, CSS, Javascript,
AWS CloudFront,
AWS S3




Skillshare

- Online learning community
- **Learning from educational videos**
- Available through subscription
- Focus on interaction, with the primary goal of **learning by completing a project.**





The screenshot displays the Skillshare website interface. At the top, the Skillshare logo is on the left, and navigation links for 'Browse', 'Workshops', 'Teach', and 'My Classes' are on the right. A search bar with the placeholder text 'What do you want to learn today?' is centered. Below the navigation, the main heading for the selected class is 'Illustrated Lettering: Drawing Intricate Floral Forms | Original', followed by the instructor's name 'Gemma O'Brien, Designer, Lettering Artist' and a 'Follow' button. The central part of the page features a video player showing a person's hands drawing a complex floral design on a piece of paper with a grid pattern. To the right of the video player is a list of 12 lessons, each with a checkmark, a title, and a duration. The fifth lesson, '5. Photographing & Composition', is highlighted with a play button icon and a duration of 4:29. Below the video player, a small icon and text indicate '60 students are watching this class'.

SKILL SHARE Browse ▾ Workshops Teach ▾ My Classes 

Illustrated Lettering: Drawing Intricate Floral Forms | Original

Gemma O'Brien, Designer, Lettering Artist • [Follow](#)



 60 students are watching this class

12 Lessons (1h 8m) [View My Notes](#)

✓ 1. Getting Started	2:30
✓ 2. Gathering Flowers	6:16
✓ 3. Researching Typefaces	5:53
✓ 4. Sketching Letters	9:22
5. Photographing & Composition	4:29
6. Collage & Design	5:59
7. Sketching	5:36
8. Inking: Hero Flower	8:03
9. Inking: Details	11:06



The inspiration for the project

1

Skillshare wants to go beyond English spoken users.

1

Online translation for transcriptions in several languages.

2

Videos and transcriptions are in English.

2

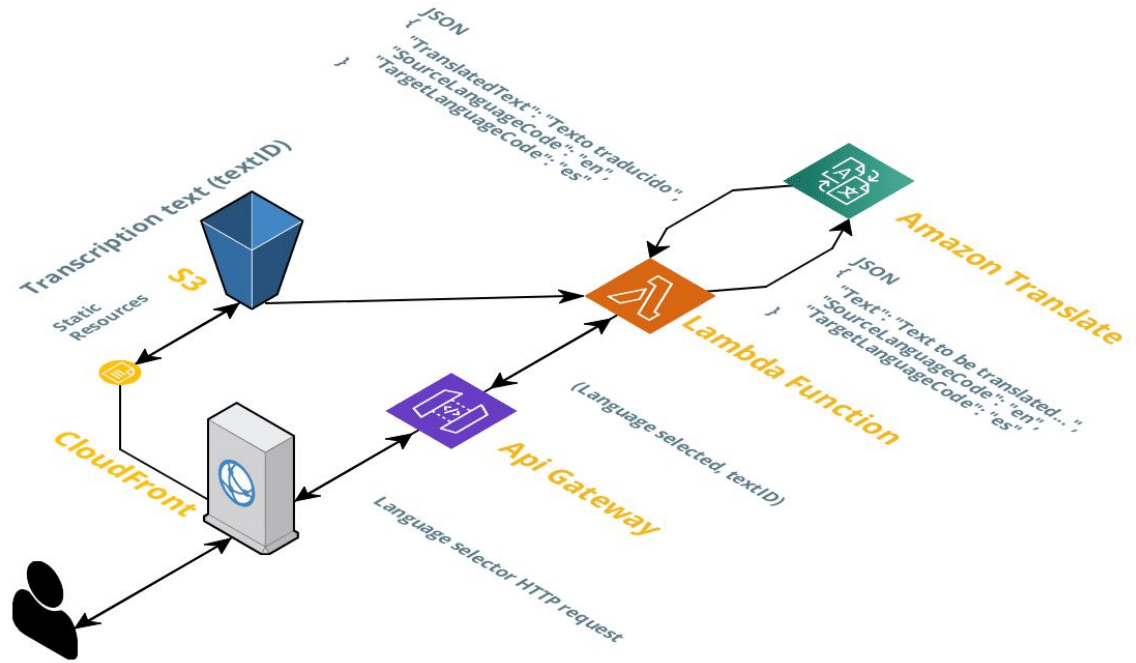
Automatic process without affecting user experience.

Technologies & Architecture

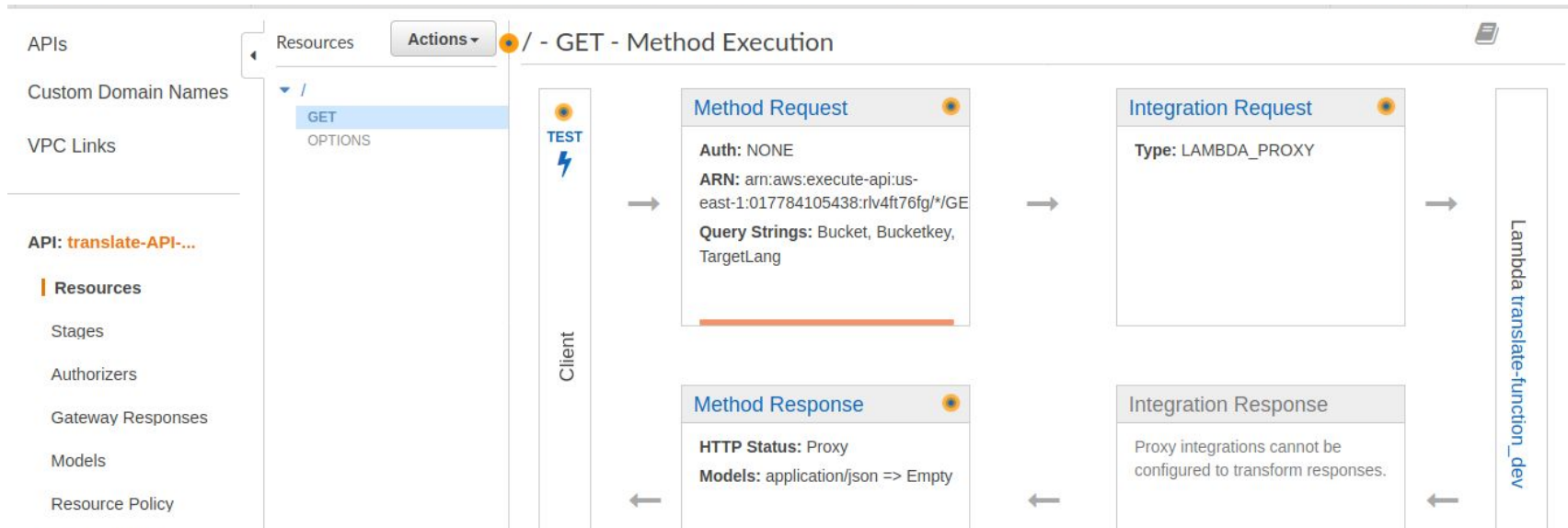
AWS - Amazon Web Services:

- Lambda function (Python)
- S3 objects
- API Gateway
- Translate
- Comprehend

JSON format



Core algorithms/code snippet





lambda_function.py

```
from boto3 import client
from json import dumps
from os import environ
```

```
def lambda_handler(event, context):
```

```
    """
    This function is called from an API Gateway (translate-API),
    receive three parameters inside an array to get a translation.
    """
```

```
    key = event['queryStringParameters']['Bucketkey']
    targetLang = event['queryStringParameters']['TargetLang']
```

```
    return lambdaFunction(key, targetLang)
```

```
def lambdaFunction(key, targetLang):
```

```
    resp['headers']['targetKey'] = key[:key.rfind('.')] + '_' + targetLang + '.srt'
```

```
    """ IF ORIGINAL TRANSCRIPTION STORED IN S3 IS NOT FOUND, RETURN 404 """
```

```
    if not keyExist(key):
```

```
        resp['body'] = dumps("Video transcription not found")
```

```
        return resp
```

```
    """GET STORED TRANSLATION, IF NOT, TRANSLATE TRANSCRIPTION AND STORE IT"""
```

```
    if keyExist(resp['headers']['targetKey']):
```

```
        translationFile = s3Client.get_object(Bucket=bucket, Key=resp['headers']['targetKey'])
```

```
        resp['body'] = translationFile['Body'].read().decode("utf-8")
```

```
        resp['statusCode'] = 200 # Ok
```

```
        return resp
```

```
    else:
```

```
        # To get S3 object with original transcription
```

```
        response = s3Client.get_object(Bucket=bucket, Key=key)
```

```
        text = str(response['Body'].read().decode('utf-8'))
```

```
        # Detect source language by sending first 100 characters to check
```

```
        responseLanguage = comprehendClient.detect_dominant_language(Text=text[:100])
```

```
        if responseLanguage['Languages'][0]['Score'] <= languageThreshold:
```

```
            resp['statusCode'] = 412 # Precondition Failed
```

```
            resp['body'] = dumps("Video language not supported")
```

```
            return resp
```

```
        sourceLang = responseLanguage['Languages'][0]['LanguageCode']
```



lambda_function.py

```
def translateBatches(sourceLang, targetLang, text):
    """ Translate text in batches of step size but with cutting rows of file """
    lenn = len(text)
    if lenn <= step:
        try:
            translatedText = translateClient.translate_text( Text = text,
                                                         SourceLanguageCode = sourceLang,
                                                         TargetLanguageCode = targetLang
                                                         )
        except translateClient.exceptions.ClientError as e:
            if e.response['Error']['Code'] == "500":
                resp['statusCode'] = 500 # Internal error
            return None
    return translatedText.get('TranslatedText')
```

Collaboration

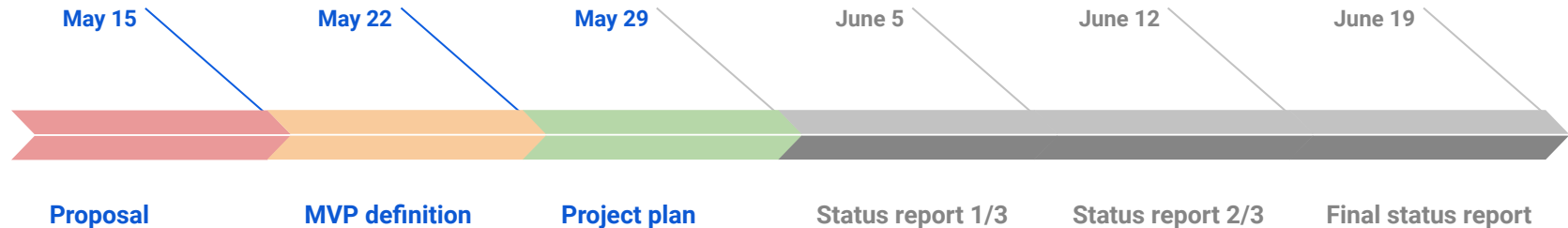


Scrum based process

- The Sprint: weekly, Friday to Thursday
- Sprint Planning: each friday
- Daily Scrum: Mo-Th, 11 am.
- Sprint Review: each thursday
- Sprint Retrospective: each friday

- All the teamwork was performed remotely
- Individual accounts for Skillshare and AWS
- Development repository in GitHub
- Internal team rules:
 - Daily meeting with the Client representative.
 - Two daily team meetings
 - To share individual research
 - To do pair programming to integrate codes.
 - Special rule to avoid migration to production code without the approval of all the team.

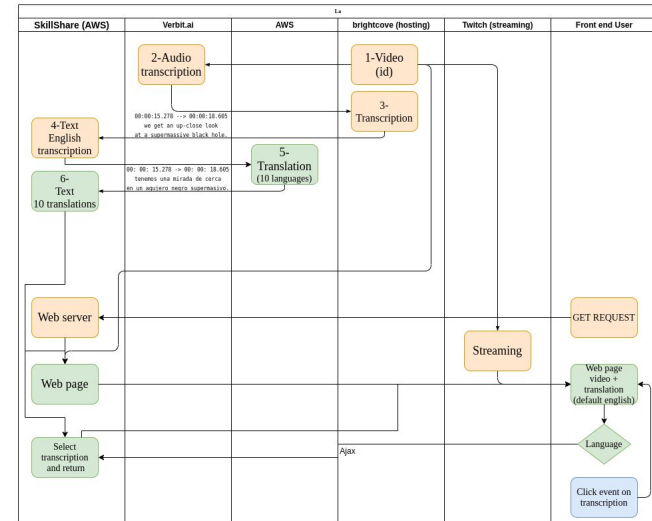
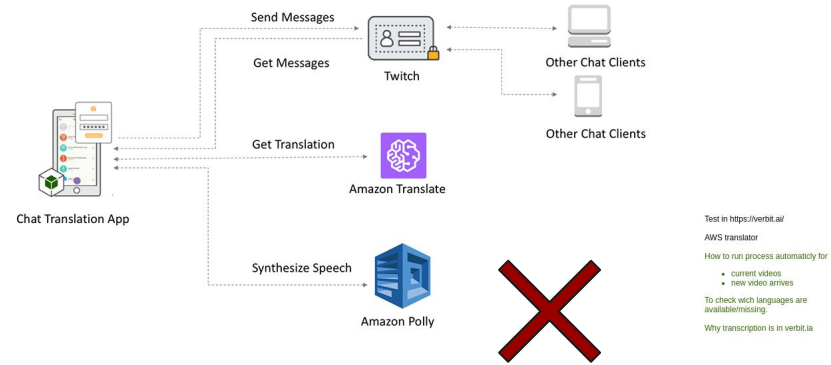
Timeline



Main challenge overcome

Initial architecture definition has struggled because of technical details not well known by the client representative.

However, his trust in our decisions helped.





Live demo

Landing page:

<https://transcriptions-translations.s3.amazonaws.com/landing/index.html>

MVP:

<https://transcriptions-translations.s3.amazonaws.com/index.html>

API test:

<https://ytuln8zsz1.execute-api.us-east-1.amazonaws.com/translate/text?TargetLang=es&Bucketkey=formats/video1t.srt>



Learning lessons

- 01 | To define architecture with an extended team.
- 02 | Scrum helps to organize, prioritize, and develop.
- 03 | Well defined team rules help during the journey.
- 04 | To use the right technology rather than use “the best”.
- 05 | We are capable to learn what we need.





Thank you.

Holberton School
Silvia Guzmán
Johanna Alzate
Laura C. Sánchez
Andrés Barreto
Cohort 10 Cali peers

We made it, keep moving.

