

소개

2024년 1학기에 수강한 클라우드 컴퓨팅의 텀프로젝트이다.

프로젝트명

AWS Transcribe를 사용한 음성을 글자로 변환(Speech-To-Text)해주는 웹사이트 제작

- AWS Transcribe를 사용한 음성노트(SpeechNote) 웹사이트 제작

웹사이트 결과물 링크 : <https://d3yte1ymgtgsl.cloudfront.net/>

프로젝트 멤버 및 역할 분담

name	part
김태훈	AWS 전체 구축, 파일 업로드 lambda 함수(file_upload_backend)구현
박민재	이메일 전송 lambda 함수(save_text, send_email) 구현
김성문	오디오 텍스트 변환 lambda 함수(audio_transcribe) 구현
박재열	웹페이지(프론트엔드) 구현

프로젝트 소개

AWS Transcribe를 사용한 음성을 텍스트로 변환(STT)해주는 서비스를 만드는 프로젝트입니다. 사용자가 음성 파일을 이메일 주소, 언어코드와 함께 웹사이트를 통하여 업로드하면, 텍스트로 변환되어 사용자의 이메일로 전송됩니다.

사용되는 클라우드 서비스는 Amazon CloudFront, Amazon S3, Amazon API GateWay, AWS Lambda, Amazon RDS, AWS Transcribe, AWS CloudWatch입니다. 이 프로젝트에서 각각의 용도는 다음과 같습니다.

Amazon CloudFront

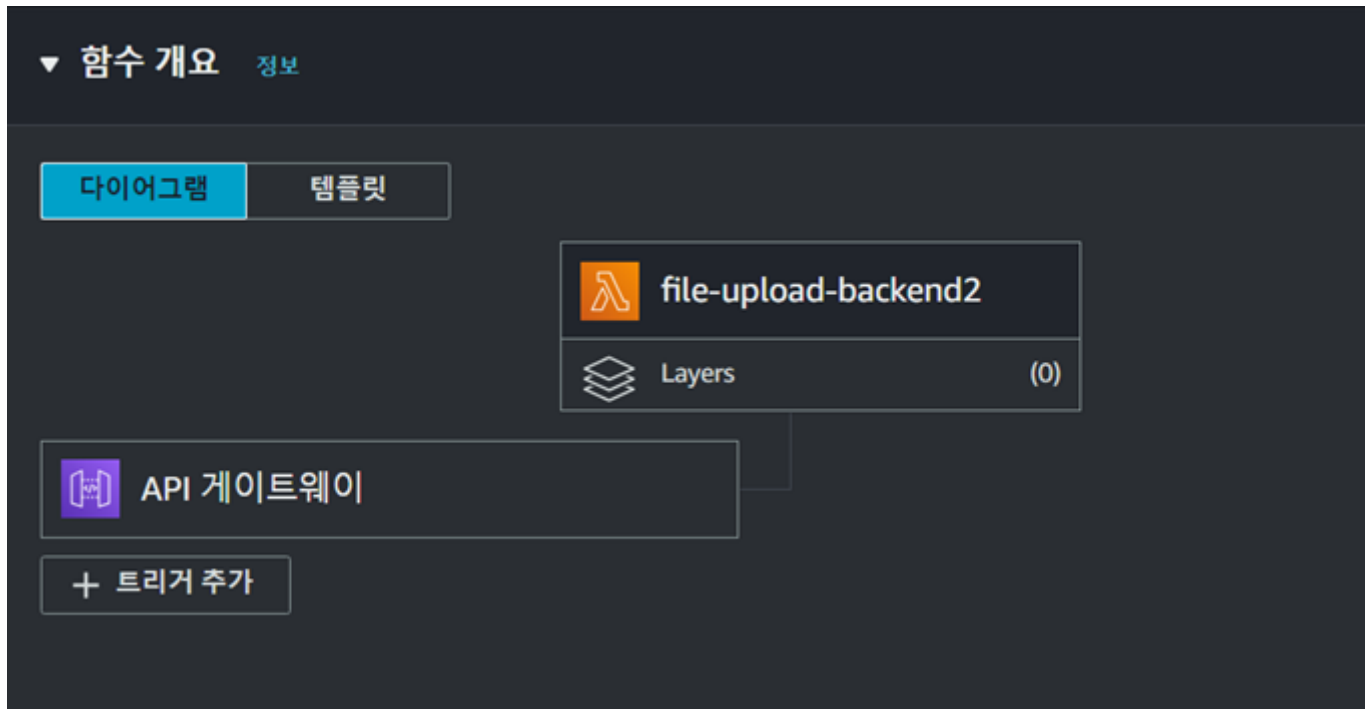
S3에 저장된 웹페이지를 캐싱하여, 사용자가 S3에 직접 접근하지 않고, 근처에 있는 CloudFront에서 웹페이지 파일을 다운로드받을 수 있게 한다. S3 퍼블릭 액세스를 차단할 수 있으므로 보안을 강화할 수 있고, https가 적용된다.

Amazon S3

정적 웹페이지에 필요한 파일(html 등)을 저장하고, 사용자가 업로드한 오디오 파일과 텍스트로 변환된 결과를 저장한다. 만약 s3에 오디오 파일이 업로드 되면, 텍스트 변환을 담당하는 lambda 함수를 호출하게 된다.

Amazon API GateWay

사용자와 lambda 함수를 연결시켜준다. 사용자가 해당 API주소로 오디오 파일과 이메일 정보 등을 전송하면, API GateWay는 전송받은 정보와 함께 lambda 함수를 호출한다. lambda 함수에서 반환된 값은, 사용자에게 전달한다.



위 그림과 같이 API GateWay를 통해 lambda 함수가 호출된다.

AWS Lambda

Lambda는 서버리스 서비스로서, 서버 없이 함수로만 구성되어 있다. 이 함수는 event를 매개변수로 받아 response를 반환한다. nodejs, python, java 등으로 함수를 구성할 수 있으며, 시스템의 lambda 함수들의 언어가 같을 필요가 없다. 이 프로젝트에서, lambda 함수를 이용하여 오디오 파일을 S3에 저장하고, 이메일 정보 등을 DB에 저장하고, transcribe를 요청하며, 텍스트로 변환된 파일을 이메일로 전송한다.



위 그림은 python으로 제작한 간단한 lambda 함수이다. event를 받아 응답코드 및 json을 반환한다.

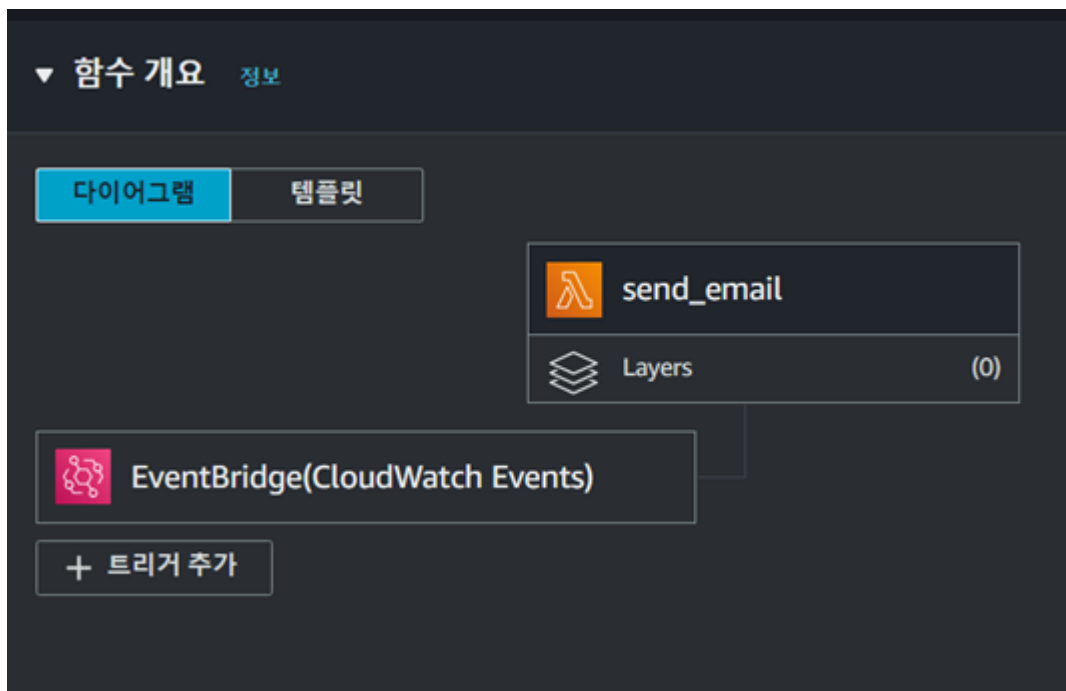
Amazon RDS

사용자에게서 받은 오디오 파일과 관련된 정보를 저장한다. 이 프로젝트에서는 mysql을 사용하며, 서로 다른 lambda 함수끼리 정보(이메일, 언어 코드 등)를 공유할 때 필요하다. table의 구성은 아래와 같다.

```
fileTable(  
  email VARCHAR(320) NOT NULL,  
  filename VARCHAR(600) NOT NULL,  
  language VARCHAR(10) NOT NULL,  
  time TIMESTAMP NOT NULL  
)
```

AWS CloudWatch and EventBridge

AWS에서 제공하는 모니터링 서비스이다. 원래는 리소스 사용률 등을 모니터링할 수 있고, 일정 사용률 초과 등의 이벤트가 발생하면 사용자에게 알리는 기능을 가지고 있다. 이 프로젝트에서는 이를 활용하여, 오디오 텍스트 변환이 완료되면, 이를 CloudWatch가 탐지한다. EventBridge가 CloudWatch에서 텍스트 변환 완료 이벤트가 생성된 것을 확인하고, 이메일을 보내는 lambda 함수를 호출하게 된다. Transcribe를 요청하는 lambda 함수에서 이메일까지 처리하지 않는 이유는, lambda 함수의 동작 시간은 최대 15분인데, 변환은 그보다 오래 걸릴 수 있기 때문이다.



cloudwatch에서 이벤트를 감지하면, EventBridge가 lambda 함수를 호출한다.



EventBridge가 lambda 함수를 호출할 이벤트 패턴은 위와 같다.

필요성 및 활용방안

회의, 강의, 강연 등에서 녹음을 통해 다시 내용을 정리하거나 복습하곤 한다. 그러나 몇시간에 가까운 녹음 파일을 듣기에는 시간이 많이 걸린다. 따라서 이것을 텍스트로 변환하여 볼 필요가 있다. 따라서 AWS Transcribe 서비스를 이용한 STT웹서비스를 구축하려고 한다. 이 서비스는 사용자가 텍스트로 변환하고 싶은 음성 파일이 있을 경우 웹에서 음성 파일을 이메일과 언어 정보와 함께 업로드하면, 사용자의 이메일로 변환된 텍스트가 파일(xxx.txt)형태로 전송된다.

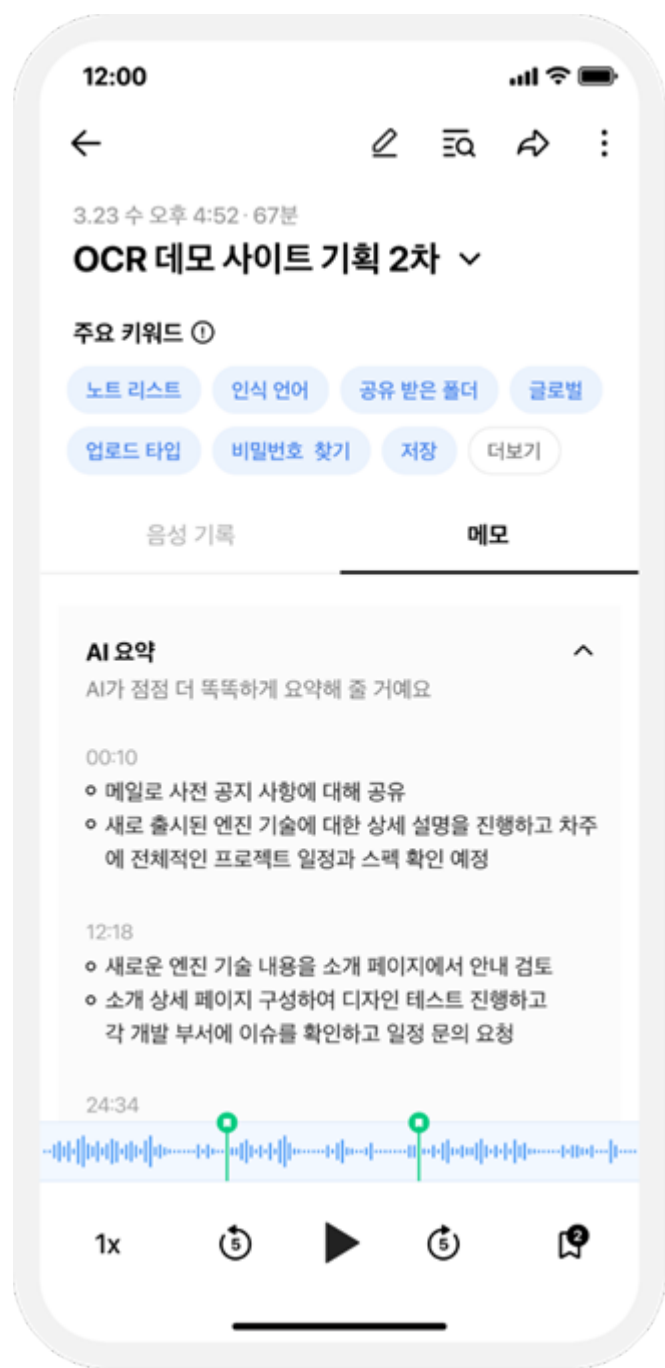
선행기술 조사

네이버에서 개발한 [클로바노트\(CLOVA note\)](#)는 웹이나 모바일 앱에서 음성을 녹음하거나 음성파일을 업로드하면, 그것을 텍스트로 변환해주고, 이것을 시간과 화자와 함께 보여준다.

The screenshot shows a web application interface. On the left is a sidebar with a home icon and the following menu items: 홈, 전체 노트, 주간 회의, 리서치, 공유 받은 노트, 미팅, 상담, 공유한 노트, 휴지통. The main content area has a top navigation bar with tabs: 웰컴 메시지, 전 연령 대상, 장소 제외, 빈티지, 스태프 모집, 더보기. Below this is a section titled '음성 기록' with a '편집' button and two buttons on the right: '점유율' and '참석자 숨기기'. The main content area lists six audio notes, each with a speaker icon, a name, a timestamp, and a text preview. The notes are: 최민아 (00:10), 박규리 (00:20), 이세진 (00:22), 박규리 (00:24), 홍주연 (00:26), and 박규리 (00:27). The text previews are: '사람들이 보편적으로 접할 수 있는 플리마켓, 버스킹 문화를 참고해서 문화 마켓이라는 주제로 이벤트를 준비해보려고 합니다.', '문화 마켓에서 책이나 빈티지 소품을 판매하거나, 먹거리와 음악 공연을 준비해서 진행하는 것은 어떨까요?', '지난 주에 진행한 이벤트 기획 회의에서 다양한 의견이 나왔는데요. 사람들의 참여를 많이 이끌어내는 것에 중점을 두고 기획팀에서 주제를 조금 더 구체화해보았습니다.', '사람들이 보편적으로 접할 수 있는 플리마켓, 버스킹 문화를 참고해서 문화 마켓이라는 주제로 이벤트를 진행해 보고자 합니다.', '세진님 구체적인 예시도 공유 부탁드립니다.', '책이나 빈티지 소품은 사람들이 관심 있어 할 만한 물건이고, 먹거리와 음악 공연은 사람들을 즐겁게 할 수 있는 요소이니까요.'

(그림 1 클로바노트에서 음성 기록이 생성된 화면 출처)

녹음 파일 업로드 시 월 300분의 제한이 있으며, 녹음 파일의 길이는 180분으로 제한된다. 앱이나 웹에서 녹음 시, 1회 녹음은 180분으로 제한되지만, 텍스트 변환은 제한이 없다. 다만 300분 초과시 텍스트 변환 속도가 매우 느려진다(최대 24시간). 아직 유료 요금제는 제공되지 않는다. 클로바노트는 추가로 AI요약 기능도 있다. AI 요약기능은 월 10회 제공하며 아래와 같이 요약하여준다.



(그림 2 클로바노트에서 AI요약을 실행한 장면 [출처](#))

프로젝트 개발 결과물

Project Tree

```

📁 stt-service
├── .idea
├── LICENSE
├── README.md
├── README.pdf
├── diagram //diagram folder
│   ├── diagram.drawio
│   └── diagram.pdf
```

```

├─ diagram.png
├─ diagram.svg
├─ presentation //presentation folder
├─ 7조_발표.pdf
├─ 7조_발표.pptx
├─ 7조_발표동영상.mp4
├─ src
├─ audio-transcribe //변환요청하는 람다함수
├─ README.md
├─ lambda_function.py
├─ pymysql
├─ file-upload-backend //API GateWay를 통해 들어온 파일과 이메일 정보를 s3와 DB에
저장하는 람다 함수
├─ README.md
├─ file-processing-service.mjs
├─ index.mjs
├─ node_modules
├─ package-lock.json
├─ package.json
├─ util.mjs
├─ frontend // 프론트엔드 정적 웹페이지(html/css/js(vue.js))
├─ README.md
├─ css
├─ about.2e450432.css
├─ about.7a0336af.css
├─ app.9e5c990d.css
├─ chunk-vendors.a078eef8.css
├─ favicon.ico
├─ index.html
├─ js
├─ about.348ee159.js
├─ about.348ee159.js.map
├─ about.de8a5e62.js
├─ about.de8a5e62.js.map
├─ app.2ed68f2d.js
├─ app.2ed68f2d.js.map
├─ app.e8578df5.js
├─ app.e8578df5.js.map
├─ chunk-vendors.0cdd5c29.js
├─ chunk-vendors.0cdd5c29.js.map
├─ chunk-vendors.fb3498cd.js
├─ chunk-vendors.fb3498cd.js.map
├─ save-text //transcribe로 부터 변환결과를 받아 s3에 저장하는 함수
├─ README.md
├─ lambda_function.py
├─ send-email //s3에 저장된 변환결과(.txt 파일)를 이메일로 전송
├─ PyMySQL-1.1.1.dist-info
├─ README.md
├─ lambda_function.py
├─ pymysql
├─ sql // RDS DB table 생성에 사용된 SQL
├─ createTable.sql
├─ 시스템구축.pdf

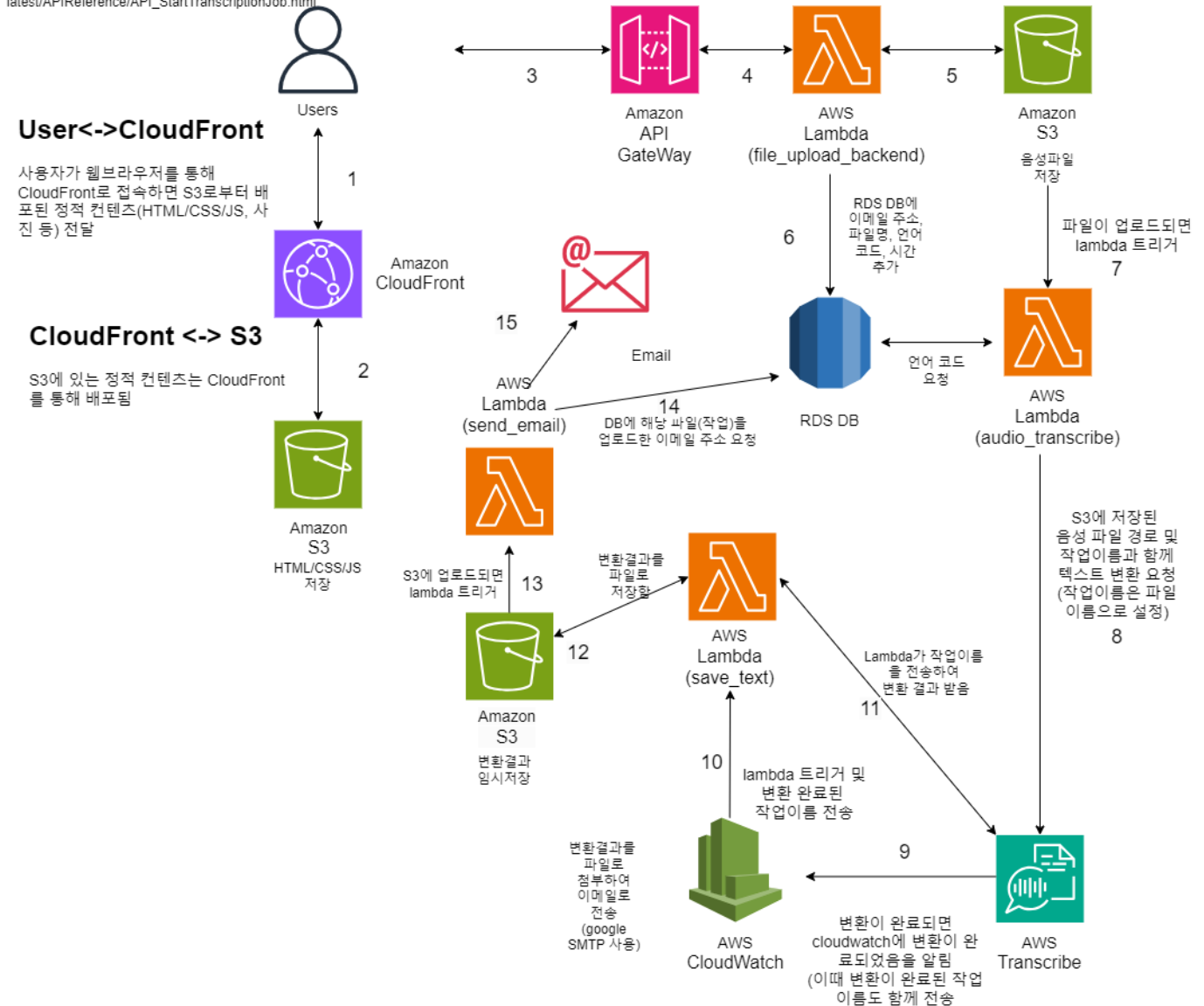
```

©generated by Project Tree Generator

구상도

언어코드

언어(영국) : en-GB
 언어(미국) : en-US
 한국어(한국어) : ko-KR
 한국어(일본어) : ja-JP
 중국어(간체) : zh-CN
 중국어(번체) : zh-TW
 자동감지 : "auto"(IdentifyLanguage=True)
https://docs.aws.amazon.com/transcribe/latest/APIReference/API_StartTranscriptionJob.html



1. 사용자가 cloudfront를 통해 정적 웹페이지에 접속한다.
2. cloudfront는 s3에 저장된 컨텐츠(html,css,js 등)를 가져와서 사용자에게 전달하거나 캐싱된 웹페이지를 전달한다.
3. 사용자가 웹페이지에서 이메일, 언어코드, 음성파일을 첨부하여 API로 POST요청을 보낸다.
4. API Gateway는 사용자에게서 받은 정보와 파일을 request에 담아 file_upload_backend를 호출한다.
5. 해당 lambda 함수는 이메일, 언어코드, 파일명, 시간을 RDS에 기록한다.

6. 해당 lambda 함수는 S3에 파일을 업로드한다.
7. S3에 파일이 업로드되면, S3는 audio_transcribe 람다 함수를 호출한다(이때 업로드된 파일명이 전달된다).
8. RDS에서 업로드된 파일명에 해당하는 언어 코드를 전달받고, S3에 저장된 음성파일 경로를 작업이름, 언어코드와 함께 transcribe에 텍스트 변환 요청한다(작업이름은 파일이름으로 설정)
9. 변환이 완료되면, cloudwatch가 이를 감지
10. eventbridge에서 save_text 람다 함수를 호출함(작업이름과 함께 호출).
11. save_text이 transcribe에 작업이름을 전송하여 변환 결과를 받는다.
12. 변환 결과를 s3에 파일로 저장한다.
13. s3에 변환결과가 업로드되면 send_email 람다 함수를 호출한다.
14. RDS DB에서 파일이름에 해당하는 이메일 주소 얻음
15. google SMTP를 사용하여 변환 결과를 파일 형태로 전송

구성 방법: [pdf 참고](#)

결과물 소개

웹사이트 결과물 링크 : <https://d3yte1ymgtgsl.cloudfront.net/>

동작 영상 : [youtube link](#)

녹음파일 출처 : [link](#)

웹페이지

[!NOTE] AWS lambda 함수는 6MB를 초과하는 요청은 거절하도록 설정되어있어 약 6MB이상의 음성파일은 업로드할 수 없습니다. 참고: [aws lambda docs](#)



(사진1 : 웹페이지를 들어가면 홈 화면은 다음과 같다.)

CLOUD-TERM Home Post

Welcome to PSQLI service!

This is the service to convert audio to text
If you want to submit your own problems, send email to admin(admin@example.com)

(사진2 : 웹페이지 홈 화면을 확대한 모습. 왼쪽 상단에 Post를 눌러 음성파일(mp3)을 업로드할 수 있는 창으로 전환할 수 있다.)

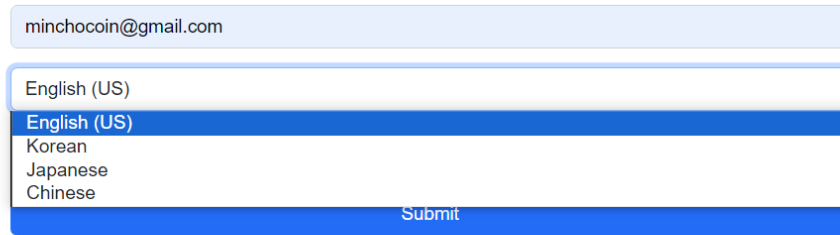
CLOUD-TERM Home Post 선택된 파일 없음

(사진3 : 전환한 모습. 위쪽부터 이메일, 언어(영어,한국어, 중국어 일본어), 파일을 선택하고, Submit 버튼을 누르면 된다.)

CLOUD-TERM Home **Post** 006_DAY 005.mp3

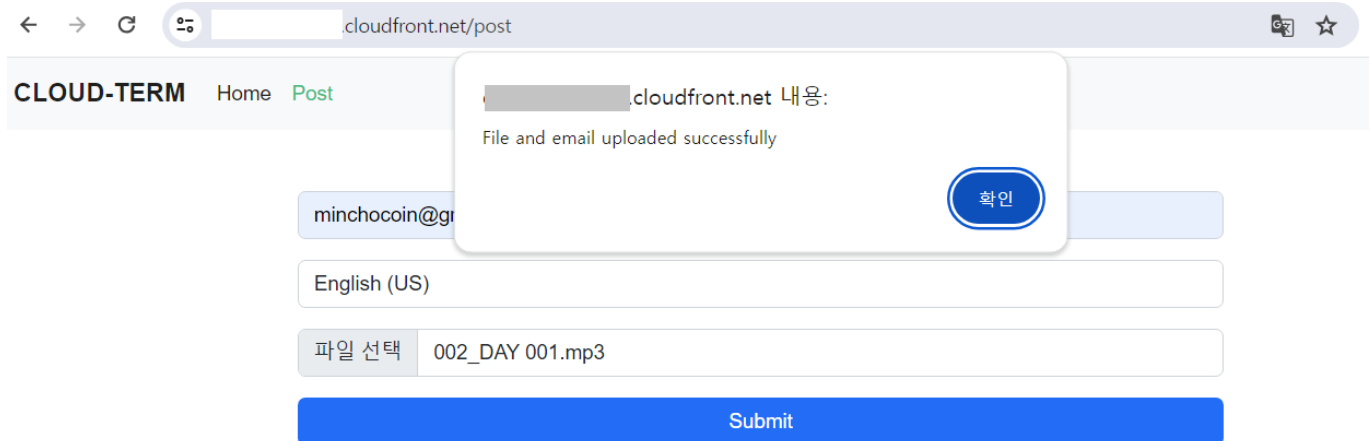
(사진4 : 임의로 입력한 모습)

CLOUD-TERM Home Post



A screenshot of a web form for language selection. At the top is a text input field containing 'minchocoin@gmail.com'. Below it is a dropdown menu currently showing 'English (US)'. The dropdown is open, showing a list of options: 'English (US)', 'Korean', 'Japanese', and 'Chinese'. At the bottom of the form is a blue 'Submit' button.

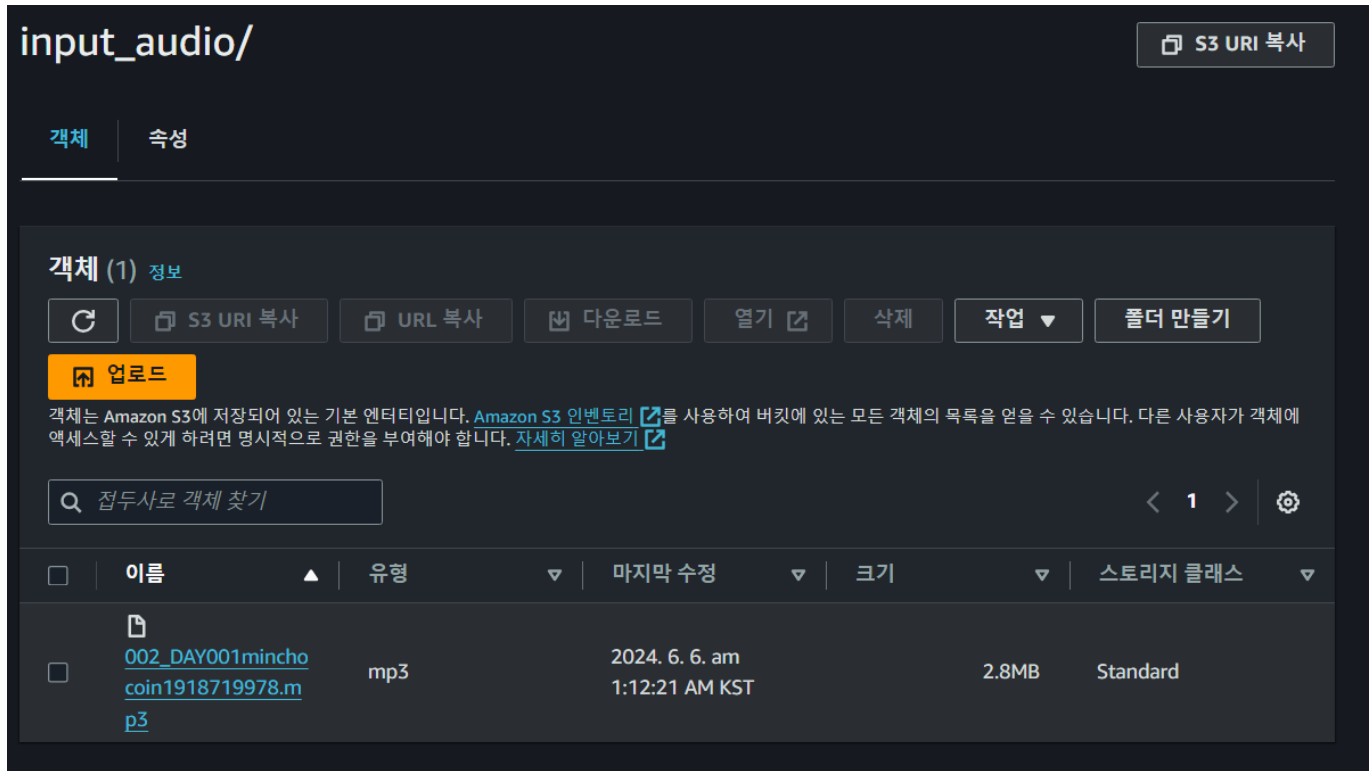
(사진5 : 지원되는 언어는 영어(미국), 한국어, 일본어, 중국어 이다.)



A screenshot of the Cloud-Term application interface. At the top, there's a navigation bar with 'CLOUD-TERM', 'Home', and 'Post'. Below it, a modal dialog box is displayed with the title '[redacted].cloudfront.net 내용:' and the message 'File and email uploaded successfully'. To the right of the message is a blue button labeled '확인' (Confirm). Below the modal, the form from the previous screenshot is visible, but the 'Submit' button is now a solid blue button. The text input field still contains 'minchocoin@gmail.com', the language dropdown is set to 'English (US)', and the file selection field shows '파일 선택' and '002_DAY 001.mp3'.

(사진6 : Submit 버튼을 누르면 잠시 후 File and email uploaded successfully 알람이 뜬다. 프론트엔드에서 AWS API GateWay 주소로 formdata 형식으로 요청을 보낸다. API GateWay는 file_upload_backend 함수와 연결되어 있으며, 이 함수는 s3에 파일을 업로드한다. 람다 함수의 특성상 약 6MB 이상의파일은 업로드할수 없다.)

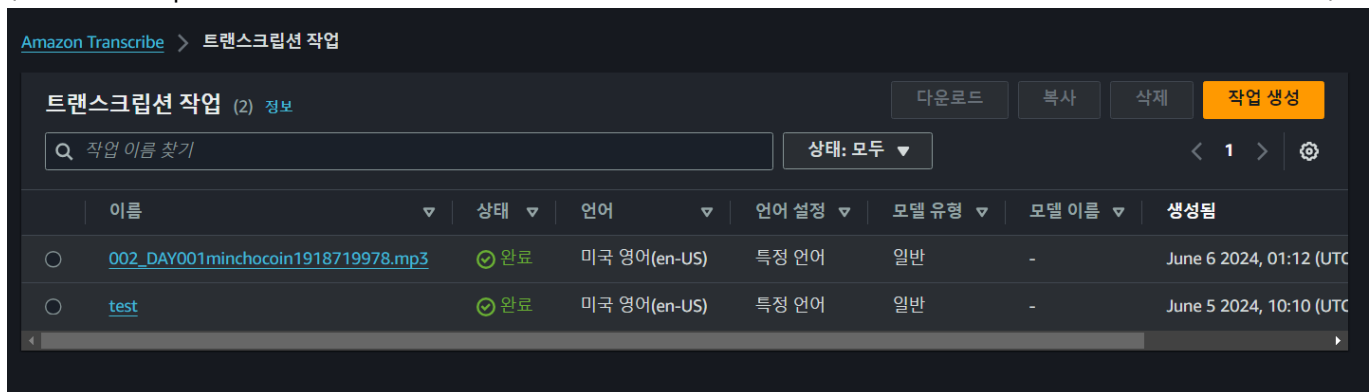
백엔드



(사진7 : s3://bucket/input_audio에 업로드한 오디오파일이 저장된다. 이름 중복을 막기 위해 padding 및 nonce를 추가하여 저장한다.)

email	filename	language	time
minchocoin@gmail.com	002_DAY001minchocoin1918719978.mp3	en-US	2024-06-05 16:12:19

(사진8 : file_upload_backend 함수는 파일 이름과 그에 해당하는 이메일 주소 및 언어를 RDS DB에 저장한다.)



(사진9 : s3에 파일이 업로드되면, audio_transcribe 람다 함수가 호출되고, 이 함수는 transcribe 서비스에 변환 작업을 생성한다(RDS DB에서 언어 코드를 받아온다).)

Amazon Transcribe > 트랜스크립션 작업 > 002_DAY001minchocoin1918719978.mp3

002_DAY001minchocoin1918719978.mp3

정보

삭제복사

작업 세부 정보

<div>이름</div> <div>002_DAY001minchocoin1918719978.mp3</div>	<div>모델</div> <div>없음</div>	<div>오디오 식별</div> <div>끄기</div>	<div>입력 데이터 위치</div> <div></div>
<div>상태</div> <div>완료</div>	<div>생성됨</div> <div>2024. 6. 6. 오전 1:12:27</div>	<div>대체 결과</div> <div>끄기</div>	<div>001minchocoin1918719978.mp3</div> <div></div>
<div>언어</div> <div>미국 영어(en-US)</div>	<div>시작됨</div> <div>2024. 6. 6. 오전 1:12:27</div>	<div>사용자 지정 단어집</div> <div>없음</div>	<div>출력 데이터 위치</div> <div>서비스 관리형 S3 버킷</div>
<div>언어 설정</div> <div>특정 언어</div>	<div>종료됨</div> <div>2024. 6. 6. 오전 1:12:37</div>	<div>PII 편집</div> <div>끄기</div>	
<div>만료 정보</div> <div>트랜스크립션은 89일 더 사용할 수 있습니다.</div>	<div>입력 파일 형식</div> <div>mp3</div>	<div>단어집 필터</div> <div>-</div>	
	<div>오디오 샘플링 속도</div> <div>44100 Hz</div>		

(사진10 : 위와 같이 작업이 생성된다.)

트랜스크립션 미리 보기

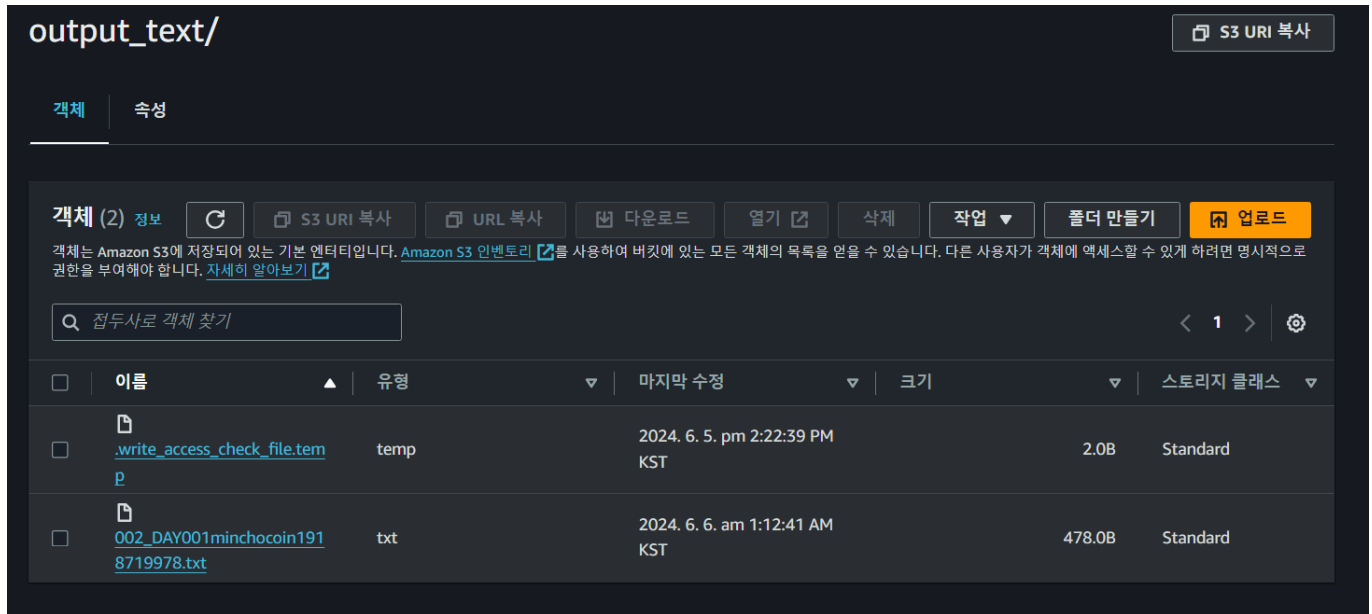
다운로드

[다운로드]를 선택하여 트랜스크립션 로컬 사본을 저장합니다.

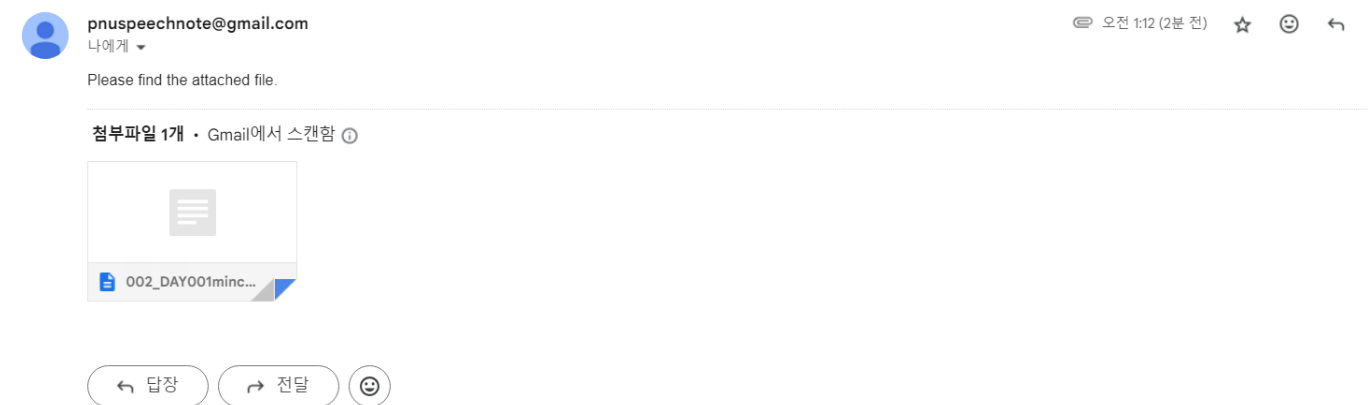
텍스트오디오 식별자막

Day one. Do you have any seats available? I'm afraid all flights are fully booked. Hello. Is this United Airlines? I'd like to book a flight to L A. I want to reserve a flight to London. Are there any seats available? How much is a ticket to L A? Are there any seats left for Friday? Do you have a direct flight to Sydney? That flight is fully booked? I'm sorry, but that flight is booked up. All seats are booked for July. Can I go stand by? Can you put me on the waiting list?

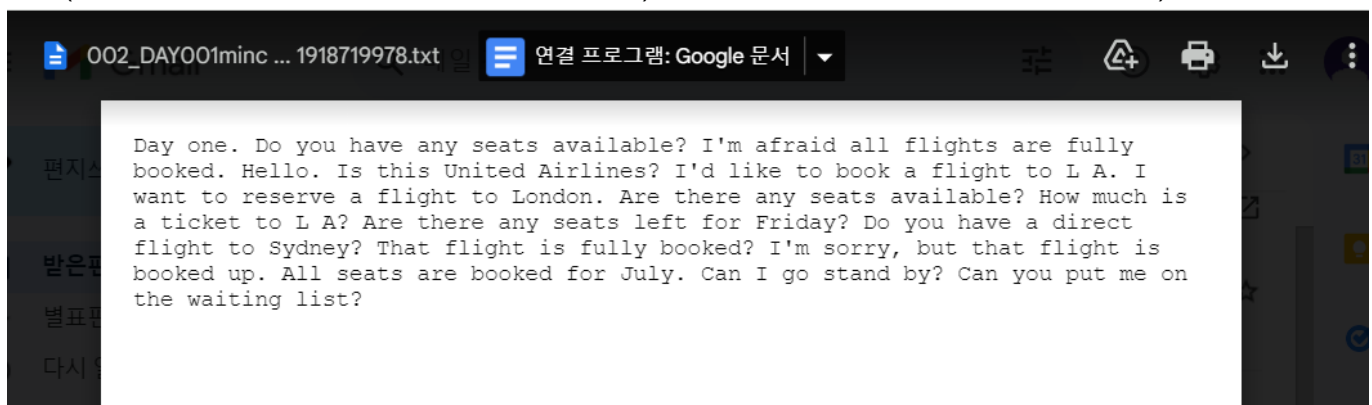
(사진11 : 작업이 완료되면 AWS transcribe 서비스 페이지에서 작업 결과를 볼 수 있다.)



(사진12 : 작업이 완료되면 eventbridge와 cloudwatch가 이를 감지하여 save_text 함수를 호출한다. 이 함수는 변환 결과를 transcribe서비스로 부터 받아서 s3에 저장한다. 즉 s3://bucket/output_text에 저장한다)



(사진13 : s3에 output_text 밑에 txt 파일이 저장되면, send_email 람다 함수가 호출된다. 이 함수는 저장된 txt 파일(파일 이름은 업로드된 음성파일의 이름과 동일)을 DB에 저장된 이메일 주소로 전송한다.)



(사진14 : 위와 같이 변환 결과를 이메일로 받을 수 있다.) 아래는 변환 결과 예시이다.

```
Day one. Do you have any seats available? I'm afraid all flights are fully booked.
Hello. Is this United Airlines? I'd like to book a flight to L A. I want to
reserve a flight to London. Are there any seats available? How much is a ticket
to L A? Are there any seats left for Friday? Do you have a direct flight to
Sydney? That flight is fully booked? I'm sorry, but that flight is booked up.
All seats are booked for July. Can I go stand by? Can you put me on the waiting
list?
```

주요 AWS 설정 설명

자세한 것은 [시스템구축.pdf](#)을 참고하세요.

s3 정적 웹사이트 호스팅

정적 웹 사이트 호스팅 편집 정보 정보

정적 웹 사이트 호스팅

이 버킷을 사용하여 웹 사이트를 호스팅하거나 요청을 리디렉션합니다. [자세히 알아보기](#)

정적 웹 사이트 호스팅

☐ 비활성화

☒ 활성화

호스팅 유형

☒ 정적 웹 사이트 호스팅

버킷 엔드포인트를 웹 주소로 사용합니다. [자세히 알아보기](#)

☐ 객체에 대한 요청 리디렉션

요청을 다른 버킷 또는 도메인으로 리디렉션합니다. [자세히 알아보기](#)

❗ 고객이 웹 사이트 엔드포인트의 콘텐츠에 액세스할 수 있게 하려면 모든 콘텐츠를 공개적으로 읽기 가능하도록 설정해야 합니다. 이렇게 하려면, 버킷에 대한 S3 퍼블릭 액세스 차단 설정을 편집하면 됩니다. 자세한 내용은 [Amazon S3 퍼블릭 액세스 차단 사용](#) 참조하십시오.

인덱스 문서

웹 사이트의 홈 페이지 또는 기본 페이지를 지정합니다.

index.html

오류 문서 - 선택 사항

오류가 발생하면 반환됩니다.

error.html

웹사이트를 호스팅할 s3를 하나 만들고, 정적 웹사이트 호스팅을 하도록 설정한다. 접속하면 index.html을 반환하도록 한다.

cloudfront 설정

동작 편집

설정

경로 패턴 | 정보

기본값(*)

원본 및 원본 그룹

s3-website.ap-northeast-2.amazonaws.com ▼

자동으로 객체 압축 | 정보

- ☐ No
☒ Yes

뷰어

뷰어 프로토콜 정책

- ☐ HTTP and HTTPS
☒ Redirect HTTP to HTTPS
☐ HTTPS only

허용된 HTTP 방법

- ☒ GET, HEAD
☐ GET, HEAD, OPTIONS
☐ GET, HEAD, OPTIONS, PUT, POST, PATCH, DELETE

뷰어 액세스 제한

뷰어 액세스를 제한하는 경우 뷰어는 CloudFront 서명된 URL 또는 서명된 쿠키를 사용하여 사용자의 콘텐츠에 액세스해야 합니다.

- ☒ No
☐ Yes

캐시 키 및 원본 요청

캐시 정책 및 원본 요청 정책을 사용하여 캐시 키 및 원본 요청을 제어할 것을 권장합니다.

- ☒ Cache policy and origin request policy (recommended)
☐ Legacy cache settings

캐시 정책

기존 캐시 정책을 선택하거나 새 캐시 정책을 생성합니다.

CachingOptimized

Policy with caching enabled. Supports Gzip and Brotli compression.

Recommended for S3 ▼



cloudfront를 사용하여 s3에 있는 정적 웹사이트를 호스팅한다. cloudfront를 사용하면 https가 적용된다.

API GateWay 생성 및 람다 함수 연결


메서드 생성

메서드 세부 정보

메서드 유형
메서드 유형 선택 ▼

통합 유형

☒ **Lambda 함수**
API를 Lambda 함수와 통합합니다.


☐ **HTTP**
기존 HTTP 엔드포인트와 통합합니다.


☐ **Mock**
API Gateway 매핑 및 변환을 기반으로 응답을 생성합니다.


☐ **AWS 서비스**
AWS 서비스와 통합합니다.


☐ **VPC 링크**
퍼블릭 인터넷을 통해 액세스할 수 없는 리소스와 통합합니다.


Lambda 프록시 통합
요청을 구조화된 이벤트로 Lambda 함수에 전송합니다.

Lambda 함수
Lambda 함수 이름 또는 ARN

ap-northe... ▼

arn:aws:lambda:ap-northeast-1:function:file-upload-backend
arn:aws:lambda:ap-northeast-1:function:file-upload-backend
arn:aws:lambda:ap-northeast-1:function:send_email
arn:aws:lambda:ap-northeast-1:function:audio_transcribe

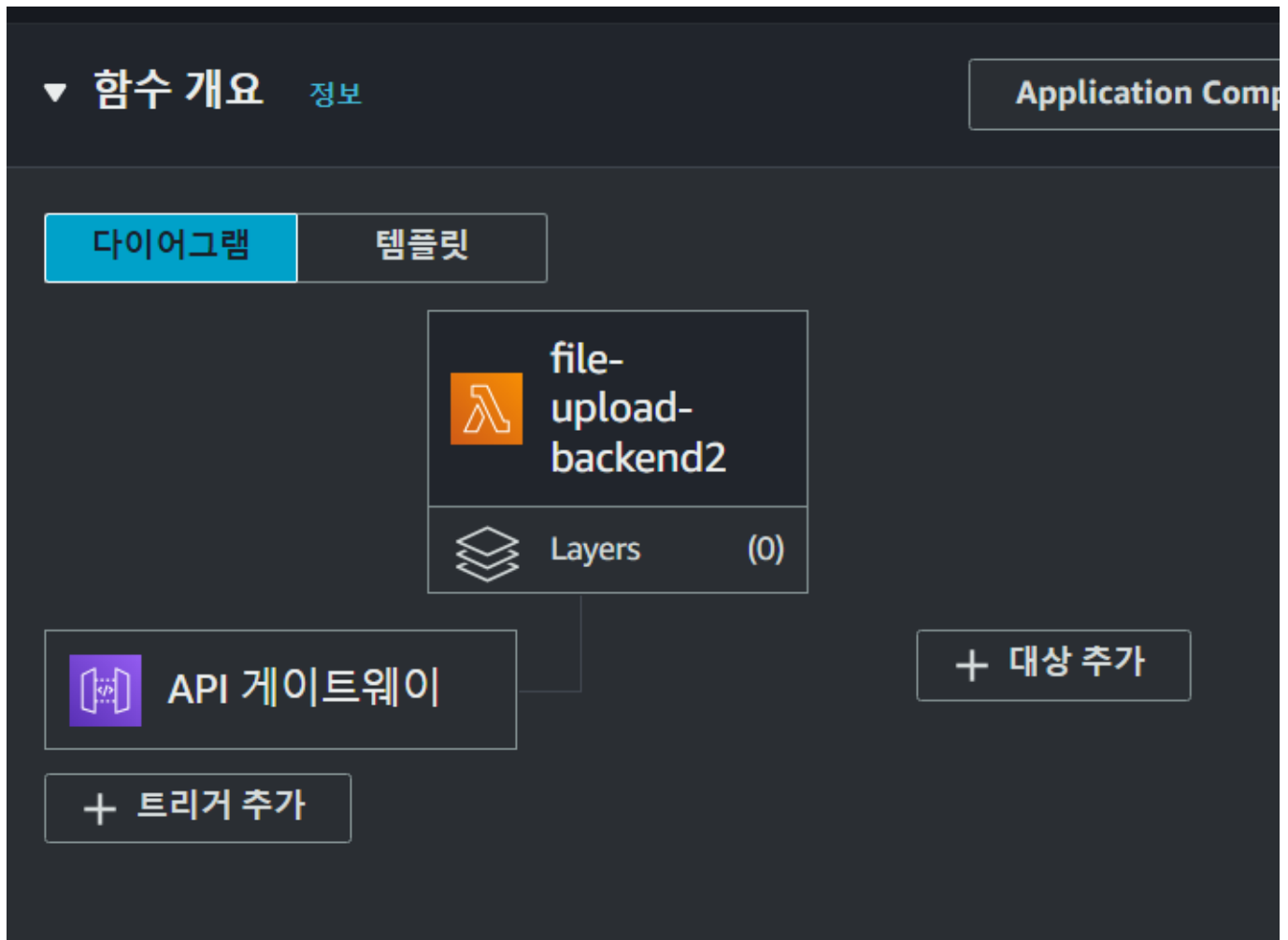
Q arn:aws:lambda:ap-northeast-1:function: X

① API Gateway에 Lambda 함수를 간접적으로 호출할 권한을 부여합니다. 이 기능을 끄려면 함수의 리소스 정책을 직접 업데이트하거나 API Gateway가 함수를 간접적으로 호출하는 데 사용하는 호출 역할을 제공하세요.

☒ **기본 제한 시간**
기본 제한 시간은 29초입니다.

☒ **Lambda 프록시 통합**
요청을 구조화된 이벤트로 Lambda 함수에 전송합니다.

API GateWay를 생성한다. API GateWay는 람다 함수와 연결하고, lambda 프록시 통합을 설정하여 formdata형식의 데이터를 정상적으로 받을 수 있도록 한다.



file_upload_backend 람다 함수는 API GateWay에 의해 트리거 된다.

s3 생성 및 람다 함수 연결

이벤트 알림 생성 정보 정보

알림을 활성화하려면 먼저 Amazon S3에서 게시할 이벤트와 Amazon S3에서 알림을 전송할 대상을 식별하는 알림 구성을 추가해야 합니다.

일반 구성

이벤트 이름

 이벤트 이름은 255자 이내여야 합니다.

접두사 - 선택 사항
 지정된 문자로 시작하는 키가 있는 객체로 알림을 제한합니다.

접미사 - 선택 사항
 지정된 문자로 끝나는 키가 있는 객체로 알림을 제한합니다.

이벤트 유형

알림을 수신할 이벤트를 하나 이상 지정합니다. 각 그룹에 대해 모든 이벤트에 적용되는 이벤트 유형을 선택하거나 하나 이상의 개별 이벤트를 선택할 수 있습니다.

객체 생성

☐ 모든 객체 생성 이벤트
 s3:ObjectCreated:*

☒ 전송
 s3:ObjectCreated:Put

☐ 게시

s3 bucket에 /input_audio 밑에 .mp3 파일이 업로드(put)되면 람다 함수를 호출하는 것으로 설정한다.

대상

① Amazon S3가 대상에 메시지를 게시하려면, 먼저 관련 API를 호출하여 SNS 주제, SQS 대기열 또는 Lambda 함수에 메시지를 게시하는 데 필요한 권한을 Amazon S3 보안 주체에 부여해야 합니다. [자세히 알아보기](#)

대상
 이벤트를 게시할 대상을 선택합니다. [자세히 알아보기](#)

☒ **Lambda 함수**
 S3 이벤트를 기반으로 Lambda 함수 스크립트를 실행합니다.

☐ SNS 주제
 연결 처리를 위해 시스템으로 메시지를 편아웃하거나 사람에게 직접 메시지를 편아웃합니다.

☐ SQS 대기열
 서버에서 읽을 수 있도록 SQS 대기열로 알림을 전송합니다.

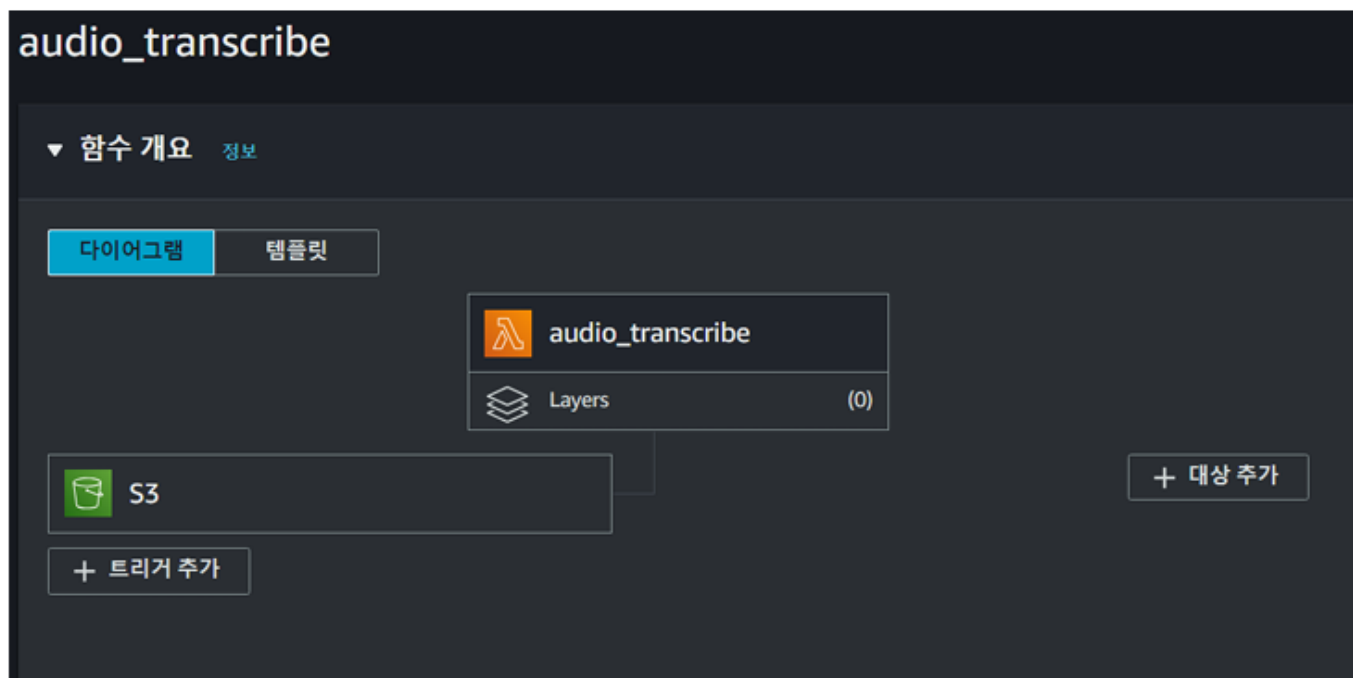
Lambda 함수 지정

☒ Lambda 함수에서 선택

☐ Lambda 함수 ARN 입력

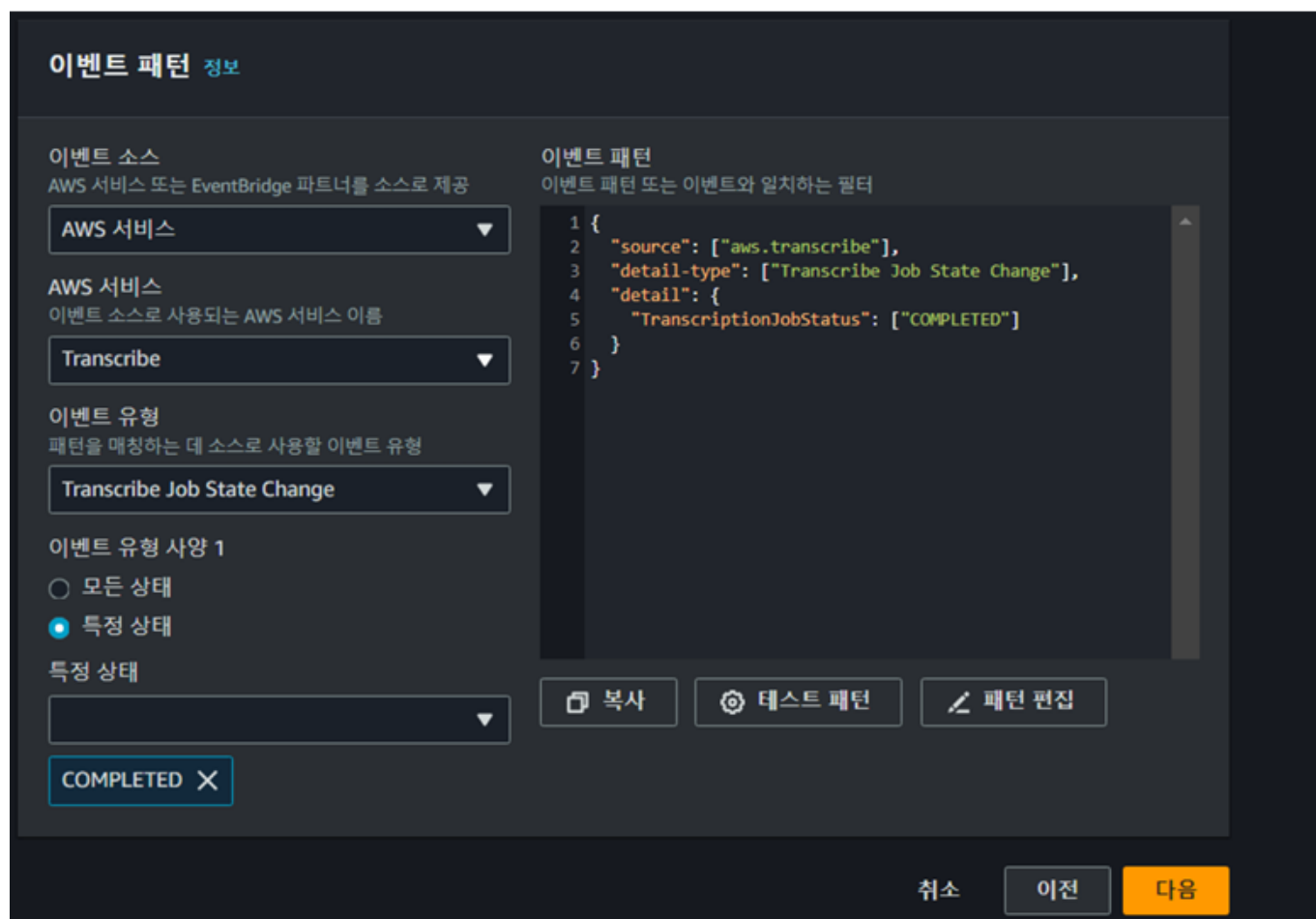
Lambda 함수

호출할 람다 함수는 업로드된 오디오파일을 transcribe에 변환 요청하는 audio_transcribe 람다 함수로 설정한다.



s3에 의해 audio_transcribe가 트리거 됨을 알 수 있다.

eventbridge 설정



eventbridge에서 transcribe 작업이 완료됨을 감지하도록 설정한다.

대상 선택



권한

참고: EventBridge 콘솔을 사용할 때, EventBridge는 선택된 대상에 대한 적절한 권한을 자동으로 구성합니다. AWS CLI, SDK 또는 CloudFormation을 사용하는 경우 적절한 권한을 구성해야 합니다.

대상 1개

대상 유형

EventBridge 이벤트 버스, EventBridge API 대상(SaaS 파트너) 또는 다른 AWS 서비스를 대상으로 선택합니다.

- ☐ EventBridge 이벤트 버스
- ☐ EventBridge API 대상
- ☒ AWS 서비스

대상 선택 | 정보

이벤트가 이벤트 패턴과 일치하거나 일정이 트리거될 때 호출할 대상을 선택합니다(규칙당 5개의 대상으로 제한).

Lambda 함수

함수

save_text



해당 이벤트가 발생하면 save_text 람다 함수를 호출하도록 설정한다.

save_text

▼ 함수 개요 정보

다이어그램

템플릿



save_text



Layers

(0)



EventBridge(CloudWatch Events)

+ 대상 추가

+ 트리거 추가

코드

테스트

모니터링

구성

별칭

버전

이제 eventbridge에 의해 save_text가 호출됨을 알 수 있다.

핵심 코드

```
transcribeClient = boto3.client('transcribe')

try:
    transcribeClient.start_transcription_job(
        TranscriptionJobName=jobName,
        LanguageCode=languageCode,
        MediaFormat='mp3',
        Media={
            'MediaFileUri': f's3://{bucketName}/{baseFileName}'
        },
    )
    ...
except Exception as e:
    ...
```

python의 boto3 라이브러리를 이용하여 transcription job을 생성한다. jobname과 언어코드, 포맷(mp3), 음성 파일 위치를 전달하여 생성한다.

참고 : save_text 및 send_email 함수 설명

- save_text 설명 : [save_text/README.md](#)
- send_email 설명 : [send_email/README.md](#)

사용 방법

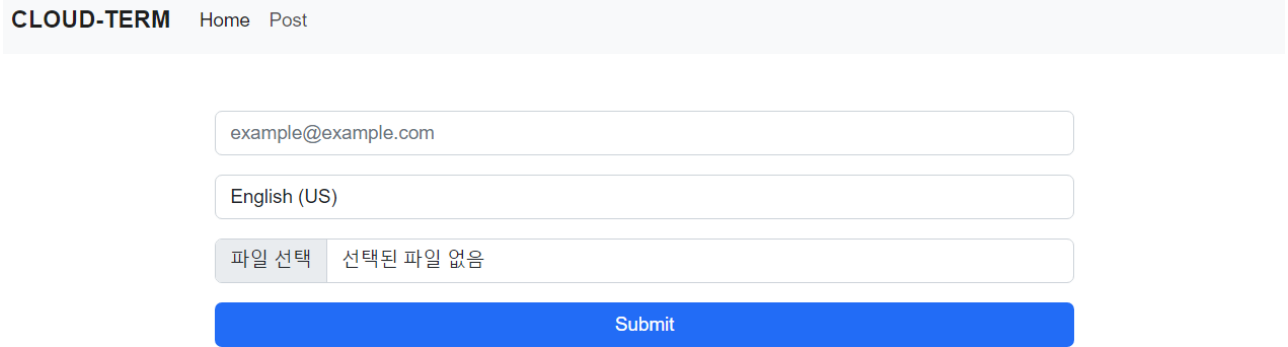
1. 먼저 <https://d3yte1ymgtgsl.cloudfront.net/> 에 접속하세요. 접속하면 아래와 같이 나옵니다.



Welcome to PSQLI service!

This is the service to convert audio to text
If you want to submit your own problems, send email to admin(admin@example.com)

2. 왼쪽 상단에 post를 누르면 아래 화면과 같이 이메일 주소와 파일을 업로드할 수 있는 form이 나옵니다.



CLOUD-TERM Home Post

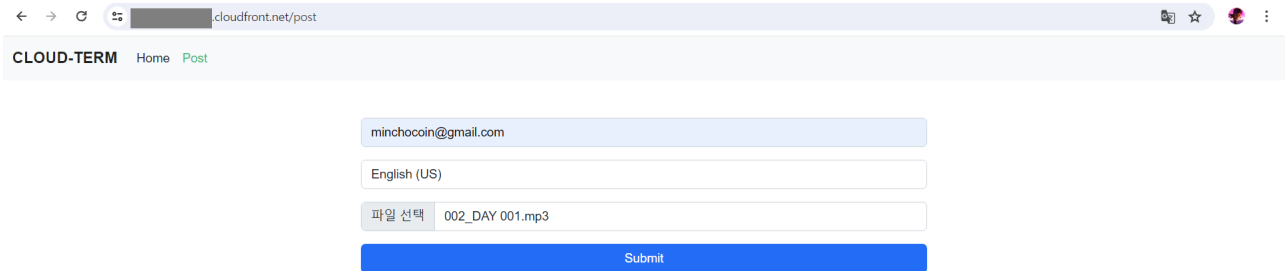
example@example.com

English (US)

파일 선택 선택된 파일 없음

Submit

3. 이메일 주소와 언어, 파일을 선택합니다. 람다 함수의 특성상 6MB 이상의 파일은 업로드 할수 없습니다.



CLOUD-TERM Home Post

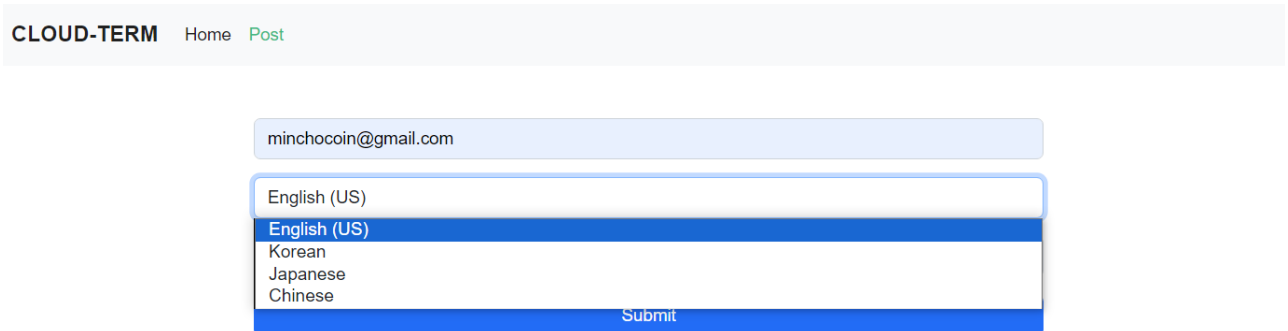
minchocoin@gmail.com

English (US)

파일 선택 002_DAY 001.mp3

Submit

4. 사용할 수 있는 언어는 영어(미국), 한국어, 일본어, 중국어 입니다.



CLOUD-TERM Home Post

minchocoin@gmail.com

English (US)

English (US)

Korean

Japanese

Chinese

Submit

5. Submit 버튼을 누르면 잠시 후 File and email uploaded successfully 알람이 발생합니다. 변환이 완료되고 이메일로 전송되기 까지 시간이 약간 걸릴 수 있습니다.

← → ↺ d3yte1ymgtgsl.cloudfront.net/post

CLOUD-TERM Home Post

d3yte1ymgtgsl.cloudfront.net 내용:
File and email uploaded successfully

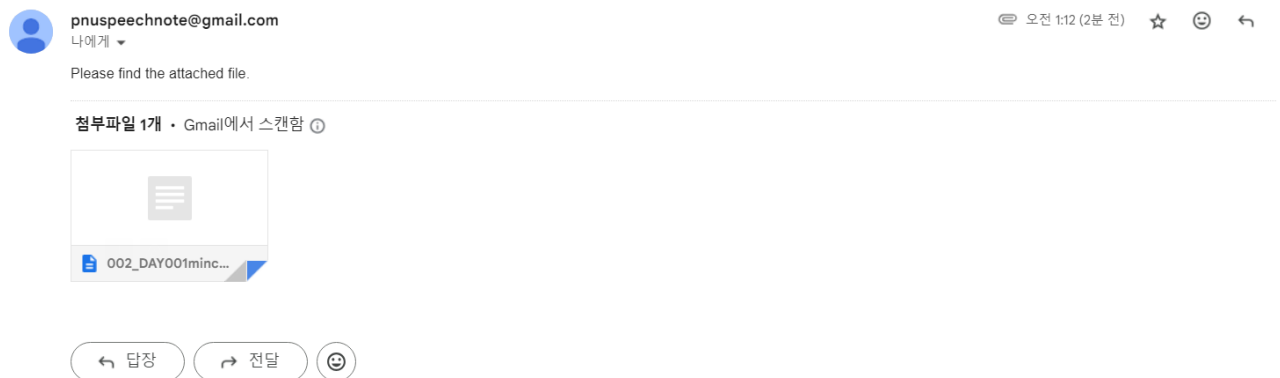
minchocoin@gr...

English (US)

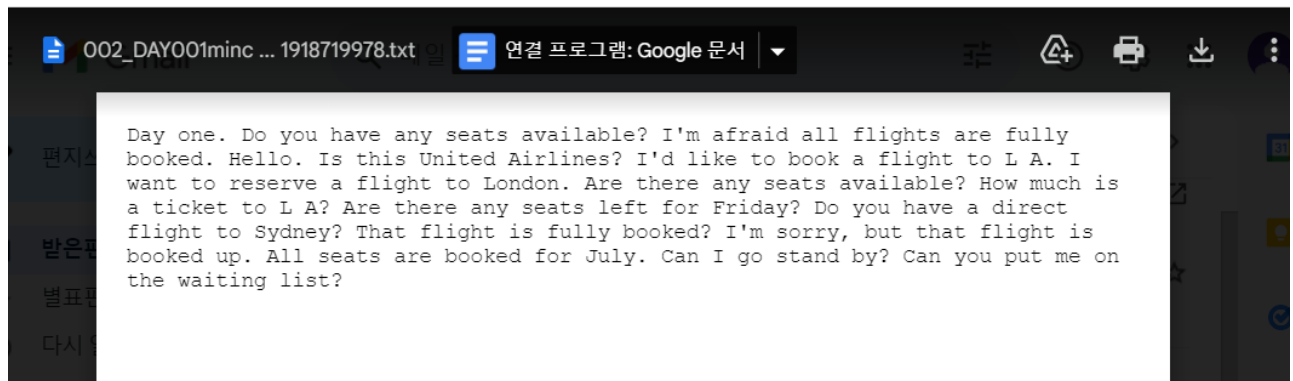
파일 선택 002_DAY 001.mp3

Submit

6. 제출한 이메일로 파일이 전송됩니다. 텍스트파일 형태로 전송됩니다.



7. 전송된 텍스트 파일에서 변환된 결과를 확인할 수 있습니다.



아래는 변환 결과 예시이다.

Day one. Do you have any seats available? I'm afraid all flights are fully booked. Hello. Is this United Airlines? I'd like to book a flight to L A. I want to reserve a flight to London. Are there any seats available? How much is a ticket to L A? Are there any seats left for Friday? Do you have a direct flight to Sydney? That flight is fully booked? I'm sorry, but that flight is booked up. All seats are booked for July. Can I go stand by? Can you put me on the waiting list?

활용방안

1. 강의 녹음을 글로 변환 강연이나 강의 녹음을 들으면서 정리할 필요없이 이 서비스를 이용해 글로 변환 하여 읽으면서 복습할 수 있습니다.

[강연 노트 예시](#)를 변환해보겠습니다. 동작 영상은 [link](#)에서 볼 수 있습니다.

아래는 이메일로 전송된 파일 예시입니다.(일부 후략)(lecturenote2minchocoin576724248.txt)

인공 지능은 일반적으로 인간의 지능이 필요하거나 인간이 분석할 수 있는 것보다 규모가 큰 데이터를 포함하는 방식으로 추론, 학습 및 행동할 수 있는 컴퓨터 및 기계를 구축하는 것과 관련된 과학 분야입니다. 먼저, 인공지능 학습 모델에 대해 말씀드리겠습니다. 비즈니스에서 AI 를 이야기할 때 학습 데이터에 관해 이야기하는 경우가 많습니다. 그것은 무엇을 의미할까요? 제한된 메모리 인공지능은 새로운 데이터로 학습 하여 시간이 지남에 따라 개선되는 AI 입니다. 머신러닝 은 알고리즘으로 데이터를 학습 하여 결과를 얻는 인공 지능의 허위 집합 입니다. 일반적으로 머신러닝 에는 세 가지 유형의 학습 모델이 사용됩니다. 지도 학습은 라벨이 지정된 학습 데이터를 사용하여 특정 입력을 출력에 매칭 하는 머신러닝 모델입니다. 간단히 말해, 고양이 사진을 인식 하도록 알고리즘을 훈련 시키려면 고양이라는 라벨이 지정된 사진을 필드 합니다 비 지도 학습은 라벨이 지정되지 않은 데이터를 기반으로 패턴을 학습하는 머신러닝 모델입니다.

2. 음성 메모 메모할 내용을 녹음하고, 이 서비스를 이용해 텍스트로 변환하여 타이핑 없이 메모할 수 있습니다.

[음성 메모 예시](#)를 변환해보겠습니다. 동작 영상은 [여기](#)서 볼 수 있습니다.

아래는 이메일로 전송된 파일 예시입니다.

올해 목표 메모 육 개월 이내에 오픽 IH 이상 달성 한다. 매일 한 시간 이상 회화 연습을 하고 주 삼회 영어 회화 수업에 참여한다. 이를 위해 오픽 책과 영어, 회화, 책, 영어 단어, 책을 구매한다. 단어는 하루에 서른 개를 외우는 것을 목표로 한다. 그리고 개인 블로그를 개설하여 학교에서 배운 개념을 정리한다. 잘 이해가 되지 않는 운영 체제는 컴퓨터 구조 위주로 정리하여 일주일에 한 개 이상 포스팅 하는 것을 목표로 한다.

개선 방향

- 앱으로 만들어 파일 업로드 등의 관점에서 접근성을 높인다.
- 24시간 또는 1개월 당 변환할 수 있는 음성 파일 시간을 제한한다.
 - 이를 위해 로그인 기능을 구현하여 계정마다 시간을 제한한다.
- 이메일을 보내는 방식이 아니라 웹에서 바로 확인할 수 있도록 한다.
- clova note와 같이 AI 요약 서비스 제공
- 사용자 지정 단어집 제공
- 파일을 s3에 바로 업로드하여 길이제한 없애기 : [link](#)
 - 현재는 약 6MB의 제한이 있음

references

- <https://devocean.sk.com/blog/techBoardDetail.do?ID=163934>
- <https://aws.amazon.com/ko/what-is/serverless-computing/>
- <https://docs.aws.amazon.com/transcribe/latest/dg/getting-started.html>

- cloudfront 사용 : [link](#)
- 파일 업로드 웹 페이지 구축 : [link](#)
- s3 trigger lambda(and lambda call transcribe) : [link](#)
- lambda language code : [link](#)
- transcribe API guide: [link](#)
- sending email : [link](#)
- using CloudWatch : [link](#)
- connect lambda and RDS DB : [link](#)
 - <https://jane-aeiou.tistory.com/51>
- google SMTP 사용 : <https://icedhotchoco.tistory.com/entry/sending-email-with-python>
- vpc ipv6 setting : <https://hossamelshahawi.com/2021/10/24/fix-errno97-address-family-not-supported-by-protocol/>

License

[Apache License v2.0](#)