



AWS Transcribe를 사용한 SpeechNote 웹서비스

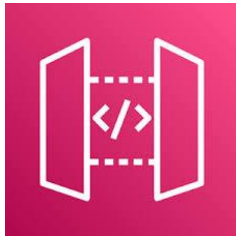
2024학년도 1학기 클라우드컴퓨팅 텀프로젝트

7조. 김태훈, 박민재, 김성문, 박재열



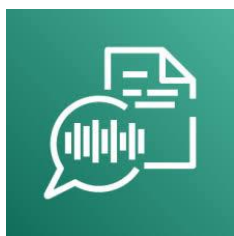
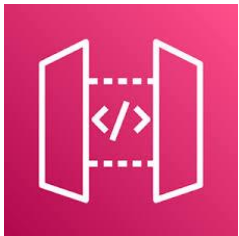
프로젝트 소개

- AWS Transcribe를 사용한 음성을 글자로 변환(Speech-To-Text)해주는 웹사이트 제작
- AWS Transcribe를 사용한 음성노트(SpeechNote) 웹사이트 제작
- 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.
- 사용된 AWS 서비스: S3, CloudFront, API GateWay, Lambda, Transcribe, Eventbridge, RDS



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.
- 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(클로바노트)

- 네이버에서 개발한 클로바노트(CLOVA note)는 웹이나 모바일 앱에서 음성을 녹음하거나 음성파일을 업로드하면,
- 그것을 텍스트로 변환해주고, 이것을 시간과 화자와 함께 보여준다.
- 녹음 파일 업로드 시 월 300분의 제한,
- 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 SI요약 기능도 있다.

The screenshot displays the CLOVA Note application interface. On the left is a sidebar menu with options: 홈 (Home), 전체 노트 (All Notes), 주간 회의 (Weekly Meeting), 리서치 (Research), 공유 받은 노트 (Shared Notes), 미팅 (Meeting), 상담 (Consultation), 공유한 노트 (Shared Notes), and 휴지통 (Trash). The main area shows a list of notes under the '음성 기록' (Audio Record) tab. Each note entry includes a speaker icon, a name, a timestamp, and a text preview. The notes are as follows:

- 최민아 00:10**: 사람들이 보편적으로 접할 수 있는 플리마켓, 버스킹 문화를 참고해서 문화 마켓이라는 주제로 이벤트를 준비해보려고 합니다.
- 박규리 00:20**: 문화 마켓에서 책이나 빈티지 상품을 판매하거나, 먹거리와 음악 공연을 준비해서 진행하는 것은 어떨까요?
- 이세진 00:22**: 지난 주에 진행한 이벤트 기획 회의에서 다양한 의견이 나왔는데요. 사람들의 참여를 많이 이끌어내는 것에 중점을 두고 기획팀에서 주제를 조금 더 구체화해보았습니다.
- 박규리 00:24**: 사람들이 보편적으로 접할 수 있는 플리마켓, 버스킹 문화를 참고해서 문화 마켓이라는 주제로 이벤트를 진행해 보고자 합니다.
- 홍주연 00:26**: 세진님 구체적인 예시도 공유 부탁드립니다.
- 박규리 00:27**: 책이나 빈티지 상품은 사람들이 관심 있어 할 만한 물건이고, 먹거리와 음악 공연은 사람들을 즐겁게 할 수 있는 요소이니깐요.

역할 분담

이름	역할
김태훈*	AWS 전체 구축 파일 업로드 람다 함수 구현 (file_upload_backend)
박민재	이메일 전송 람다 함수 구현 (save_text, send_email)
김성문	오디오 텍스트 변환 람다 함수 구현 (audio_transcribe)
박재열	프론트엔드 구현 (HTML/CSS/JS(Vue.js))

언어코드

영어(영국) : en-GB
 영어(미국) : en-US
 한국어 : ko-KR
 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

https://docs.aws.amazon.com/transcribe/latest/APIReference/API_StartTranscriptionJob.html

User<->CloudFront

사용자가 웹 브라우저를 통해 CloudFront로 접속하면 S3로부터 배포된 정적 콘텐츠(HTML/CSS/JS, 사진 등) 전달

CloudFront <-> S3

S3에 있는 정적 콘텐츠는 CloudFront를 통해 배포됨

User<-> API GateWay

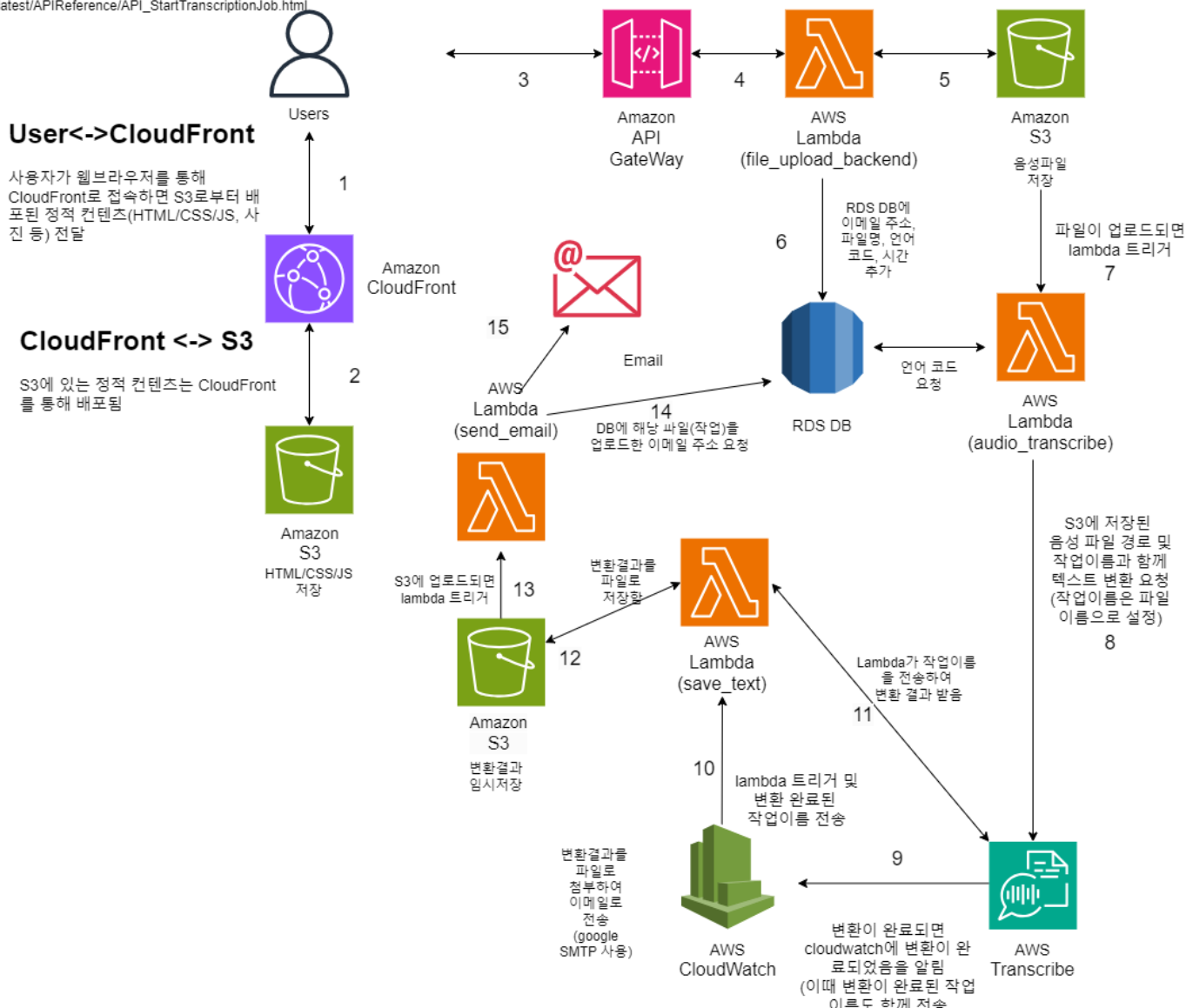
사용자는 API GateWay로 파일과 이메일 주소, 언어 코드 전달
 GateWay는 Lambda로부터 받은 응답 코드 전달

API <-> Lambda

API GateWay는 파일, 이메일 주소와 함께 Lambda 함수를 호출하여 S3에 업로드 요청 (파일명에 임의의 문자 및 메일 주소 (id-example)를 추가하여 다른 파일과 구분)
 Lambda는 API GateWay에 응답코드 반환

Lambda(file...) <-> S3

Lambda는 S3에 파일 업로드



결과물 소개



- 홈페이지



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)



결과물 소개



- 파일 업로드 화면

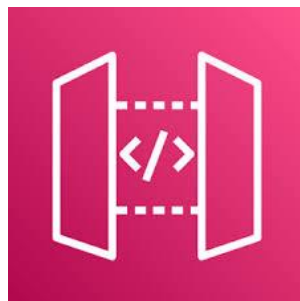
CLOUD-TERM Home Post

example@example.com

English (US)

파일 선택 선택된 파일 없음

Submit



결과물 소개



- 간단한 동작 영상 : <https://youtu.be/KWpaLcgWzTE>

CLOUD-TERM Home Post

minchocoin@gmail.com

English (US)

파일 선택 006_DAY 005.mp3

Submit

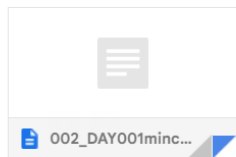


pnuspeechnote@gmail.com

나에게 ▾

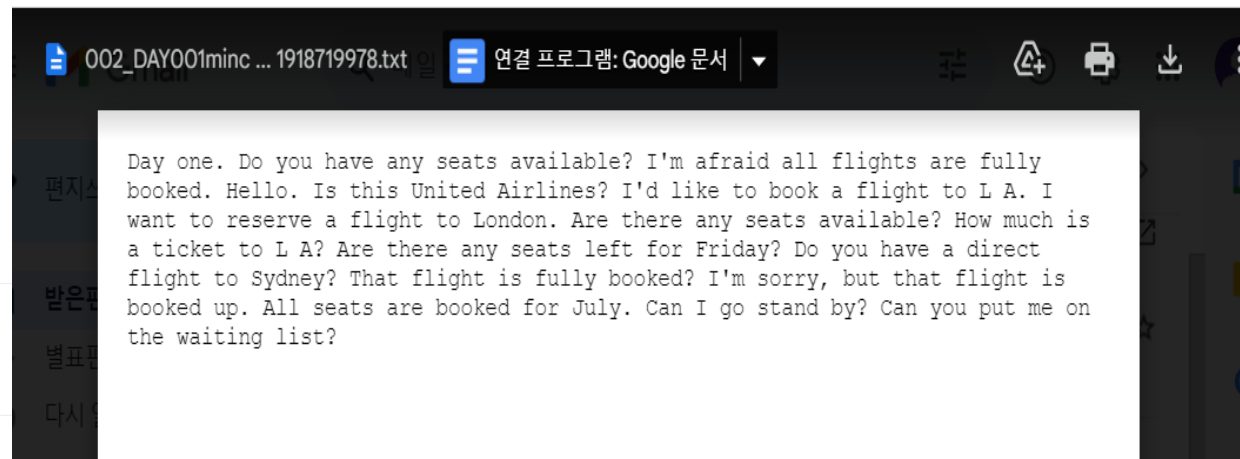
Please find the attached file.

첨부파일 1개 • Gmail에서 스캔함 ⓘ



← 답장

→ 전달



결과물 소개

- 지원되는 언어 : 영어(en-US), 한국어(ko-KR), 중국어(zh-CN), 일본어(ja-JP)

CLOUD-TERM [Home](#) [Post](#)

minchocoin@gmail.com

English (US)

English (US)

Korean

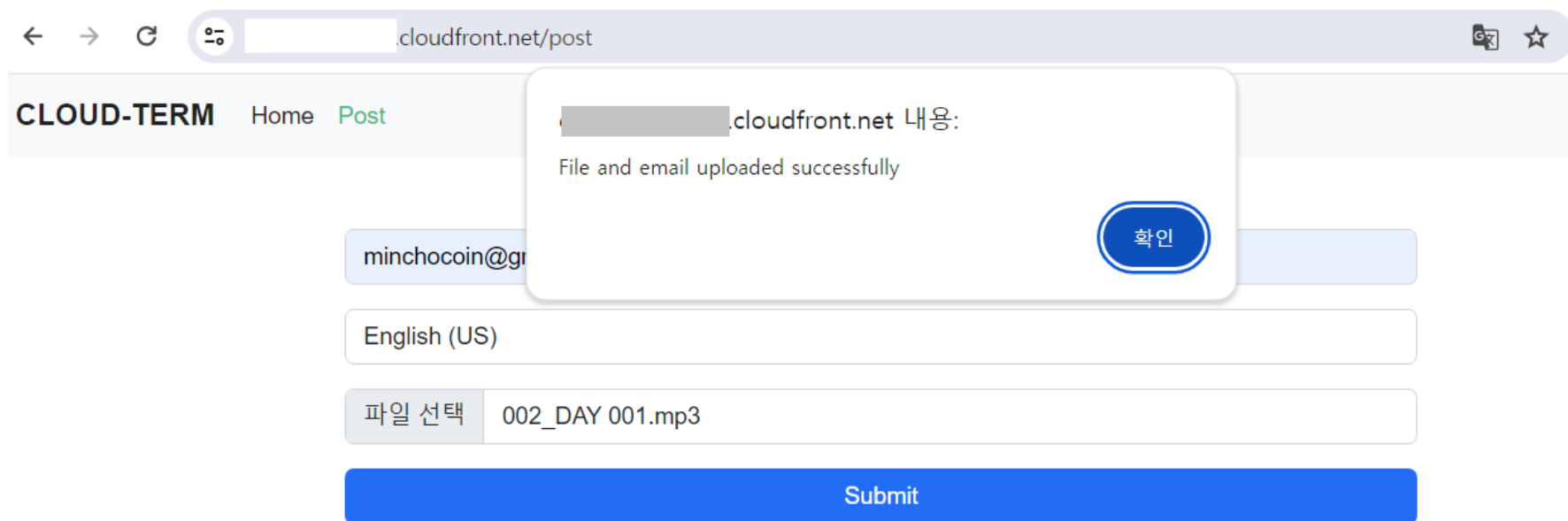
Japanese

Chinese

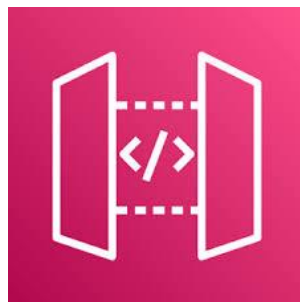
Submit

결과물 소개

- Submit 버튼 누를 시 API GateWay 주소로 전송됨

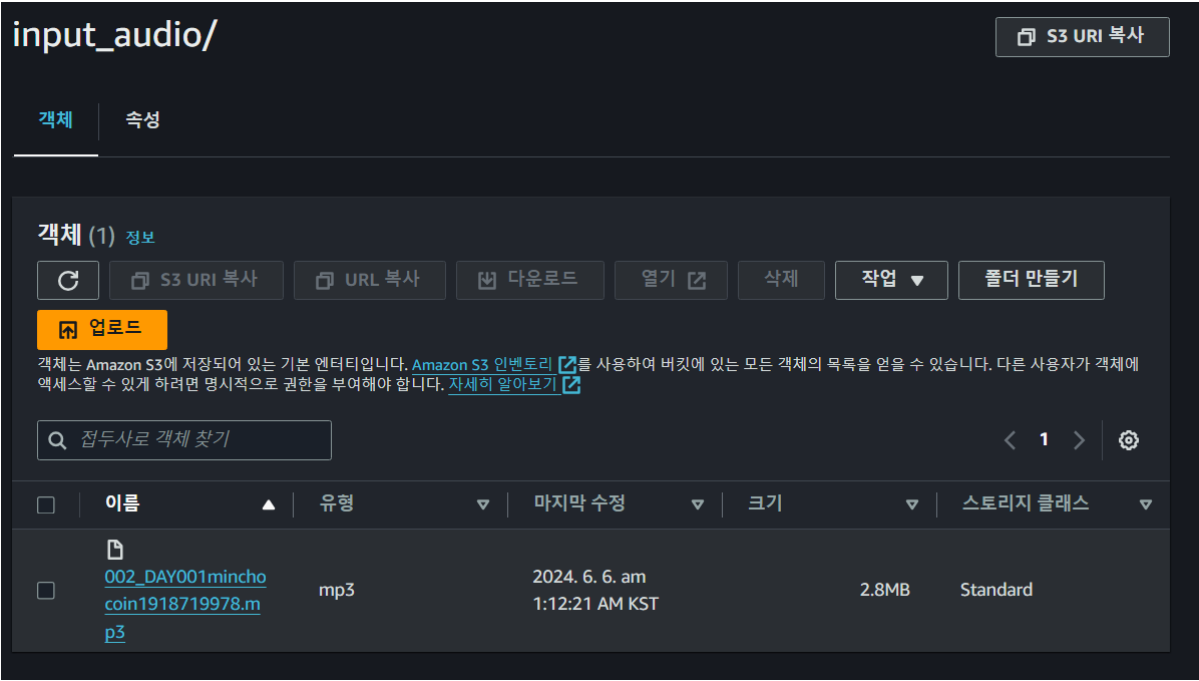


The screenshot shows a web browser window with the address bar displaying a CloudFront URL. The page header includes 'CLOUD-TERM' and navigation links for 'Home' and 'Post'. A modal dialog box is centered on the screen, displaying a success message: 'File and email uploaded successfully'. Below the message is a blue button labeled '확인' (Check/Confirm). The form fields visible include an email address 'minchocoin@gr...', a language selection dropdown set to 'English (US)', and a file selection field showing '002_DAY 001.mp3'. A large blue 'Submit' button is at the bottom of the form.



결과물 소개(내부)

- File_upload_backend 함수 : s3/bucket/input_audio에 파일을 업로드하고, 테이블에 이메일과 언어코드 기록



```
fileTable;
+-----+-----+-----+-----+
| email | filename | language | time |
+-----+-----+-----+-----+
| minchocoin@gmail.com | 002_DAY001minchocoin1918719978.mp3 | en-US | 2024-06-05 16:12:19 |
+-----+-----+-----+-----+
1 row in set (0.0105 sec)
```



결과물 소개(내부)

- Audio_transcribe함수 : AWS transcribe에 변환 요청

```
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:
    ...
```

Amazon Transcribe > 트랜스크립션 작업

트랜스크립션 작업 (2) 정보

다운로드

복사

삭제

작업 생성

🔍 작업 이름 찾기

상태: 모두 ▼

< 1 > ⚙️

	이름 ▼	상태 ▼	언어 ▼	언어 설정 ▼	모델 유형 ▼	모델 이름 ▼	생성됨
<input type="radio"/>	002_DAY001minchocoin1918719978.mp3	✅ 완료	미국 영어(en-US)	특정 언어	일반	-	June 6 2024, 01:12 (UTC)
<input type="radio"/>	test	✅ 완료	미국 영어(en-US)	특정 언어	일반	-	June 5 2024, 10:10 (UTC)



결과물 소개(내부)

- Save_text함수 : s3/bucket/output_text에 변환결과를 .txt 형태로 저장
- Send_email 함수 : DB에 저장된 이메일로 text 파일 첨부하여 전송

output_text/ S3 URI 복사

객체 속성

객체 (2) 정보 🔄 S3 URI 복사 URL 복사 다운로드 열기 삭제 작업 ▼ 폴더 만들기 업로드

객체는 Amazon S3에 저장되어 있는 기본 엔터티입니다. [Amazon S3 인벤토리](#)를 사용하여 버킷에 있는 모든 객체의 목록을 얻을 수 있습니다. 다른 사용자가 객체에 액세스할 수 있게 하려면 명시적으로 권한을 부여해야 합니다. [자세히 알아보기](#)

🔍 접두사로 객체 찾기

<input type="checkbox"/>	이름	유형	마지막 수정	크기	스토리지 클래스
<input type="checkbox"/>	.write_access_check_file.tem	temp	2024. 6. 5. pm 2:22:39 PM KST	2.0B	Standard
<input type="checkbox"/>	002_DAY001minchocoin1918719978.txt	txt	2024. 6. 6. am 1:12:41 AM KST	478.0B	Standard

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?

pnuspeechnote@gmail.com
나에게 ▼

Please find the attached file.

첨부파일 1개 • Gmail에서 스캔함

002_DAY001minc...

← 답장 → 전달 😊

002_DAY001minc ... 1918719978.txt 📄 연결 프로그램: Google 문서 ▼

편집
받은
별표
다시


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

- 강의/강연 녹음파일 변환
- <https://youtu.be/SEC13EslJKc>

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 모델 입니다. 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 머신 러닝 모델입니다. ...



 lecturenote2.mp3

2024-06-06 오후 12:51



pnuspeechnote@gmail.com

나에게 ▾

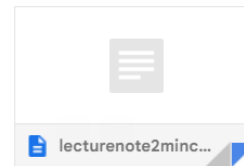


한국어로 번역



Please find the attached file.

첨부파일 1개 • Gmail에서 스캔함 ⓘ



minchocoin@gmail.com

Korean

파일 선택

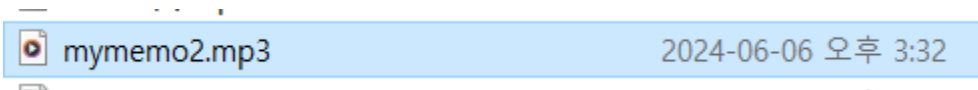
lecturenote2.mp3

Submit

활용방안

- 음성메모를 텍스트로 변환
- https://youtu.be/-CuXJ7cM_Ck

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.



Your file has been uploaded [받은편지함](#) ×



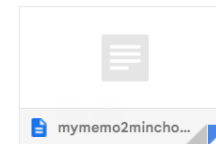
pnuspeechnote@gmail.com

나에게 ▾

한국어로 번역 ×

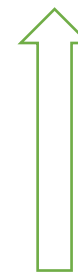
Please find the attached file.

첨부파일 1개 • Gmail에서 스캔함 ⓘ



← 답장

→ 전달



minchocoin@gmail.com

Korean

파일 선택 mymemo2.mp3

Submit

개선방안

- 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
- 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>
- <https://duckgugong.tistory.com/336> : cloudfront 사용법
- <https://heytech.tistory.com/403> : 파일업로드 웹페이지 구축
- <https://medium.com/@manishdiddi03/automating-audio-transcription-how-to-use-aws-api-gateway-s3-눈-lambda-and-transcribe-to-20c220b1e77f> : apiGateway-s3-lambda-transcribe 구축
- https://docs.aws.amazon.com/transcribe/latest/APIReference/API_StartTranscriptionJob.html : transcribe API guide
- <https://medium.com/analytics-vidhya/transcribing-audio-files-with-amazon-transcribe-lambda-s3-474dc9a1ced7> : cloudwatch 구축
- <https://velog.io/@chaduri7913/AWS-RDS-Lambda-%EC%97%B0%EA%B2%B0%ED%95%98%EA%B8%B0> : lambda와 DB 연결
- <https://icedhotchoco.tistory.com/entry/sending-email-with-python> : google SMTP 사용

최종 결과물 깃허브 링크

- <https://github.com/minchoCoin/stt-service>



stt-service Public

main 1 Branch 0 Tags

Go to file Add file Code

minchoCoin Update README.md 0cc4d93 · 1 hour ago 69 Commits

.idea	feat: S3 Transcribe 결과를 이메일로 전송하는 Lambda 함수 ...	19 hours ago
diagram	change diagram	4 days ago
sql	Create createTable.sql	17 hours ago
src	remove space from filename	16 hours ago
LICENSE	Update LICENSE	2 months ago
README.md	Update README.md	1 hour ago
시스템구축.pdf	Create 시스템구축.pdf	15 hours ago

README Apache-2.0 license

stt-service / src

minchoCoin remove space from filename

Name

- ..
- audio-transcribe
- file-upload-backend
- frontend
- save-text
- send-email



<https://github.com/minchoCoin/stt-service/issues>