

A 3D scene featuring several translucent cubes and cylinders in shades of blue, orange, and brown, arranged in a dynamic, floating composition. A bright yellow sphere is positioned near the center-right. A thin white horizontal line spans the width of the slide, passing behind the text.

Hi Azure!

Helping video creators who are visually-impaired edit videos using Azure Cloud AI!

01 아이디어 구상 및 제안 배경

02 제안하는 AI 모델 및 서비스

03 추후 고도화 및 실현 가능성

A black and white photograph of a man sitting on a light-colored floor, leaning over a laptop. He is holding a smartphone in his right hand, pointing at the laptop screen. A video game controller is on the floor next to the laptop. The background is a simple room with a couch and a lamp.

Part 1 아이디어 구상 및 제안 배경

(1) 제안배경 1: 통계조사

- 시각장애가 없는 사람들은 관심분야에 대해 크리에이터 활동을 자유롭게 꿈꿀 수 있고 활동에 제약이 없음

성인남녀 10명 중 6명, 유튜버 꿈꾼다!

성인남녀 3,543명 설문조사 [자료제공: 사람인]

유튜버 도전 의향 **있다 63%**

없다 37%

* 가장 하고 싶은 유튜브 분야

일상 **31.1%**

게임 13.9%

먹방 10.4%

요리 6.3%

음악 5.9%



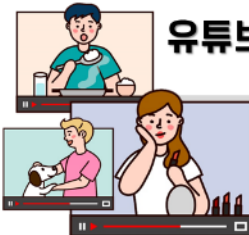
사람인

직장인 4명중 1명 이상 '유튜브 운영중'... 평균 150만원 투자했다

조사대상: 성인남녀 735명

설문기간: '20년 9월 25 ~ 29일(닷새간)

Q. 개인 유튜브 채널 운영하고 있습니까?



유튜브 운영한다

전체 응답자 **32.3%**

직장인 **29.3%**

운영하는
이유는?

- 취미생활 및 일상 기록 (47.0%)
- 수익창출, 부업 (36.7%)
- 포트폴리오 구축 차원 (10.2%)

Q. 채널운영에 얼마나 투자하셨습니까?

콘텐츠 준비비용 평균 74만원

방송장비 구매비용 평균 56만원

인력비용 평균 31만원



총 투자비용 평균 **150만원**

자료제공: Incruit

초등학생 희망직업 상위 10위

단위: %

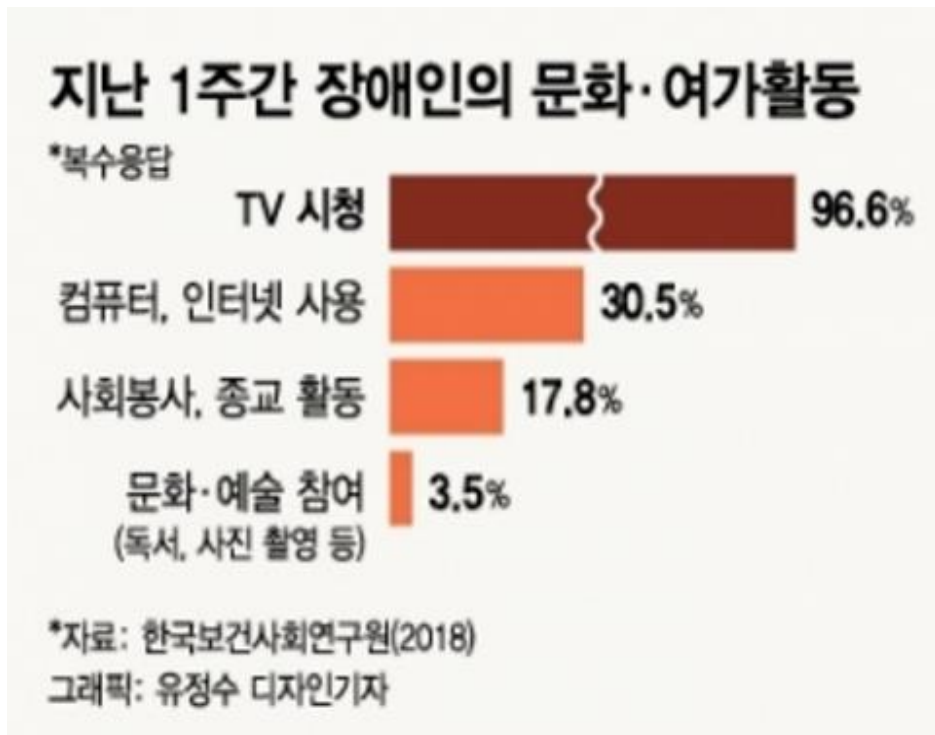
순위	2019	2020	2021	
1	운동선수	운동선수	운동선수	8.5
2	교사	의사	의사	6.7
3	크리에이터	교사	교사	6.7
4	의사	크리에이터	크리에이터	6.1
5	조리사(요리사)	프로게이머	경찰관/수사관	4.2
6	프로게이머	경찰관	조리사(요리사)	4.1
7	경찰관	조리사(요리사)	프로게이머	3.7
8	법률전문가	가수	배우/모델	3.3
9	가수	만화가(웹툰작가)	가수/성악가	2.9
10	뷰티디자이너	제과·제빵사	법률전문가	2.7

자료: 한국직업능력개발원

The JoongAng

(1) 제안배경 1: 통계조사

- 하지만, 시각장애인들의 문화 및 여가활동의 절대 다수는 TV시청등의 정적인 활동임
- 이는 시각장애인들이 가진 끼와 재능을 더 다채롭게 펼치기에 많은 현실적인 제약이 있다고 볼 수 있음



(2) 제안배경 2: 실제 시각장애 크리에이터들의 인터뷰



“편집은 혼자하기 어려워 친구가 도와줍니다. 누군가 편집을 어떻게 하냐고 물어보면 ‘좋은 친구를 사귀면 된다.’라고 말하죠. 그 친구에게 고마운 점이 많아요.”

영상 크리에이터 ‘원샷한솔’ 한국시작장애인복지관 인터뷰 中



“아무래도 시각장애가 있어서 혼자 컴퓨터나 이런 걸로 편집프로그램을 사용하기는 어려움이 있어요. 그래서 저는 도와주는 PD 언니한테 스크립트를 전달해서 영상을 만들고 있습니다.”

영상 크리에이터 ‘우령의 유디오’ TBS 인터뷰 中



편집은 어떻게 찍혔는지 기억해가며 진행한다.

이렇다 보니 6 ~ 7분가량의 영상을 만드는 데 3일이 넘는 시간이 걸릴 때도 있다.

영상 크리에이터 ‘브레드박’ 한국일보 기사 中

(2) 제안배경 2: 실제 시각장애 크리에이터들의 인터뷰



“편집은 혼자하기 어려워 친구가 도와줍니다. 누군가 편집을 어떻게 하냐고 물어보면
‘좋은 친구를 사귀면 된다.’라고 말하죠. 그 친구에게 고마운 점이 많아요.”

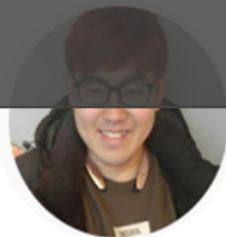
영상 크리에이터 ‘원샷한술’ 한국시각장애인복지관 인터뷰 中



“시각장애 크리에이터들은 콘텐츠를 제작할 때
주변의 도움 없이는 청각정보에만 의존할 수 밖에 없음”

“시각장애 크리에이터들은 콘텐츠를 제작할 때
주변의 도움 없이는 청각정보에만 의존할 수 밖에 없음”

영상 크리에이터 ‘우령의 유디오’ TBS 인터뷰 中



편집은 어떻게 찍혔는지 기억해가며 진행한다.

이렇다 보니 6 ~ 7분가량의 영상을 만드는 데 3일이 넘는 시간이 걸릴 때도 있다.

영상 크리에이터 ‘브레드박’ 한국일보 기사 中

(3) Hi Azure!

- 조용한 공원에서 풍경을 촬영하거나 영상 배경이 다양하게 바뀌는 vlog 제작 등, 청각 정보에만 의존하기 어려운 영상들은 혼자 편집하기에 매우 어려울 수 있음
- 이를 해결하기 위해, **Hi Azure!** 는영상의 시각 정보를 묘사한 음성을 출력하여 콘텐츠 제작에 도움을 줌



사용자가 편집 중인 영상

>>



>>

‘an office with
several computers’

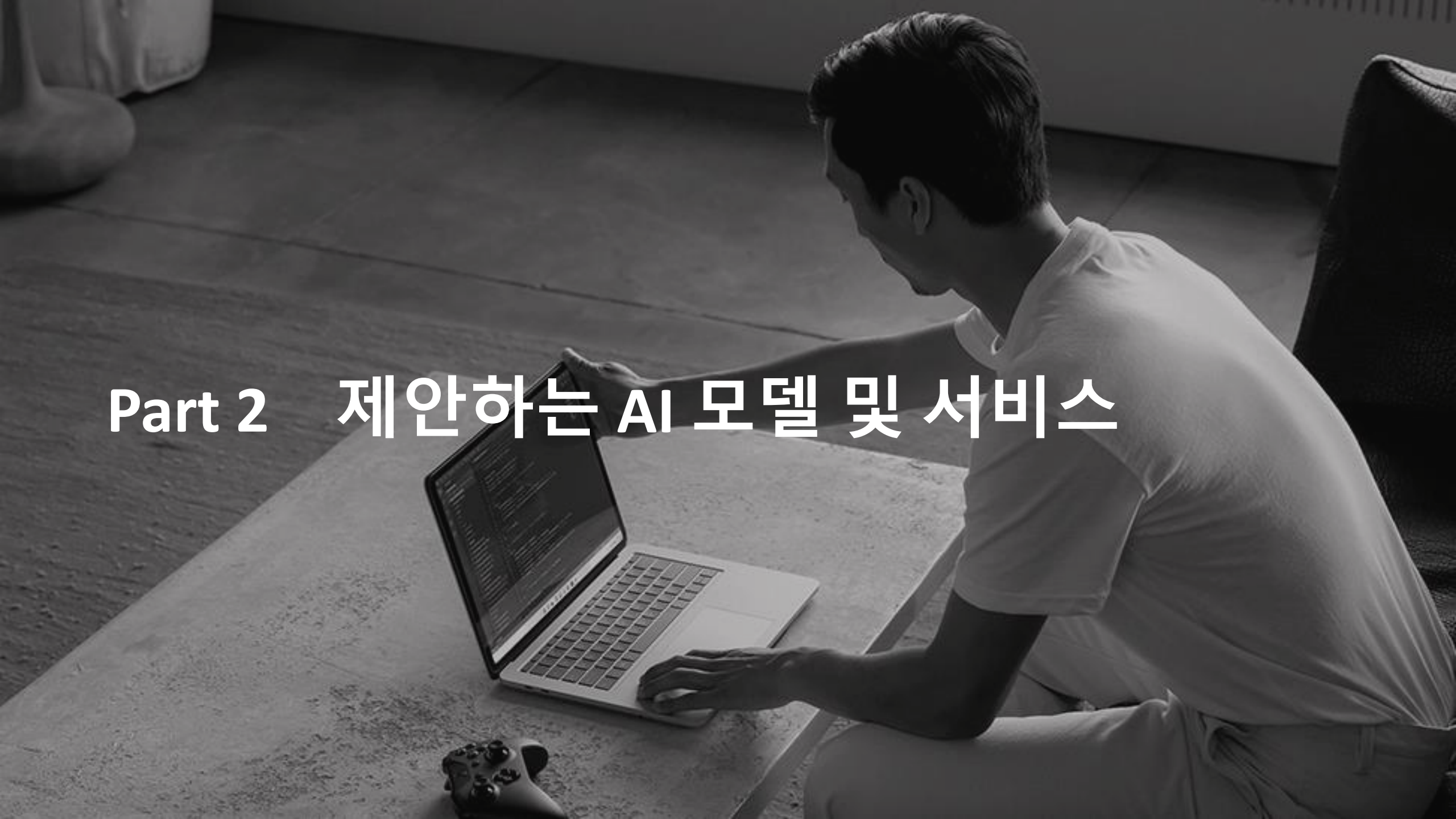
Visual feature를 분석하여 caption 생성



“an office with
several computers”

음성 출력

Part 2 제안하는 AI 모델 및 서비스



(1) 제안 목적



→ AI for Accessibility: For video creators who are visually impaired

- ✓ 영상 제작의 진입장벽이 높음
- ✓ 청각정보에만 의존하여 콘텐츠를 제작할 수 밖에 없음

영상 편집 및 콘텐츠 제작에 대한
접근성을 보다 높임

영상의 청각정보 + 시각정보

Hi Azure!

Part 2

제안하는 AI 모델 및 서비스

(2) 제안하는 AI 모델

Video Frames
(e.g. Daily Vlog)



Input
➔

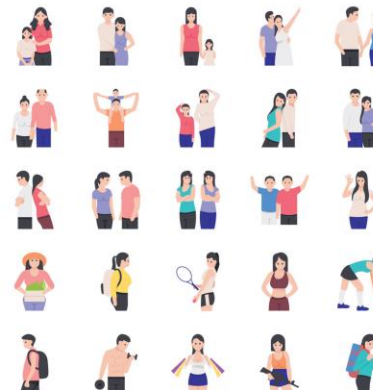
Hi Azure!
(Our Service model)



Video Description
Using Visual-Audio Feature



Video Creators
Who Are Visually Impaired



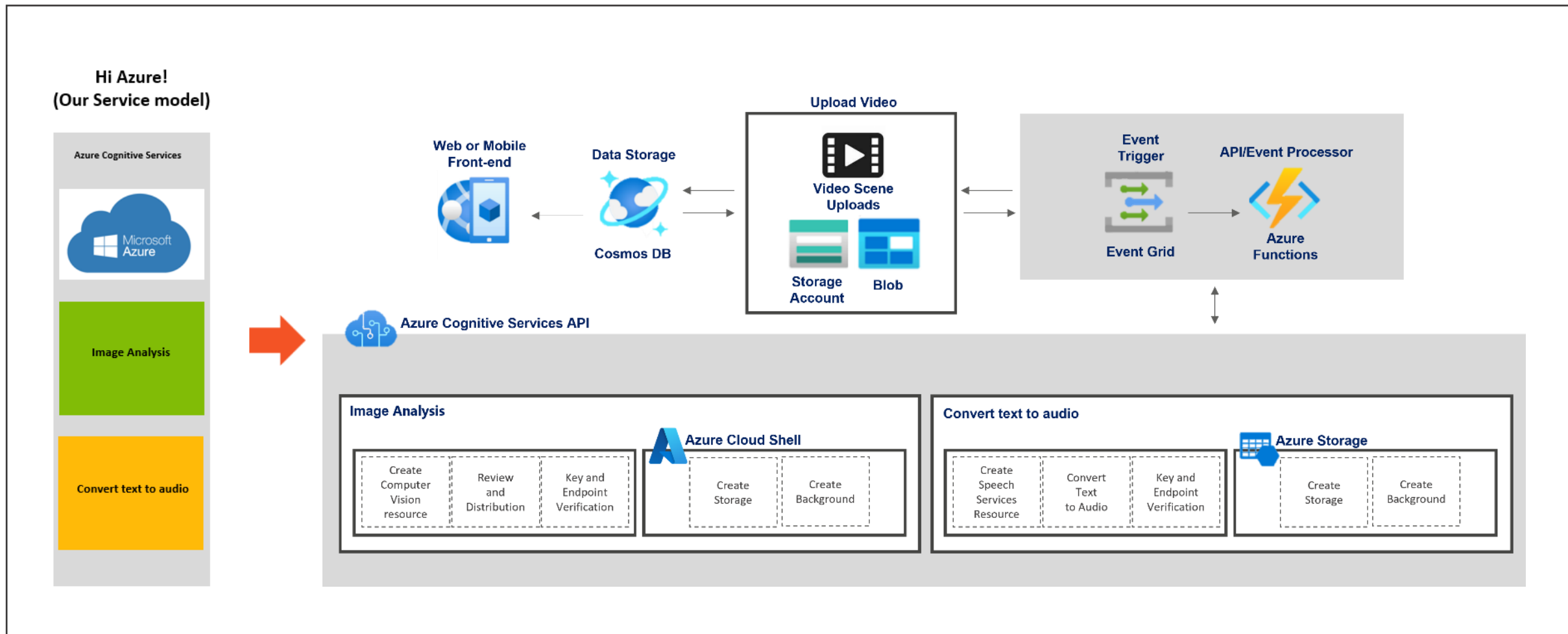
Video Frames
with user caption



User Caption:
“한번에 누를 수 있게 됐어요”

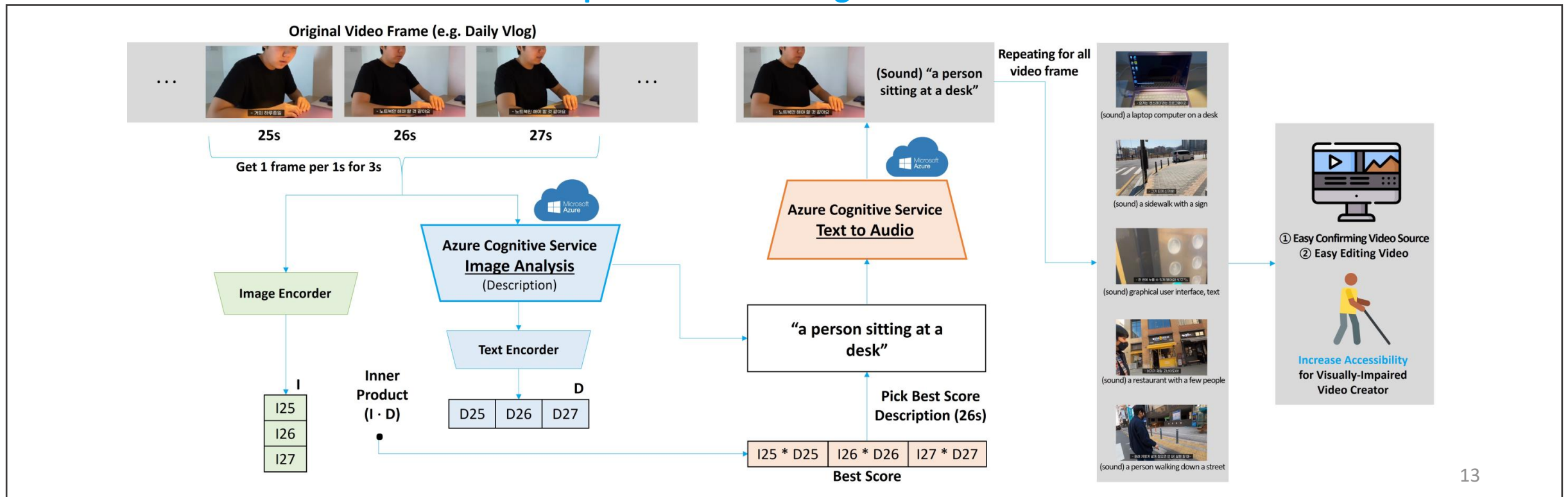
(2) 제안하는 AI 모델

- ① Azure Cognitive Services API ② Image Analysis ③ Convert Text to Audio 를 활용하여 서비스를 제작



(2) 제안하는 AI 모델

- 영상을 프레임 단위로 나누고 나뉜 프레임 단위마다 Azure를 활용한 Image Analysis module로 영상을 분석
- 분석 결과 Azure를 사용한 Convert Text to Audio module로 음성으로 변환함
- **영상의 caption을 생성할때, 출력 결과를 그대로 사용하지 않고, context window를 설정하여 context 내에서 현재 video frame과 가장 연관이 있는 caption을 출력하여 generation error를 최대한 방지함**



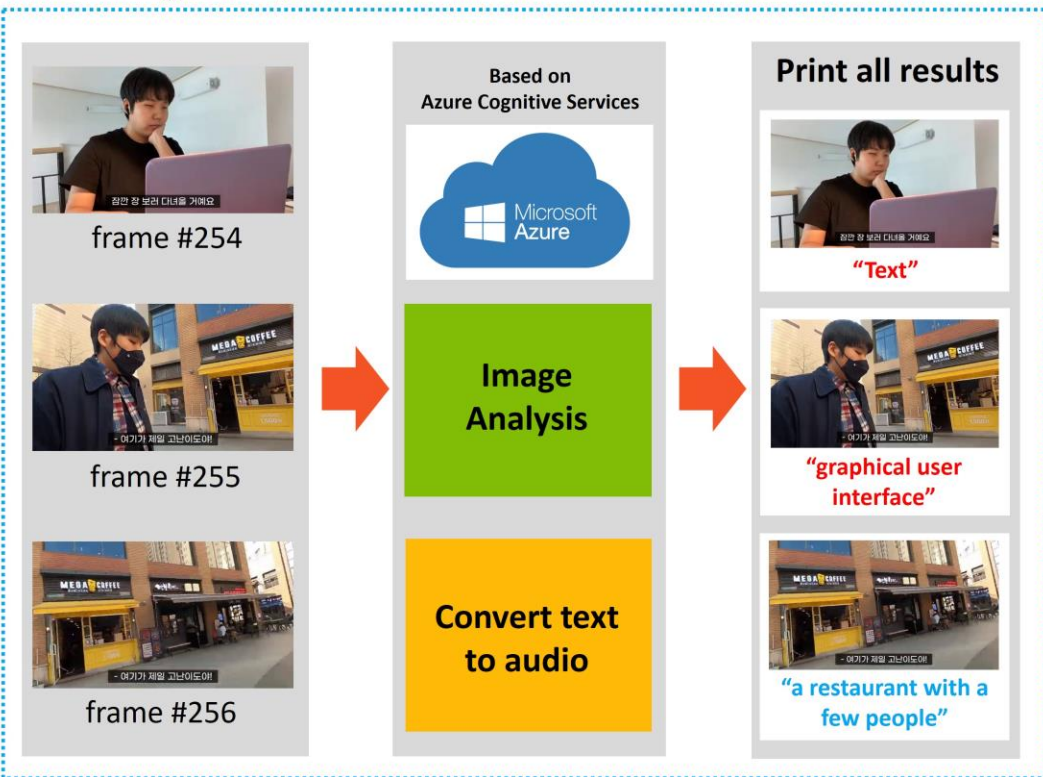
Part 2

제안하는 AI 모델 및 서비스

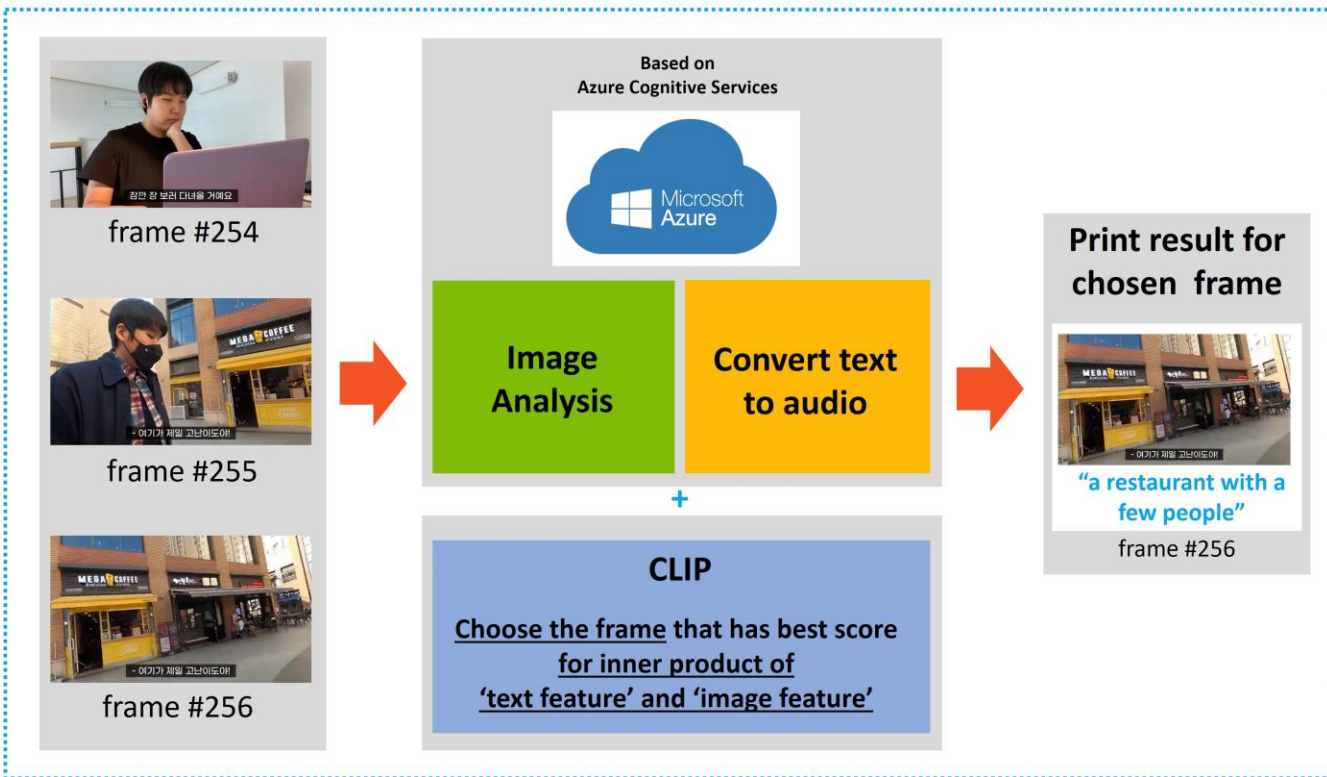
(3) 제안하는 서비스

- ① Azure Cognitive Services API ② Image Analysis ③ Convert Text to Audio ④ Dense Captioning을 활용

Azure Image Analysis, TTS

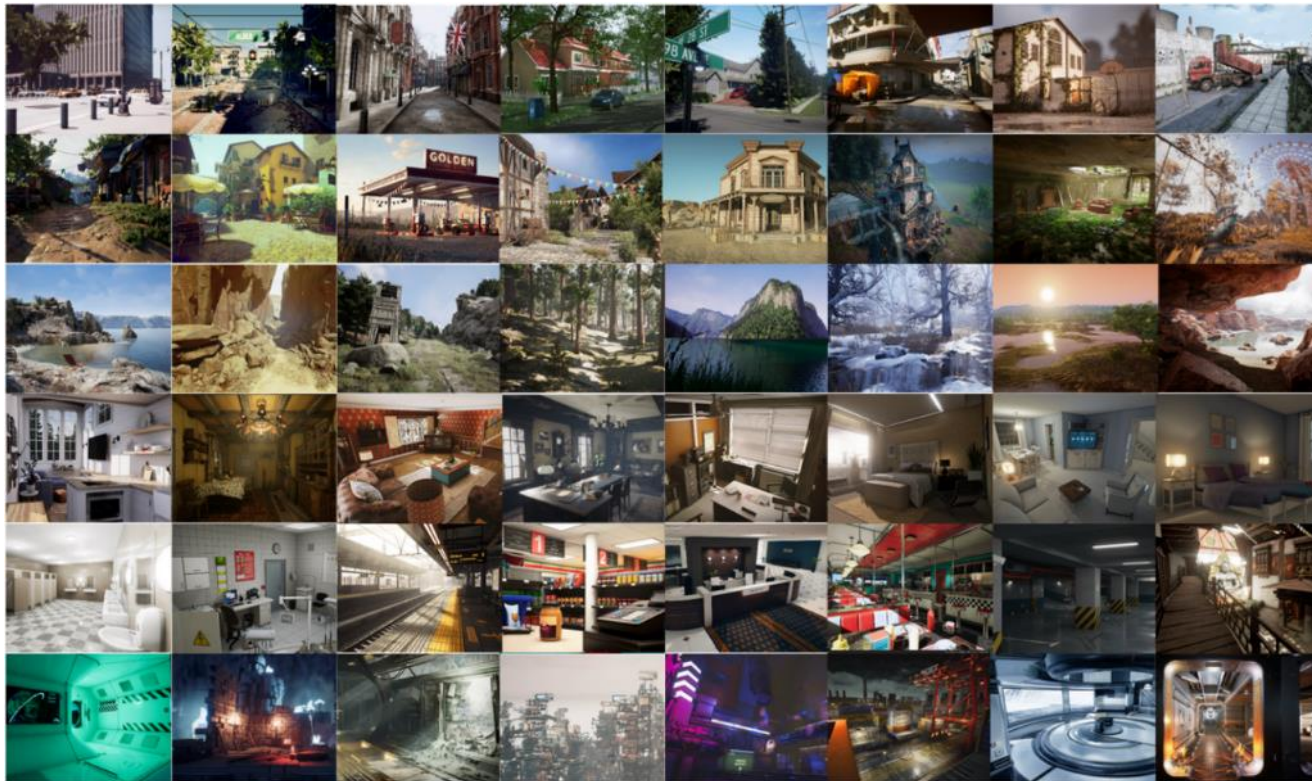


Azure Image Analysis, TTS + CLIP

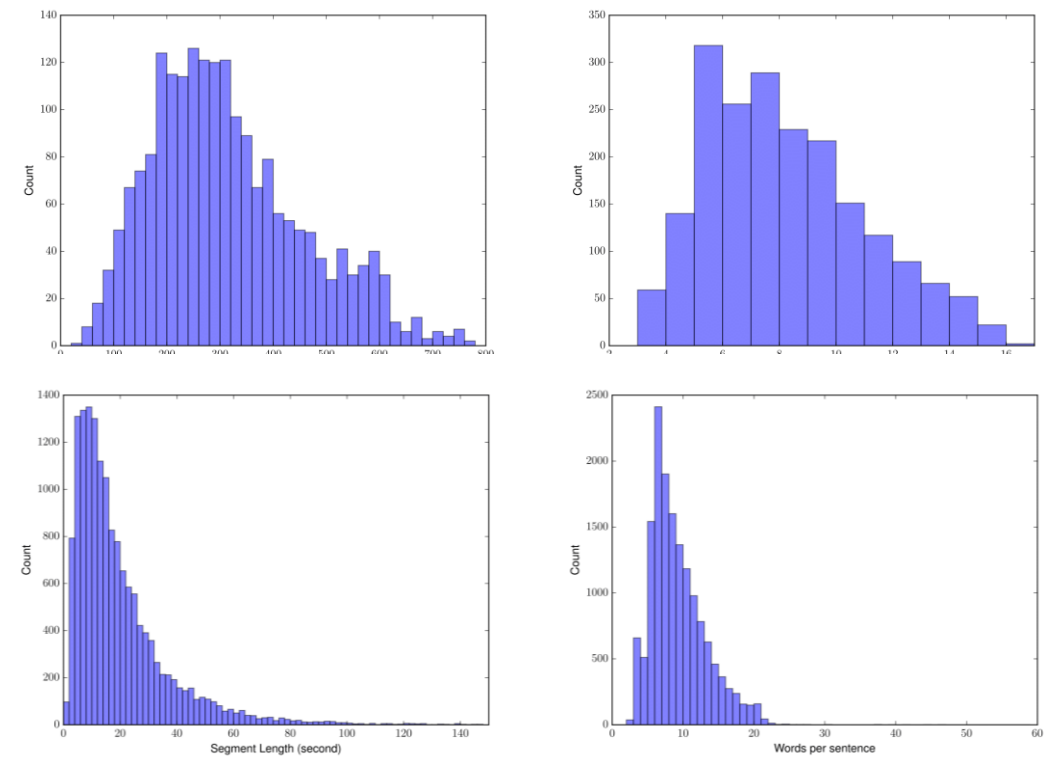


(4) 사용하는 데이터셋에 대한 설명 및 활용성

- TartanAir dataset (Azure open dataset) : 연속 프레임 데이터셋, 모델 성능 분석 및 검증에 활용
- Youcook2 dataset : 비디오-캡션 데이터셋, 모델 성능 분석 및 검증에 활용



TartanAir dataset (Azure open dataset)



Youcook2 dataset

(5) Microsoft Azure 활용성

- Azure Machine Learning Studio 가상 공간에 코드와 이미지가 있는 git repository를 clone 해서 사용함.
- 깃 레포지토리 안에 Azure Cognitive Services에서 제공하는 코드를 활용함.
- Key 값, Endpoint, Running code 세 가지를 불러와서 사용함.



Input images

키 표시

키 1

키 2

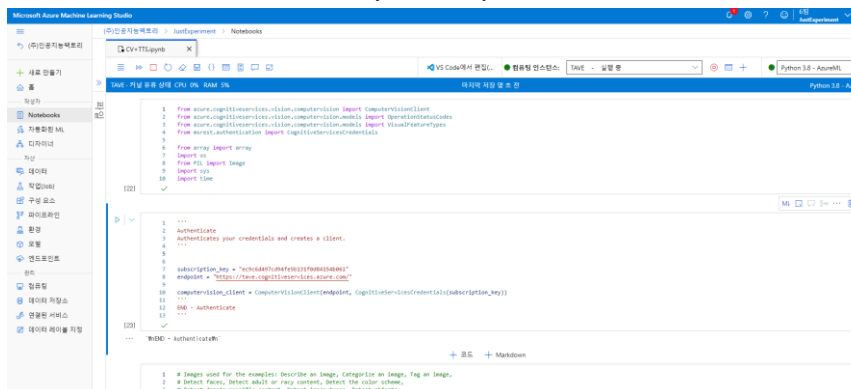
위치/지역 ①

koreacentral

엔드포인트

https://tave.cognitiveservices.azure.com/

Key 값, Endpoint



Azure Cognitive Services에서 제공한 코드를 활용한 실행창

MicrosoftLearning / AI-900-AIFundamentals (Public)
generated from MicrosoftLearning/INF99X-SampleCourse

Includes labs for AI Fundamentals.

microsoftlearning.github.io/ai-900-ai-fundamentals/

MIT license

☆ 29 stars 🍴 45 forks

☆ Star

🔔 Notifications

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#)

main ▾

Go to file

GraemeMalcolm ...

✓ on 15 Apr 🕒

[View code](#)

Microsoft Azure AI Fundamentals Exercises

This repository contains the hands-on lab exercises for Microsoft course [AI-900 Microsoft Azure AI Fundamentals](#) and the [self-paced modules on Microsoft Learn](#). The labs are designed to accompany the learning materials and enable you to practice using the technologies they describe.

- [Download Latest Student Handbook and AllFiles Content](#)
- [Are you a MCT?](#) - Have a look at our [GitHub User Guide for MCTs](#)
- [Need to manually build the lab instructions?](#) - Instructions are available in the

[readme.md](#)

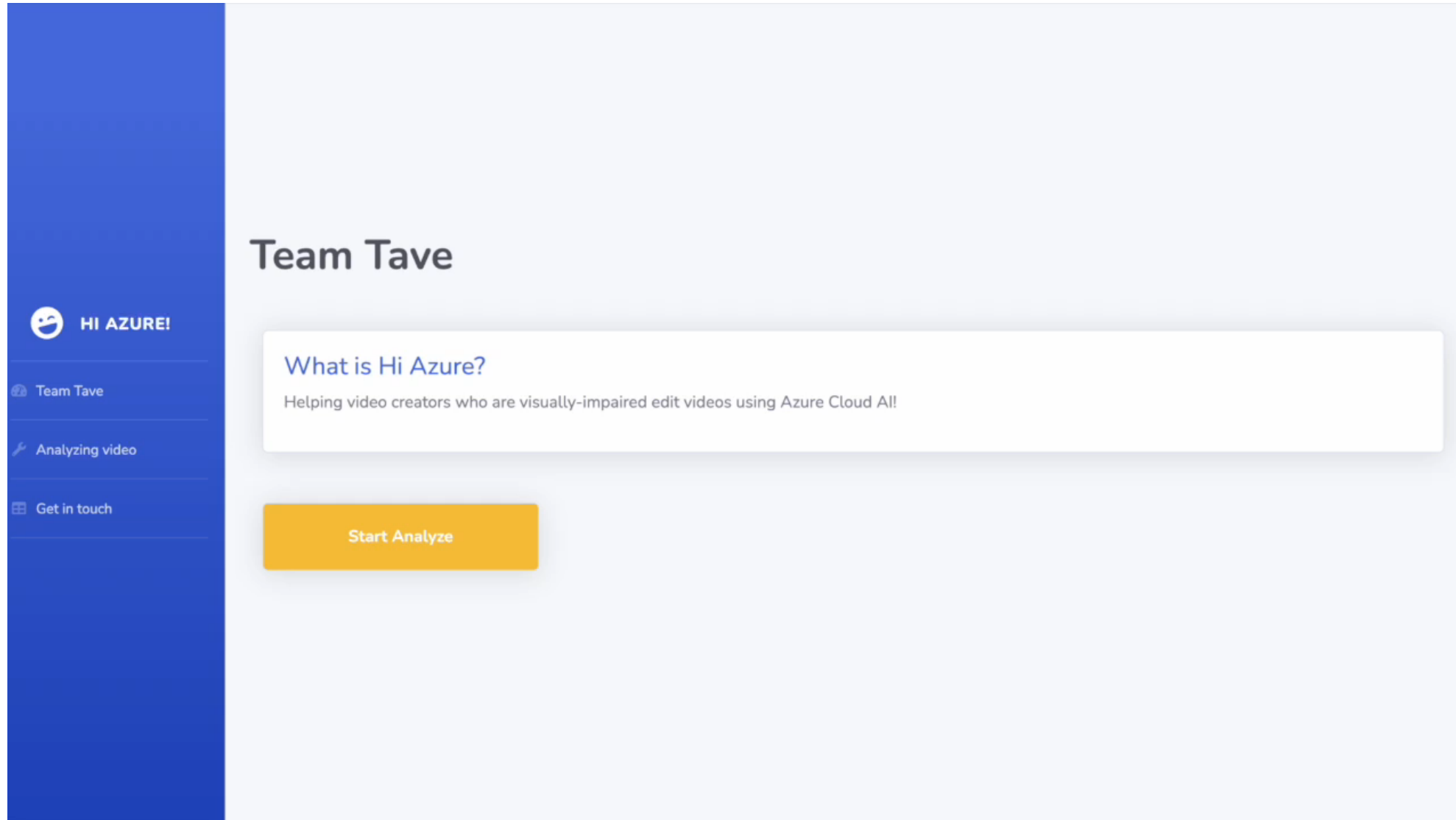
What are we doing?

Azure Cognitive Services에서 제공한 git repository

Part 2

제안하는 AI 모델 및 서비스

(6) Demo

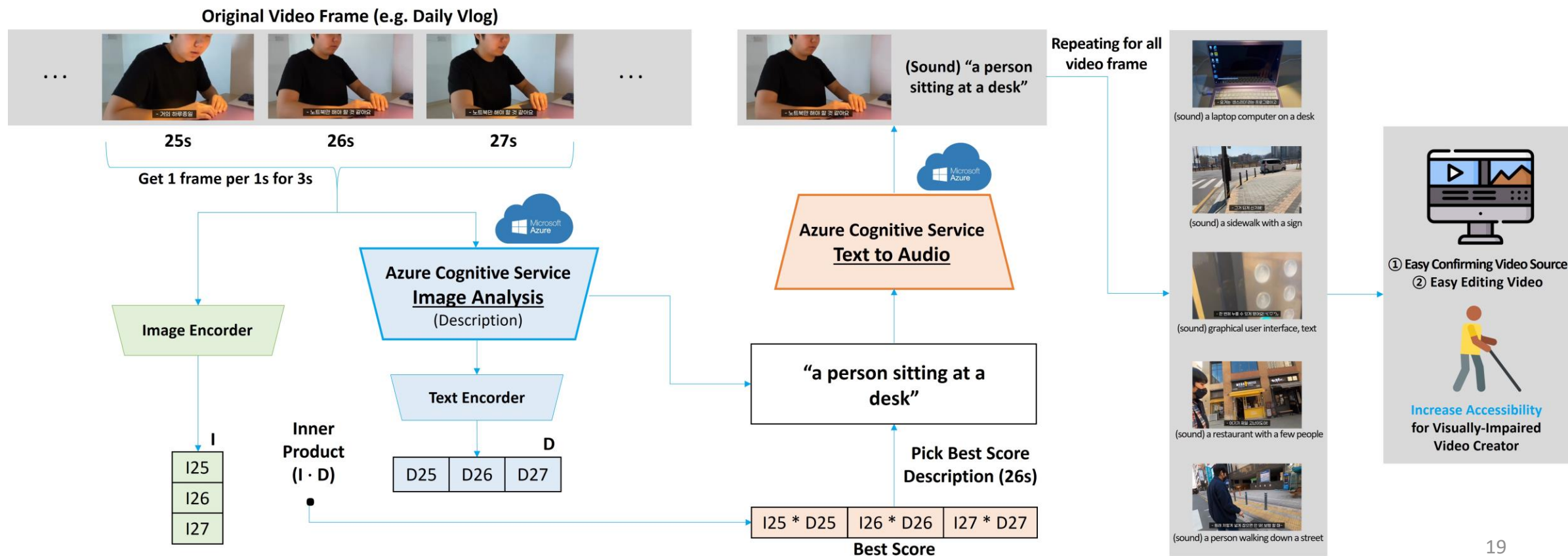


Part 3 추후 고도화 및 실현 가능성



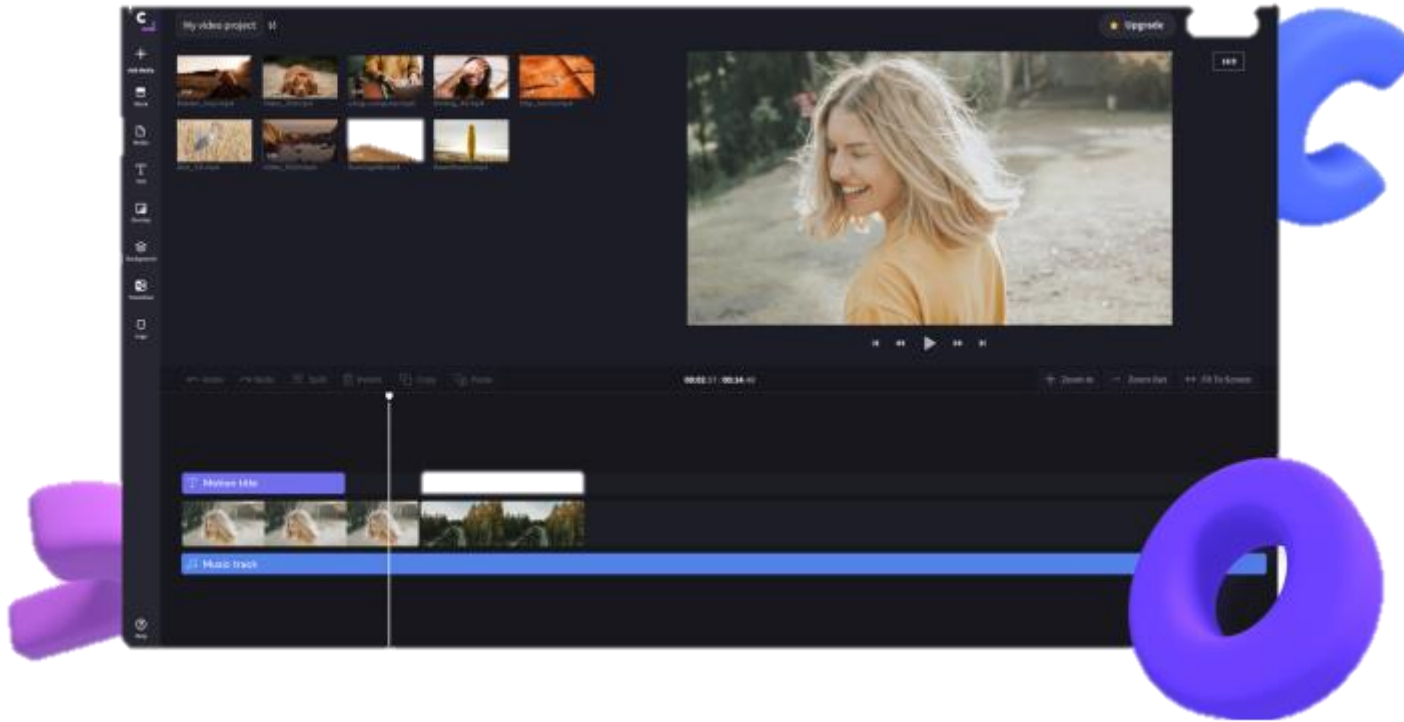
(1) 추후 고도화 및 실현 가능성: 기술적인 관점

- 영상 내의 다양한 modality (e.g. RGB features, Optical Flow, audio features)를 활용하여 인물의 세밀한 표정변화등도 묘사할 수 있도록 기술의 고도화 진행



(2) 추후 고도화 및 실현 가능성: 실제 적용의 관점

- 'Microsoft Clipchamp' 등 영상 편집 프로그램 적용으로의 실현 가능성



현재 널리 쓰이고 있는 전문가용 영상 편집 프로그램
'Clipchamp' 등에 보편적 접근성 강화를 위한 기능으로서 사용 가능

(1) 파급효과 1: 시각 장애인의 사회 참여 및 삶의 질 향상

적용 전

1. 취미로서 오디오 청취, 점자책
독서, 사색, 명상 등 혼자만의
단방향 취미
2. 외부활동이 극히 적어,
인간 관계에 부족함을 느낌
3. 주변의 시선에 괜히 움츠러들음



Hi Azure!

적용 후

1. 본인의 이야기, 전문지식을
담은 영상을 외부와 공유,
소통하여 독립, 사회 참여율 상승
2. 영상 작업을 통해 자연스럽게
외부로 신체적 확장 가능 및
경험할 수 있는 공간 증가

(2) 파급효과 2: 일반인의 장애인에 대한 인식 변화 유도

- 타 매체 대비 사용율이 높은 영상매체에서의 장애인의 삶에 대한 자연스러운 노출 상승
- 장애인에 대한 배려의식 및 함께 사는 삶에 대한 인식 변화 유도
- 갑작스럽게 생긴 시각장애를 겪는 분들에게 사회 적응, 일반 생활 영위에 도움을 줄 수 있도록 함

시사 > 전체기사

장애예술인 활동 동영상 “670만회 기록 대박”

| 10일 장애예술인지원법을 시행 내년 기대감 고조, 해외에서도 관심 높아

입력 : 2020-12-05 20:27



국민일보 정창교 기자



유튜버 '우리정인이STUDIODEAR'



유튜버 '원샷한솔'

시각장애 유튜버의 관점

Thank you



임예람



이문기







허주희



박제윤

프로젝트를 위한 노력

- Cloud Skills Challenge: 52명이 참가하여 9명 완주. TAVE팀 전원 완주 (9/52 = 17.3%, Track 2)

참가자 52명		모듈 완료
 임예람		21/21
 이문기		21/21
 Juhui Heo		21/21
 박제윤[대학원석사과정졸업 / 컴퓨터학과]		21/21