


캡스톤 디자인 I 종합설계 프로젝트

프로젝트 명	수화 人(in) 플레이어
팀 명	전지적 수화 시점 팀
문서 제목	수행계획서

Version	2.3
Date	2019-APRIL-18

팀원	윤지은 (조장)
	정창희
	이희지
	배진영
	축이림

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


CONFIDENTIALITY/SECURITY WARNING

이 문서에 포함되어 있는 정보는 국민대학교 소프트웨어융합대학 소프트웨어학부 및 소프트웨어학부 개설 교과목 캡스톤 디자인 수강 학생 중 프로젝트 “수화 人(in) 플레이어”를 수행하는 팀 “전지적 수화 시점”의 팀원들의 자산입니다. 국민대학교 소프트웨어학부 및 팀 “전지적 수화 시점”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

문서 정보 / 수정 내역


Filename	수행 계획서-수화 인 플레이어
원안작성자	윤지은, 정창희, 이희지, 배진영, 측이림
수정작업자	윤지은, 정창희, 이희지, 배진영, 측이림

수정날짜	대표수정 자	Revision	추가/수정 항목	내 용
2019-03-06	이희지	1.0	최초 작성	수행 계획서 초안 작성
2019-03-08	배진영	1.1	내용 작성	개발 목표 및 내용 작성
2019-03-08	정창희	1.2	내용 작성	프로젝트 개요 및 추진 배경 작성
2019-03-08	윤지은	1.3	내용 작성	현실적 제한 요소 및 그 해결 방안
2019-03-14	전원	1.4	내용 수정	전반적 내용 수정
2019-03-15	전원	1.5	최종 수정	수행 계획서 ver1.5 최종본 작성
2019-04-03	전원	2.0	계획서 갱신	시스템 기능/비기능 요구사항 추가
2019-04-10	배진영	2.1	내용 수정	변경된 기술 내용 수정
2019-04-15	이희지	2.2	내용 수정	계획서 수정
2019-04-18	윤지은	2.3	최종 계획서 검토	제출 전 최종 수정

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

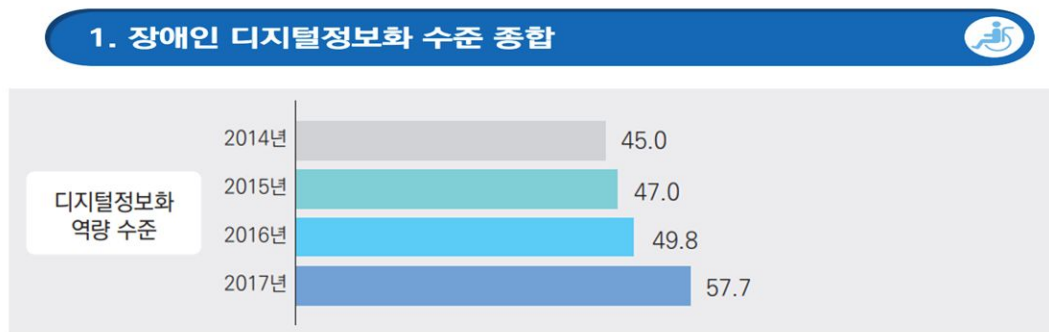
목 차

1	개요	4
1.1	프로젝트 개요	4
1.2	추진 배경 및 필요성	5
2	개발 목표 및 내용	5
2.1	목표	6
2.2	연구/개발 내용	7
2.2.1	단어 및 수화 DataSet 구축	7
2.2.1.1	Crawling	7
2.2.1.2	MongoDB	7
2.2.2	동영상의 음성에서 자막파일 추출 - STT(Speech to Text)	8
2.2.3	텍스트 수화 번역 - 형태소 분석	8
2.2.4	수화 Data Mapping - 단어간 유사도 측정	10
2.3	개발 결과	11
2.3.1	예상 시나리오	11
2.3.2	시스템 기능 요구사항	12
2.3.3	시스템 비기능(품질) 요구사항	12
2.3.4	Use Case Diagram	13
2.3.5	시스템 구조	14
2.3.6	결과물 목록 및 상세 사양	14
2.4	기대효과 및 활용방안	15
3	배경 기술	15
3.1	기술적 요구사항	15
3.1.1	개발 환경	15
3.1.2	Frontend	15
3.1.3	Backend	16
3.1.4	자막/수화 생성	16
3.2	현실적 제한 요소 및 그 해결 방안	18
3.2.1	하드웨어	18
3.2.2	소프트웨어	18
3.2.3	데이터	18
3.2.3	기타	18
4	프로젝트 팀 구성 및 역할 분담	19
5	프로젝트 비용	19
6	개발 일정 및 자원 관리	20
6.1	개발 일정	20
6.2	일정별 주요 산출물	20
6.3	인력자원 투입계획	21
6.4	비 인적자원 투입계획	22
7	참고 문헌	22

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

1. 개요

1.1. 프로젝트 개요




'2017 디지털정보격차 실태조사'결과에 따르면

장애인 정보화 역량은 비장애인의 **절반 수준**

<그림1. 장애인의 디지털 정보화 수준>

한국정보화진흥원이 조사한 '2017 디지털정보격차 실태조사'결과에 따르면 장애인 정보화 역량은 비장애인의 절반 수준으로 나타났다. 특히 청각장애인은 어떤 장애 유형보다도 정보 습득이 어렵다. 시각장애인을 위해서는 전자점자 자료 및 점자번역기 등 솔루션이 있지만 청각장애인을 위한 수화로 번역된 영상의 제작과 배포 체계는 여전히 미흡한 상황이다. 따라서 우리는 청각장애인이 동영상을 원활하게 시청할 수 있도록 수화를 함께 제공하는 동영상 플레이어(이하 수화 人 플레이어) 제작을 목표로 한다.

수화 人 플레이어는 기존의 동영상에 ①단일 자막, ②단일 수화 또는 ③두 가지 모두를 선택적으로 제공한다. 결과적으로 청각장애인이 동영상 콘텐츠를 제약없이 이용함으로써 원하는 정보를 습득할 수 있도록 한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

1.2. 추진 배경 및 필요성

[단위: 천명]

	2009	2010	2011	2012	2013	2014	2015	2016	2017
청각, 언어	262	277	279	276	273	271	269	291	322

<표 1. 연도별 장애인 현황> (출처 : 보건복지부 (시.도 장애인등록현황 자료))

최근에 사회적 약자에 대한 관심의 필요성이 대두되고 있다. 위 제시된 [표1]에 나온 바와 같이 2017년 기준 청각장애인의 수는 32만 명으로 지체장애 다음으로 많은 장애유형이며 그 수가 조금씩 증가하고 있지만, 사회적 여건은 열악하다. 이는 정보를 습득하는 과정에서도 문제점을 야기한다.

IT 기술 발달로 정보화 시대가 도래함에 따라 정보의 습득은 현대 사회에서 가장 중요한 부분으로 자리 잡았다. 정보는 TV, 신문, 영화, 라디오, 휴대폰 등 다양한 대중매체를 이용해 접근할 수 있으나 대부분의 경우 음성으로 제공한다. 이는 결과적으로 청인들과 정보격차를 발생시켜 청각장애인들을 정보 소외 계층으로 전락하게 만든다.


대규모의 세계적 영상 스트리밍 플랫폼인 Netflix와 Youtube는 자막을 제공하는데 하지만 청각장애인들은 자막만으로는 영상의 내용을 정확하게 이해하기 힘들다. 이를 뒷받침 해 줄 기사의 일부 내용을 발췌하였다.

장애인정보문화누리 함효숙 활동가는 “방송을 볼 때 수화통역을 주로 보지만 중요한 내용은 자막과 같이 비교하며 보는데, 그래야 내용을 더 정확하게 알 수 있기 때문” 이라면서 “나에게 자막과 수화통역은 자유롭게 선택되어야 하는 언어” 라고 강조했다.

이날 기자회견에서 장애인정보문화누리 안세준 고문은 “나는 어려서 청각장애를 입어 수화가 모국어나 마찬가지”라면서 “자막도 방송을 보는 데 도움이 되기는 하지만 전달이 잘 안 될 때가 잦다. 그래서 방송을 볼 때 수화통역이 있는 것을 더 좋아한다”라고 밝혔다.

(“선거방송, 수화와 자막 동시에 제공하라”, 비마이너 ,
2012년 11월 23일, (<http://beminor.com/detail.php?number=4586>))

기사 내용으로 보다시피 청각장애인이 내용을 정확하게 이해하기 위해서는 자막만이 아닌 수화를 동시적으로 보여주는 것이 필요하다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

미국은 사회적 약자를 위한 복지 서비스를 제공하고 있어 청각장애인을 위한 수화 data API나 수화 번역 툴들이 개발된 상태이지만 우리나라는 아직 사회적 약자를 위한 서비스 배포가 미흡하고 특히 청각장애인을 위한 서비스는 부족하다. 또한 이를 활용할 수화 데이터 또한 부족한 상태이다. 현재 우리나라에서 상용되고 있는 서비스로 삼성의 ‘수화번역 도우미’가 있으나 이는 표정을 제공하지 않아 정확한 의미 전달을 하지 못하며, 문장을 입력하면 수화 애니메이션을 보여주는 방식으로 청각장애인을 위한 서비스는 아니다.

우리는 이러한 환경에서 청각장애인을 위한 수화 통역 영상플레이어 제작의 필요성을 느껴 본 프로젝트를 진행하고자 한다.

2. 개발 목표 및 내용

2.1. 목표

본 프로젝트의 이름은 ‘수화인플레이어’로 사용자가 정보를 얻고자 하는 영상을 등록하면 영상의 내용을 수화로 번역하여 이를 등록된 영상과 함께 사용자에게 제공하는 것이 목표이다.


현재 시중에 나와있는 대부분의 영상매체는 자막이 함께 제공되는 영상이 적을 뿐더러 더욱이 수화 영상을 함께 제공하는 것은 뉴스나 일부 수화통역 영상 이외에 거의 찾아볼 수 없다. IT 기술 발달로 점점 영상 매체의 의존도가 높아짐에 따라 청인과 청각장애인 간의 정보 격차는 심화될 것으로 예상된다. 하여 본 프로젝트의 기대효과는 청각장애인도 인터넷 강의나 시사/교양 프로그램 등 정보를 담고 있는 영상에서 원하는 정보를 제약없이 습득할 수 있도록 하는 것이다.

목표 달성을 위한 기술적 단계는 크게 4가지로 나뉜다.

- 1) 단어 및 수화 DataSet 구축
- 2) 동영상의 음성에서 자막 파일 생성
- 3) 한국어 문법을 수화 문법 체계 형식으로 변형
- 4) 추출한 형태소와 수화 Data Mapping방식 구현

본 프로젝트의 주된 목적은 영상을 자막 및 수화로 번역하는 시스템 자체의 설계 및 구현이며, 모든 문장을 정확히 해석하는 것보다 전체적인 영상의 내용을 파악할 수 있을 정도의 효과를 목표로 한다.

따라서 본 프로젝트의 전제 조건은 다음과 같다. STT나 형태소 분석기의 오인식으로 발생하는 몇몇의 통역 오류는 현실적인 프로젝트 진행 과정을 고려하여 해결하지 않는다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

추후의 이 기술들이 개선되었을 때 자연스럽게 해결 될 문제라고 판단한다. 또한 한국어의 중의적 표현은 고려하지 않고 직역한다. 중의적 표현의 특징은 해학이나 풍자 등에 활용되며 의미의 다양성으로 문학 작품의 예술성을 높이고 싶을 때 사용한다. 하지만 중의적 표현을 전달하지 못하는 것이 전체 영상의 내용을 이해하는데 큰 장애가 되지 않기 때문에 중의적 표현의 처리 기능은 구현하지 않기로 한다.

2.2. 연구/개발 내용

2.2.1 단어 및 수화 DataSet 구축

2.2.1.1 Crawling

크롤링은 웹 페이지를 그대로 가져와 해당 페이지 내의 데이터를 추출해 내는 행위를 말한다. 본 프로젝트에서는 단어마다 단어명, 품사, 의미, 수화 영상 Data를 수집하기 위해 사용한다. 국립국어원한국수어사전 홈페이지 (<http://sldict.korean.go.kr/front/main/main.do>)의 수화 영상을 크롤링한다. 사용되는 파이썬 라이브러리로는 다음과 같다.


1. Requests - HTML을 읽어오는 과정을 수행
2. Beautiful Soup - 웹 데이터 크롤링 또는 스크래핑을 할 때 사용하는 Python 라이브러리로 HTML에서 원하는 데이터를 손쉽게 parsing할 수 있도록 한다.
3. re (regular expression) - 웹페이지에서 크롤링을 할 때 불필요한 문자를 제거하기 위해 정규표현식 라이브러리를 사용한다.
4. cv2 (openCV) - 수어사전 홈페이지에서 제공하는 동영상의 경우 1배속과 0.5배속이 더해져 있다. 따라서 openCV로 동영상의 (전체 프레임 수) / 3의 프레임 수를 구해 수화 영상 data를 재가공한다.

2.2.1.2 MongoDB

대용량 빅데이터를 처리하기 용이한 NoSQL인 MongoDB를 사용한다. DB는 수어사전 홈페이지에서 제공하는 단어를 저장하는 Dictionary, 사용자가 업로드한 Video 및 음성파일, 자막파일을 저장하는 Video로 나뉜다.

Dictionary

- 수어사전 홈페이지에서 제공하는 단어들 중 일상생활 수어 약 14,000개를 저장한다.
- Collection으로는 일상생활 단어가 담겨 있는 Basic Collection, 숫자 단어가 담겨있는 Number Collection, 지화 단어가 담겨있는 FingerLanguage Collection이 있다.

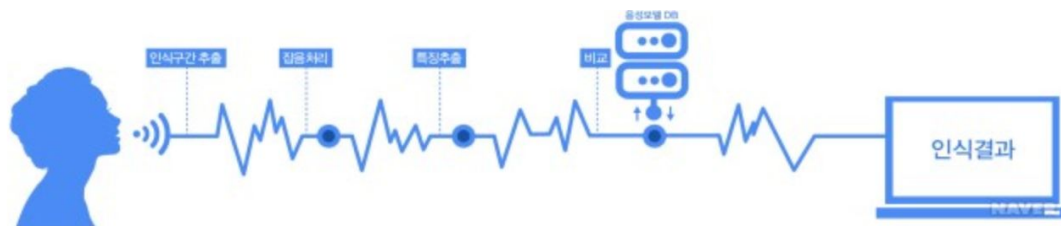
 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

- 각 Collection의 Documents에는 단어명을 나타내는 word_name, 품사정보가 담겨있는 part, 수화동영상의 url을 갖는 video_url, 프레임 수 정보를 가지고 있는 frame Fields가 있다.

Video

- Collection으로는 사용자가 upload한 동영상의 정보를 저장하는 Upload_Video가 있다.
- 사용자가 업로드한 동영상에서 음성 추출, 자막 생성을 하기 위해 불러오며 재가공하게 된다.

2.2.2 동영상의 음성에서 자막파일 추출 - STT(Speech to Text)



<그림 2. STT 기술의 일반적인 구성>


(출처: <https://terms.naver.com/entry.nhn?docId=3386818&cid=58369&categoryId=58369>)

음성인식이란 컴퓨터가 마이크와 같은 소리 센서를 통해 얻은 음향학적 신호(acoustic speech signal)를 단어나 문장으로 변환시키는 기술을 말한다. 최근에는 많은 기업들이 머신러닝 기반 음성 처리 프로그램의 개발자 환경을 지원하는 응용프로그램 인터페이스(API)를 제공하고 있다. 대표적인 API로는 구글의 Cloud Speech-to-Text, IBM의 Watson Speech-to-Text 등이 있다.

카카오나 네이버 같은 우리나라의 소프트웨어 기업들도 개발자들은 위한 음성 API를 제공하고 있지만 자막 제작을 위한 시간 정보를 제공하지 않는다. 따라서 본 프로젝트는 time stamp를 제공하는 Google Cloud Speech-to-Text를 사용한다. 입력 동영상은 음성인식 정확도를 위해 44100Hz으로 설정한다.

2.2.3 텍스트 수화 번역 - 형태소 분석

한국어와 외국어의 문법체계가 다르듯이 수화도 하나의 독립된 언어이기 때문에 그 자체만의 문법 체계를 가지고 있다. 아래는 한국수화 문형사전에서 발췌한 수화 문장 구조이다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

☑ 조사를 거의 쓰지 않거나 조사 대신 다른 명사('비교')를 사용한다.



[나는 남시보다 축구가 더 좋다]

【주어(대명사)+ 부사어(명사+ 명사)+ 서술어(서술절=주어(명사)+ 서술어(형용사)】

☑ 종결형 표지를 단어 '끝' 으로 사용한다.



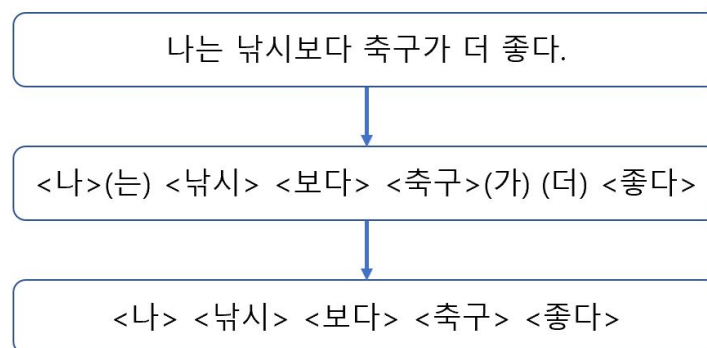
[아버지가 일본으로 건너가셨다.]

【주어(명사)+ 부사어(명사)+ 서술어(동사+ 시제 표지(과거: 끝))】

<그림 3. 수화 문장 구조>


(출처: 국립국어원 (한국 수화의 구조를 밝히다 - 『한국수화 문형사전』 발간))

위에서 보는 바와 같이 수화의 문법 구조는 조사를 거의 쓰지 않고 종결형 표지를 문장 뒤에 배치하는 등 한국어와는 문법 구조에서 다른 점이 많다. 따라서 청각장애인들이 사용하는 자연수화의 문장 구조를 이해하고 번역해야한다. 변환된 텍스트를 수화로 번역하기 위해 필요한 단계는 총 3가지 이다.



<그림 4. 텍스트 형태소 구분 및 추출>

먼저 형태소 분석 기술인 KonlPy를 이용하여 문장의 형태소를 분석한다. 예를들어 '나는 남시보다 축구가 더 좋다.'라는 문장을 <나>(는) <남시> <보다> <축구>(가) (더) <좋다>로 구분한다. 그 다음 불필요한 조사를 빼고 의미있는 <나> <남시> <보다> <축구> <좋다>의 형태소들만 남겨둔다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18




<그림 5. 형태소와 수어 매칭>

마지막으로 각 형태소들과 의미가 맞는 수화 동영상을 매칭한다.

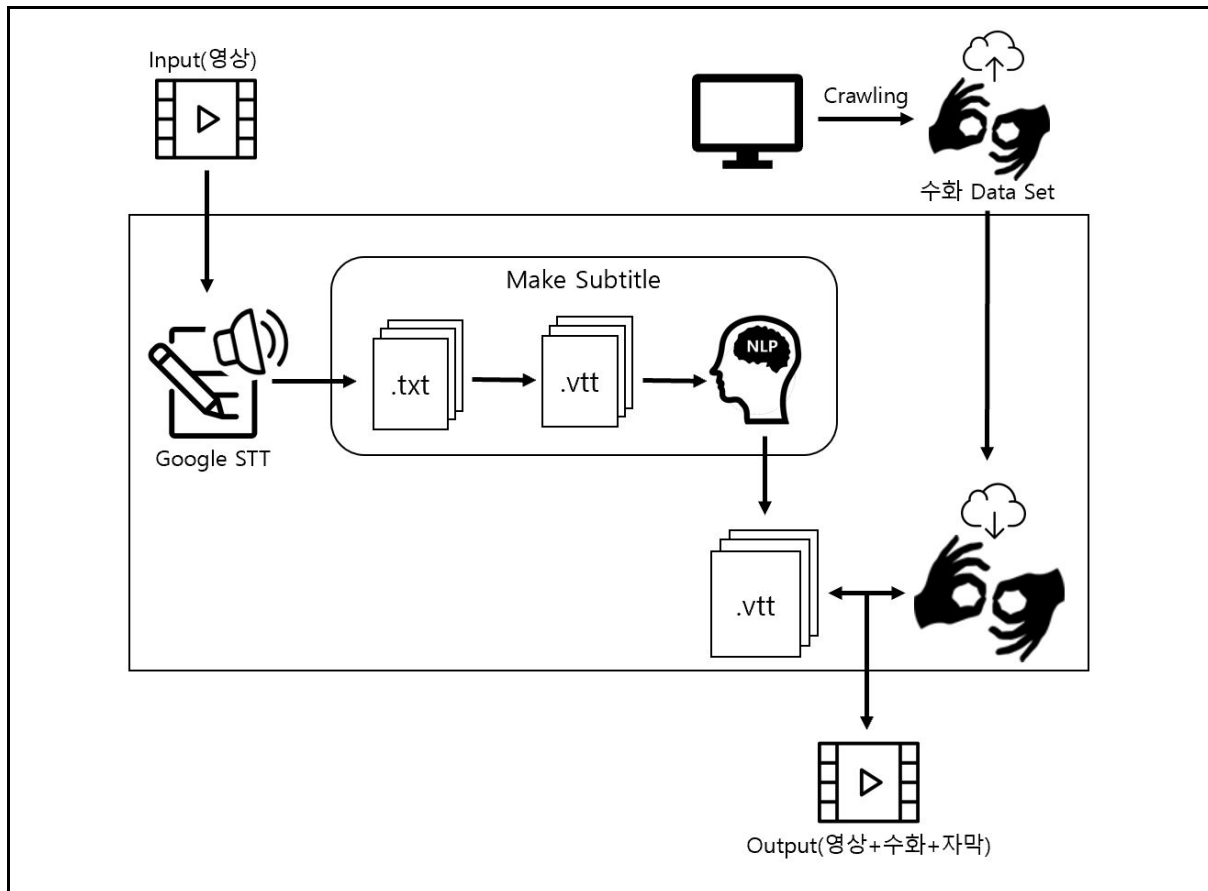
2.2.4 수화 Data Mapping - 단어간 유사도 측정

수화 문법 체계로 재구성한 형태소들에 적절한 수화를 매핑시켜주기 위해 수화 데이터베이스에서 쿼리문을 이용하여 적절한 수화동영상을 가져온다. 각 형태소에 맞는 수화 동영상을 가져와 이어붙인 뒤 동영상 자막의 싱크 정보와 수화통역 영상 길이 간의 상관관계를 계산하여 싱크를 맞춰준다. 만약 수화와 매핑을 할 때 동음이의어가 나올 경우를 대비하여 디비에 참조단어 속성을 추가적으로 저장한다. 문장의 동음이의어 앞 뒤의 단어를 가져오고 디비에 저장되어 있는 참조단어와 비교한다. 만약 참조단어에 비교 단어가 존재하면 그것을 선택하고 없다면 단어간 유사도를 계산하여 유사도가 더 높은 것으로 선택한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 인(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


2.3. 개발 결과

2.3.1. 예상 시나리오



<그림 6. 예상 시나리오>

사용자가 보고자하는 동영상을 수화 인 플레이어에 업로드하면 서버에서는 google-STT API를 사용해 동영상의 음성을 text 및 자막으로 변환한다. 변환된 text는 형태소 분석기 Konlpy를 이용하여 수화에 필요한 형태소만을 분류한다. 이후에 텍스트 및 자막과 매핑되는 수화 동영상을 연결하여 수화통역 동영상을 만든 뒤 서버 클라우드에 저장한다. 그 다음 웹페이지에 등록된 동영상과 자막, 수화통역 동영상을 전송해준다. 최종적으로 사용자가 동영상을 재생하면 선택에 맞게 자막, 수화를 영상과 함께 보여준다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


2.3.2. 시스템 기능 요구사항

번호	내용	중요도
FR1	사용자가 원하는 동영상을 웹 페이지에 업로드한다.	상
FR2	사용자가 자막만을 제공받을 지 수화와 함께 제공받을 지의 옵션을 선택한다.	하
FR3	업로드 된 동영상을 서버 클라우드 및 데이터베이스에 저장한다.	상
FR4	저장된 동영상에서 구글 STT API를 사용해 음성을 텍스트로 변환한다.	상
FR5	텍스트를 자막파일로 변환하여 저장한다.	상
FR6	자막 파일의 내용을 Konlpy를 사용하여 형태소 분석을 한 뒤 수화 문법 체계와 맞게 재구성한다.	상
FR7	재구성한 자막 파일을 바탕으로 적절한 수화 동영상과 매치 후 이어붙여 수화로 통역된 영상을 만든다.	상
FR8	사용자에게 최종적으로 사용자가 등록했던 동영상과 자막, 그리고 수화 통역 영상을 함께 제공한다.	상

<표 2. 시스템 기능 요구사항>

2.3.3. 시스템 비기능(품질) 요구사항

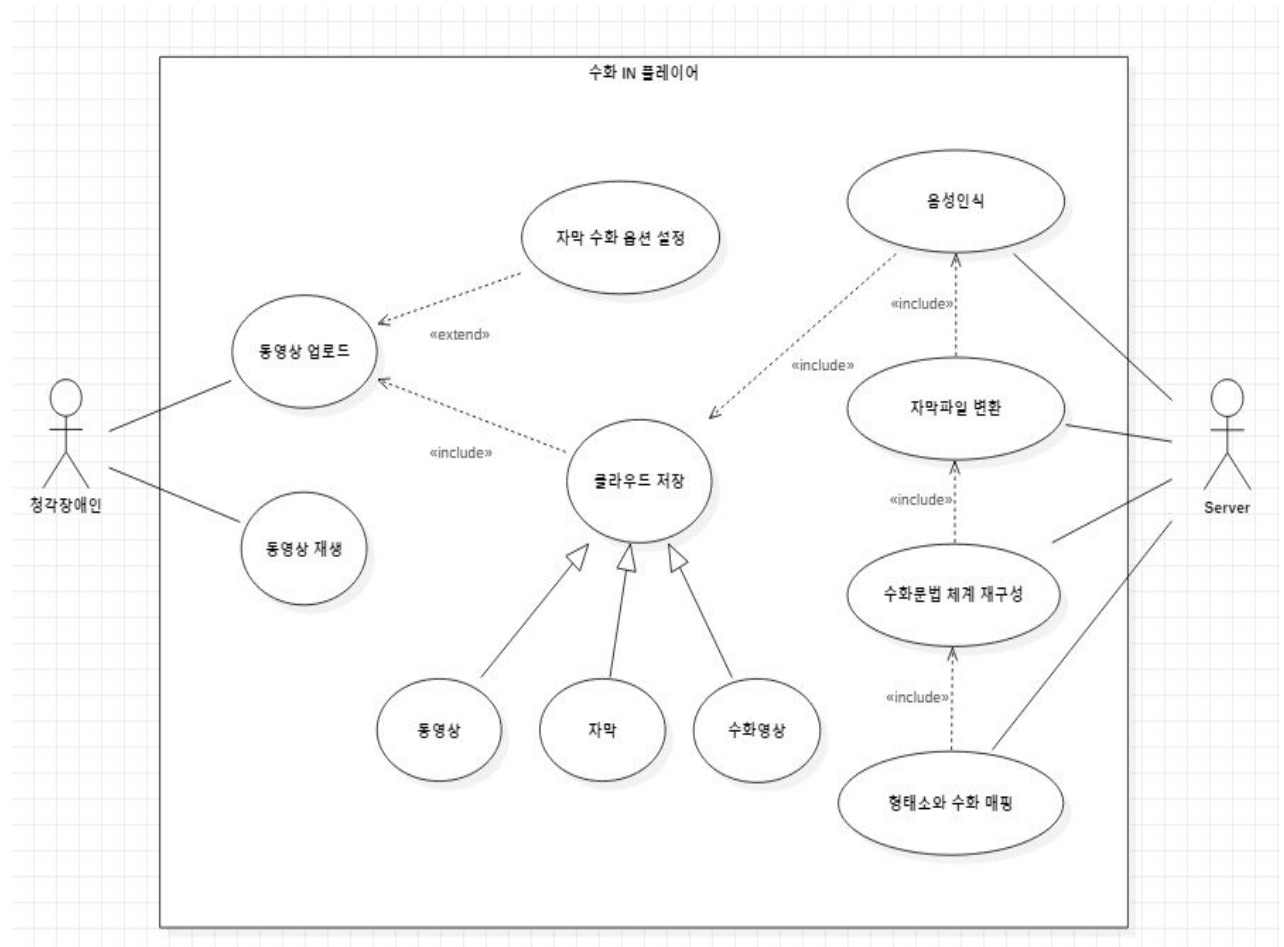
번호	내용	품질 속성
NFR1	데이터 베이스는 적어도 15000개 이상의 수화 데이터를 담을 수 있어야 한다.	Efficiency(효율성)
NFR2	시스템은 사용자에게 플레이어 매뉴얼을 제공해야 한다.	Learnability(학습성)
NFR3	사용자는 몇번의 클릭으로 웹페이지의 기능들을 파악할 수 있어야 한다.	Understandability (이해성)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


NFR4	시스템은 사용자가 등록한 영상의 정보를 외부에 유출하지 않도록 한다.	Legislative(합법성)
NFR5	시스템은 등록된 영상을 서버 클라우드에 빠르게 저장한다.	Efficiency(효율성)
NFR6	동영상 업로드 후 통역까지의 처리시간이 영상시간의 1/20을 넘지 않도록 한다.	Efficiency(효율성)

<표 3. 시스템 비기능 요구사항>

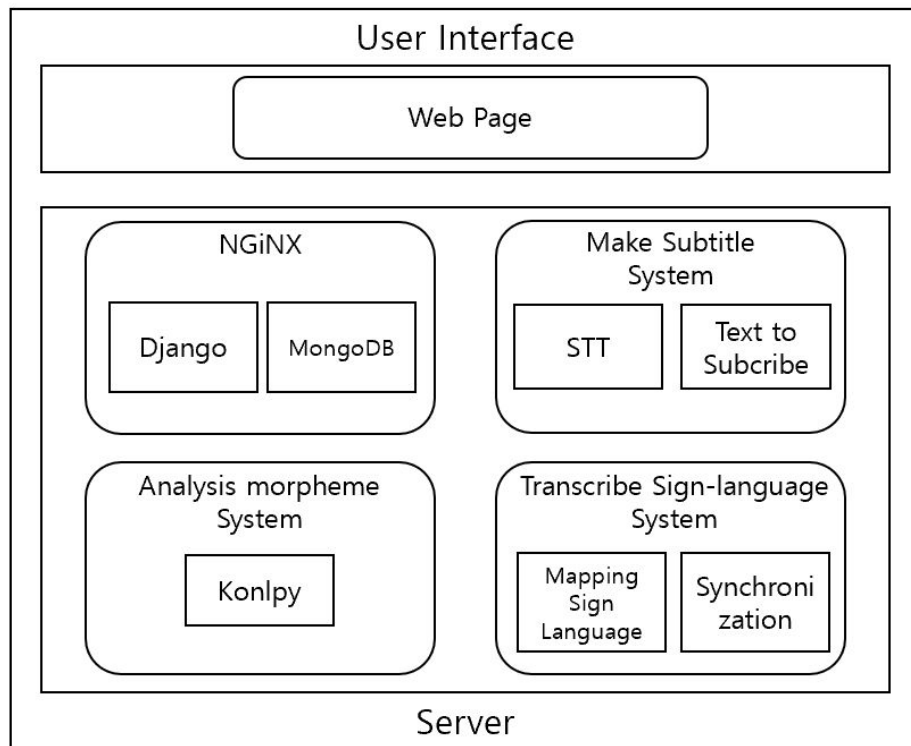
2.3.3. Use Case Diagram



<그림 7. Use case Diagram>

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


2.3.4 시스템 구조



<그림 8. 시스템 구조도>

2.3.5 결과물 목록 및 상세 사양

대분류	소분류	기능	형식	비고
프론트 엔드	UI	자막 및 수화 제공에 대한 사용자 입력을 받는다	함수	
		자막 및 수화 파일을 동영상과 함께 재생한다.	함수	
	통신	입력받은 영상을 서버에 전송		
백엔드	웹서버 구축	AWS와 Nginx를 이용한 웹 서버 구축		
	음성 텍스트화	구글 Speech-to-Text API를 이용하여 영상에서 받은 음성데이터를 텍스트로 변환하여 출력한다.	API	
	형태소 분석	한국어를 수화 문법으로 변환시키기 위해 형태소분석 기법을 활용하여 문장을 분석하고 수화 문법 구조로 재구성한다.	함수	
	DB	MongoDB를 사용하여 수집한 수화 데이터 및 업로드 영상을 저장하고 관리한다.		

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

2.4. 기대효과 및 활용방안

수화 人(in) 플레이어 프로젝트를 통해 기대하는 효과 및 활용방안은 다음과 같다. 원하는 영상을 플레이어로 재생하면 소리를 인식하여 자막과 수화를 선택적으로 서비스를 제공받을 수 있다. 즉, 청각장애인도 다양한 분야의 동영상 콘텐츠를 제한없이 이용할 수 있다. 또한 내용의 이해도가 향상되기 때문에 청인들과의 정보 격차 또한 줄어들 것으로 예상된다. 수화 人(in) 플레이어가 상용화 된다면 그에 대한 가치는 함께 발전할 수 있다고 본다.

아울러 음성을 수화로 통역해주는 시스템은 동영상에만 국한된 것이 아닌, 강의나 회의같은 상황에서 실시간으로 청인의 말을 인식하여 수화로 통역하여 청각장애인들이 청인과 수월하게 소통할 수 있도록 해줄 것이다.

3. 배경 기술

3.1. 기술적 요구사항


3.1.1. 개발 환경

- 운영체제 : Mac OS, Window 10, Ubuntu 16.04
- 개발 언어 : Python, HTML, CSS, JavaScript
- 서버 : AWS / Django / nginx
- 데이터베이스 : MongoDB
- Package: Konlpy
- IDE : Pycharm, Atom

3.1.2. Front-End

3.3.2.1 BootStrap

- 웹 디자인 개발에 사용할 시간을 최소화 하기 위해서 반응형 Front-End Framework인 부트스트랩(BootStrap)을 사용한다. 단일 코드로 데스크탑 뿐만 아니라 어떠한 모바일 환경에서도 호환가능하게 만들 수 있다.
- 사용자가 처음 웹 사이트에 접속 했을 때 사용방법을 한눈에 알아볼 수 있도록 깔끔하며 필요한 기능만 들어가 있도록 구현해야 한다. 본 프로젝트에서는 UI 구성이 깔끔한 1-page의 부트스트랩 template인 Agency를 사용하였다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

3.1.3. Back-End

3.1.3.1. NginX

- NginX는 트래픽이 많은 웹사이트를 위해 설계한 비동기 이벤트 기반 구조의 웹서버 소프트웨어로 가벼움과 높은 성능을 갖는다.
- 본 프로젝트는 다수의 사용자가 사용할 수 있는 웹사이트를 개발하는 데에 목적을 두고 있으므로 Ubuntu 16.04의 NginX 위에서 Django로 동작시킬 예정이다.

3.1.3.2. Django

- Django는 Python으로 작성된 오픈 소스 웹 애플리케이션 프레임워크로 모델-뷰-컨트롤러 패턴을 따른다.
- 웹 개발에서 번거로운 요소들을 새로 개발할 필요 없이 내장된 기능만을 이용해 빠른 개발을 할 수 있어 개발시간의 단축이 된다.


3.1.3.3. MongoDB

- 대용량 빅데이터를 처리하기 용이한 NoSQL로 관계형 데이터베이스가 아닌 다른 형태의 데이터 저장 기술을 의미한다.
- MongoDB는 Schema-less 구조를 취하여 다양한 형태로 데이터를 저장할 수 있고, Read & Write 성능이 뛰어나다.
- MongoDB는 사용자가 upload한 동영상과 동영상을 재가공해 생성된 VTT파일, 수화 영상 및 수화 사전 데이터 베이스를 구축하는데에 사용된다.

3.1.4. 자막/수화 생성

3.1.4.1. Google Cloud Speech-to-text API

- 짧거나 긴 오디오에 사용할 수 있는 머신러닝 기반 음성 텍스트 변환 기술이다.
- 사용하기 간편한 API로 강력한 신경망 모델을 적용하는 Google Cloud Speech-to-Text를 사용하면 개발자가 오디오를 텍스트로 변환할 수 있다. 이 API는 글로벌 사용자층에 대응하여 120개 이상의 언어와 방언을 인식한다. 음성 명령 및 제어 기능을 구현하고 콜센터의 오디오를 텍스트로 변환하는

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


등의 작업을 할 수 있고 Google의 머신러닝 기술을 사용하여 실시간 스트리밍 또는 사전 녹음 오디오를 처리할 수 있다.

3.1.4.2. Konlpy

- 한국어 정보처리를 위한 파이썬 패키지 이다.
- Hannanum, Kkma, Komoran, Mecab, Twitter 등 다양한 tag 패키지를 지원한다.
- 시스템 사전에 신조어나 대명사 등 사용자가 원하는 항목을 추가할 수 있어 형태소 분석 기능을 더욱 유연하게 만들 수 있다.

3.1.4.3. Gensim 패키지 Word2Vec

- word2vec은 CBOW 방식과 Skip-Gram 방식의 단어 임베딩을 구현한 C++ 라이브러리로 구글의 있던 Mikolov 등이 개발하였다.
- 파이썬에서는 gensim이라는 패키지에 Word2Vec이라는 클래스로 구현되어 있다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

3.2. 현실적 제한 요소 및 그 해결 방안

3.2.1. 하드웨어

많은 양의 수화 데이터와 유저가 업로드한 동영상을 저장하고 재가공한 영상을 배포하기 위해서 서버 컴퓨터가 따로 존재해야 한다. 이를 해결하기 위해 AWS의 EC2 인스턴스의 Ubuntu 16.04 운영체제를 활용한 서버 컴퓨터를 사용한다.

3.2.2. 소프트웨어


본 프로젝트는 영상에서 인식되는 문장을 텍스트 및 수화로 통역해야 하는데, 인식되는 문장에 맞게 수화영상을 매핑하고 연결하는 것에 대한 기술적 자문이 필요하다.

3.2.3. 데이터

대한민국에서 수화에 대한 정보를 얻을 수 있는 대표적인 웹 사이트로 ‘국립국어원 한국수어사전’이 있다. 사이트에 들어가보면 사용하기 어려운 UI 디자인과 수화 데이터의 업로드가 오래되었다는걸 확인할 수 있다. 이처럼 우리가 수화데이터를 얻기에 어려움을 겪고, 서비스를 제공하기 위한 데이터가 부족하다. 따라서 부족한 데이터는 직접 촬영하여 추가할 예정이다.

3.2.4. 기타

영상 콘텐츠의 양이 방대하기 때문에 어느 범위 까지 포함할 것인지 지속적인 논의가 필요하며, 청인들이 사용하는 한국어와 청각장애인들이 사용하는 수어의 언어체계는 차이가 있다. 이를 고려하여 한국어 텍스트를 수화로 변환할 수 있는 문장구조로 변환시키는 과정이 필요하다. 따라서 본 프로젝트에 대한 청각장애인의 견해와 수화에 대한 자세한 전문적 지식이 필요하다. 국민대학교 중앙동아리 ‘손말사랑회’ 수화 동아리에 문의하여 전문가와의 연락할 예정이다.


 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

4. 프로젝트 팀 구성 및 역할 분담

이름	역할
윤지은	<ul style="list-style-type: none"> - Team Leader - 수화 Data Structure 개발 및 데이터베이스화 - 자막 변환 기능 구현
배진영	<ul style="list-style-type: none"> - 자연어 분석 및 처리 기능 구현(형태소 분석) - 수화 문법 체계로 재구성하는 기능 구현
이희지	<ul style="list-style-type: none"> - INPUT 영상 인코딩 - 음성 인식 및 텍스트 변환(Speech-to-text) - 수화 Mapping 기능 구현
정창희	<ul style="list-style-type: none"> - 수화 단어 및 동영상 웹 크롤링 - 수화 Data Structure 개발 및 데이터베이스화 - 웹 백엔드 개발
축이림	<ul style="list-style-type: none"> - UI 구상 및 설계

5. 프로젝트 비용

항목	예상치 (MD)
아이디어 구상	35
개발 환경 구축	5
수화 Data Structure 개발 및 DB화	40
음성 인식 (STT)	20
형태소 분석	40
질의문 작성	20
웹 서버 구축	20
성능 테스트 및 오류 수정	10

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


6. 개발 일정 및 자원 관리

6.1. 개발 일정

항목	세부내용	1월	2월	3월	4월	5월	비고
요구사항분석	요구 분석						
	주제 회의						
관련분야연구	주요 기술 연구						
	관련 시스템 분석						
설계	시스템 설계						
구현	코딩 및 모듈 테스트						
테스트	시스템 테스트						
평가	최종 평가 및 전시회						

6.2. 일정별 주요 산출물


마일스톤	개요	시작일	종료일
계획서 발표	요구 사항 분석 및 관련 기술 연구 산출물 : 1. 프로젝트 수행 계획서 2. 프로젝트 발표 슬라이드쇼 3. 프로젝트 계획서 발표 자료	2019-01-01	2019-03-15
설계 완료 및 개발 환경 설치	시스템 설계 완료 개발에 필요한 기술 정보 수집 개발 환경 구축 완료 산출물 : 1. 시스템 설계 사양서 2. 개발환경	2019-03-16	2019-03-20
중간 평가	수화 단어 사전 및 수화 모델링 DB 구현 완료 Google STT API를 이용한 결과 문장 전처리 기능 구현 완료 산출물 : 1. 프로젝트 중간 평가 요약 보고서 2. 프로젝트 중간 평가 발표 자료	2019-03-21	2019-04-19

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

	3. 계획서 수정본 4. 프로젝트 중간 소스코드		
구현 완료	Text <-> 수화 Mapping 기능 구현 완료 수화 人(in) 플레이어 앱 구현 완료 산출물: 1. 데모 버전 2. 각 기능 소스코드	2019-04-20	2019-05-27
테스트	시스템 통합 테스트 산출물: 1. 최종 버전	2019-05-28	2019-05-30
최종 보고서	최종 보고 산출물: 1. 최종결과보고서 2. 최종발표자료 3. 포스터 및 소개 책자 4. 최종 발표 및 시연 동영상 5. 프로젝트 최종 소스코드	2019-05-31	2019-05-31

6.3. 인력자원 투입계획

이름	개발항목	시작일	종료일	총개발일(MD)
전원	아이디어 구상	2019-02-04	2019-03-10	35
전원	시스템 설계	2019-03-13	2019-03-17	5
윤지은/ 정창희	수화 동작 데이터 추출	2019-03-13	2019-03-20	10
정창희	데이터 베이스 설계 및 구현	2019-03-20	2019-04-03	15
정창희	웹 백엔드 개발	2019-03-13	2019-04-11	30
이희지	소리 데이터 추출	2019-03-13	2019-03-22	10
이희지	소리를 텍스트로 전환 및 기능 추가	2019-03-13	2019-03-23	10
배진영	형태소 분석	2019-03-23	2019-04-26	35
윤지은	자막파일 생성	2019-03-28	2019-04-15	25
이희지	수화데이터와 텍스트 매핑	2019-04-14	2019-04-18	15
전원	시스템 통합	2019-05-10	2019-05-14	5

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18


전원	테스트 및 수정	2019-05-15	2019-05-29	15
----	----------	------------	------------	----

6.4. 비 인적자원 투입계획

항목	Provider	시작일	종료일	Required Options
서버용 PC 1대	AWS	2019-03-16	2019-05-31	Billing Account
개발용 PC 4대	SAMSUNG HP LG APPLE	2019-03-16	2019-05-31	

7. 참고 문헌

번호	종류	제목	출처	발행년도	저자	기타
1	논문	음성인식 기반의 사이버 수화통역사 개발	경인교육대학교	2015년	김성현	
2	웹페이지	한국수어 사전	국립국어원	-	국립국어원	
3	논문	2017년 한국수어 사용 실태 조사 연구	국립국어원	2017년	이대섭	
4	기사	형식적인 자막방송만, 수화,해설방송은 가끔 시늉만	미디어오늘 (http://www.mediatoday.co.kr/?mod=news&act=articleView&idxno=132500)	2016년	금준경	
5	기사	“선거방송, 수화와 자막 동시에 제공하라”	BeMinor (http://beminor.com/detail.php?number=4586)	2012년	홍권호	
6	기사	“靑 국민청원 답변 영상에 자막,수화 넣어주세요”	국민일보 (http://news.kmib.co.kr/article/view.a	2018년	우승원	

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	수화 人(in) 플레이어	
	팀 명	전지적 수화 시점	
	Confidential Restricted	Version 2.3	2019-APR-18

			sp?arcid=0012575346)			
7	기술개발 보고서	청각 장애인용 모바일 애플리케이션 가이드북	SlideShare (https://www.slideshare.net/teaminterface/ss-27477625)	2013	지식경제부	
8	기사	한국 수화의 구조를 밝히다 - 『 한국수화 문형사전』 발간	국립국어원 (https://www.korean.go.kr/nkview/onlineletter/20070401/print/02.html)	2015	국립국어원	