

캡스톤 디자인 I 종합설계 프로젝트

프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템
팀 명	
문서 제목	계획서

Version	1.0
Date	2019-MAR-12

팀원	김 승환 (조장)
	김 병찬
	홍 일권
	김 영준
	김 태훈

CONFIDENTIALITY/SECURITY WARNING

이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인 I 수강 학생 중 프로젝트 “Voice_Book”를 수행하는 팀 “2019_23”의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 “2019_23”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

문서 정보 / 수정 내역

Filename	1 차수행계획서.doc
원안작성자	김승환
수정작업자	

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2019-03-12	팀 전원	1.0	최초 작성	

	국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
		프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
		팀 명		
		Confidential Restricted	Version 1.0	2019-MAR-12

목 차

1	개요	4
1.1	프로젝트 개요	4
1.2	추진 배경 및 필요성	4
2	개발 목표 및 내용	5
2.1	목표	5
2.2	연구/개발 내용	6
2.3	개발 결과	7
2.3.1	결과물 목록 및 상세 사양	7
2.3.2	시스템 기능 및 구조	7
2.4	기대효과 및 활용방안	7
3	배경 기술	8
3.1	기술적 요구사항	8
3.2	현실적 제한 요소 및 그 해결 방안	9
3.2.1	하드웨어	9
3.2.2	소프트웨어	9
3.2.3	기타	9
4	프로젝트 팀 구성 및 역할 분담	10
5	프로젝트 비용	10
6	개발 일정 및 자원 관리	11
6.1	개발 일정	11
6.2	일정별 주요 산출물	12
6.3	인력자원 투입계획	13
6.4	비 인적자원 투입계획	14
7	참고 문헌	15

	국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
		프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
		팀 명		
		Confidential Restricted	Version 1.0	2019-MAR-12

1 개요

1.1 프로젝트 개요

시각장애인들은 비시각장애인들에 비해 상대적으로 독서를 접할 수 있는 경로가 한정되어있고, 그마저도 콘텐츠가 다양하지 못하고 비용이 많이 든다는 한계가 있다.


이런 한계를 극복하기 위하여 인위적으로 사람의 소리를 합성, 텍스트를 음성으로 변환하는 텍스트 음성 변환(Text-To-Speech) 기술을 활용하여 시각장애인을 위한 오디오북을 개발할 것이다.

이 프로젝트는 제공하는 기능을 사용자가 별다른 학습 없이도 편하게 이용할 수 있게 음성만으로 조작이 가능한 UI와 사용자가 원하는 음성 데이터를 오디오에 입힐 수 있도록 하여 시각장애인들의 독서 진입장벽을 낮추고 편리성과 콘텐츠의 다양성을 갖춘 서비스를 제공하는 것을 목표로 한다.

중요한 점은 시각 장애인들이 주위의 도움을 받지 않고도 프로그램을 사용할 수 있도록 하는 것과 원하는 목소리로 책을 들을 수 있는 것이다. 음성 인식을 통해 접근성을 높이고 음성 합성을 통해 원하는 음성에 가장 가까운 음성으로 학습이 된 인공지능 모델이 음성을 선택하도록하여 목소리에 대한 저작권 문제를 해결(직접 그 목소리를 사용하는 것이 아니라 그 목소리와 가장 유사한 일반인 목소리를 사용하는 것이기 때문에, ex) 성대모사)하여 기존의 오디오북들과 차별성을 둔다.

세부적인 프로젝트의 구조 및 개발 내용은 다음과 같다.

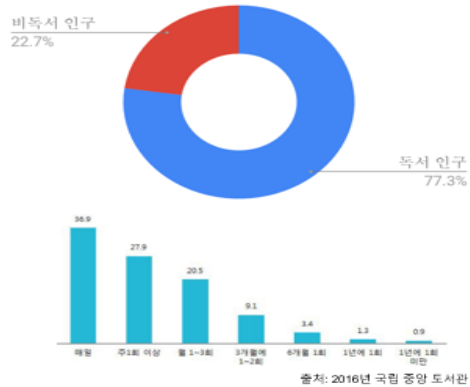
AI server에서 음성합성을 위해 deepvoice를 사용하여 인공지능 모델을 구현한다. 구현한 모델을 다양한 목소리의 음성 데이터와 텍스트 데이터를 가지고 학습시킨다. 학습시킨 모델을 API server에 전송하여 로그인, 도서 순위, 내 정보, 인공지능 모델을 적용한 음성 선택 등 필요한 기능들을 만들어 device에 적용한다.

 <div> 국민대학교 컴퓨터공학부 캡스톤 디자인 I </div>	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

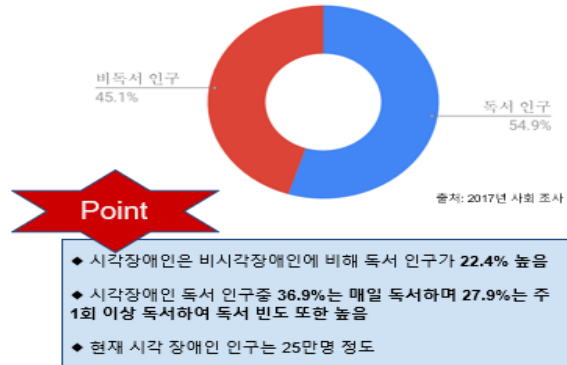
1.2 추진 배경 및 필요성

1.2.1 시각장애인 독서 시장의 크기

시각장애인 독서 현황과 빈도



비시각장애인 독서 현황



시각장애인의 인구 대비 독서 시장의 규모는 상당히 크다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

1.2.2 기존 시각장애인 독서 문제점



기존 시각장애인 독서 시스템은 시각장애인들이 접근하기 쉽지 않고, 콘텐츠 제작과 이용에 있어서 소요가 크다.

 <div> 국민대학교 컴퓨터공학부 캡스톤 디자인 I </div>	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

1.2.3 기 개발된 시스템 현황과 한계



(네이버랩스에서 음성합성 기술을 활용하여 만든 오디오북)

기존 시스템은 시각장애인의 접근이 어렵다(**UI**의 편리성, 접근성의 부재), 시스템에서 제공하는 음성만 이용할 수 있다(다양성의 부재)

2 개발 목표 및 내용

2.1 개발 목표

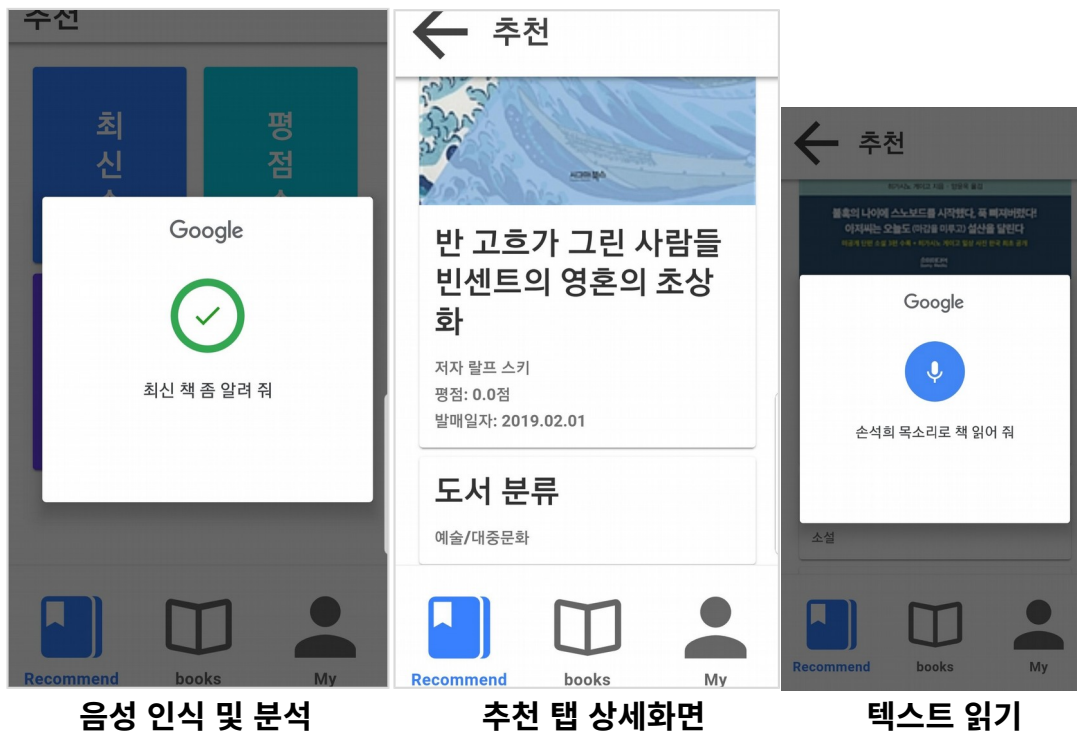
시각장애인의 독서 진입장벽을 낮추고 서비스 이용의 접근성과 편리성을 갖춘 **UI**를 구축, 원하는 음성데이터를 오디오에 적용하여 사용자 니즈에 맞춘 오디오북 기능과 이 밖에 독서에 도움을

 <div> 국민대학교 컴퓨터공학부 캡스톤 디자인 I </div>	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

줄 수 있는 다양한 편의기능을 제공하는 것을 목표로 한다.

2.2 연구/개발 내용

1) 음성 인식을 활용한 사용자 인터페이스 구현



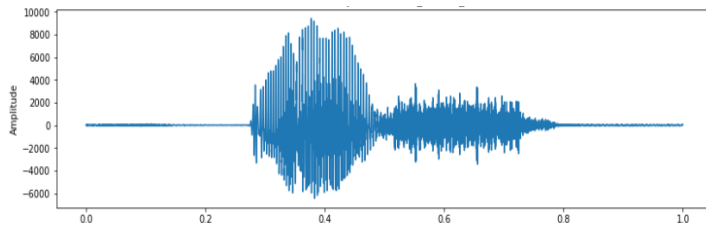
구글 음성인식 서비스를 연동해서 시각장애인들이 사용하기 쉽게 모든 **User Interface** 를 구축합니다.

2) 음성인식 기술을 활용한 AI Text Reading 연구

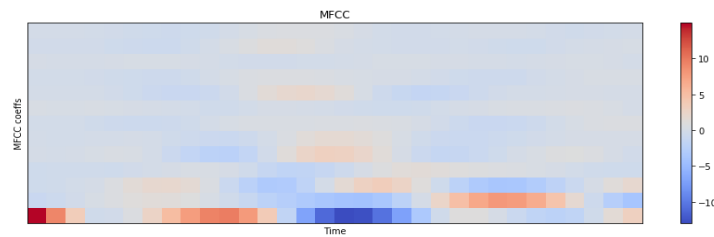
음성 데이터 입력



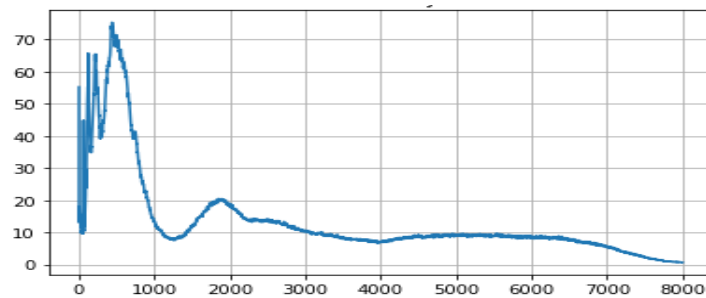
계획서		
프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
팀 명		
Confidential Restricted	Version 1.0	2019-MAR-12



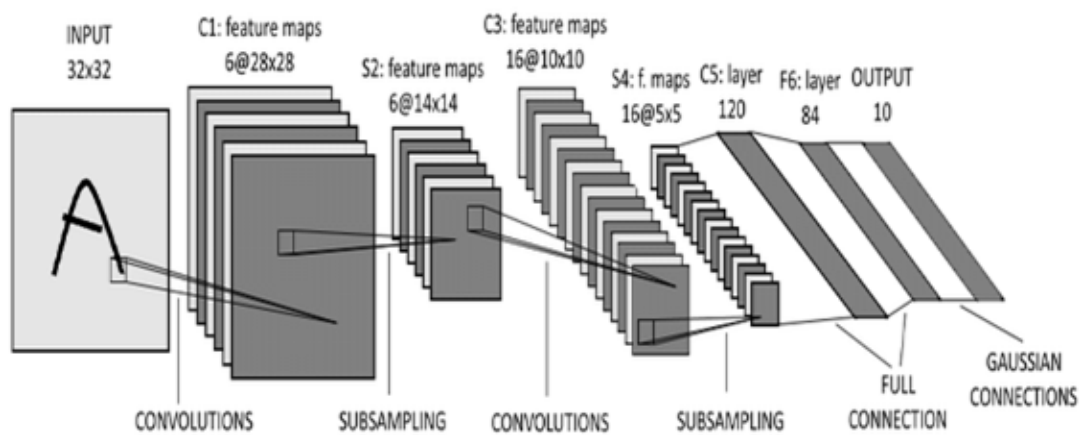
음성데이터 추출



음성 특징 추출(MFCC)



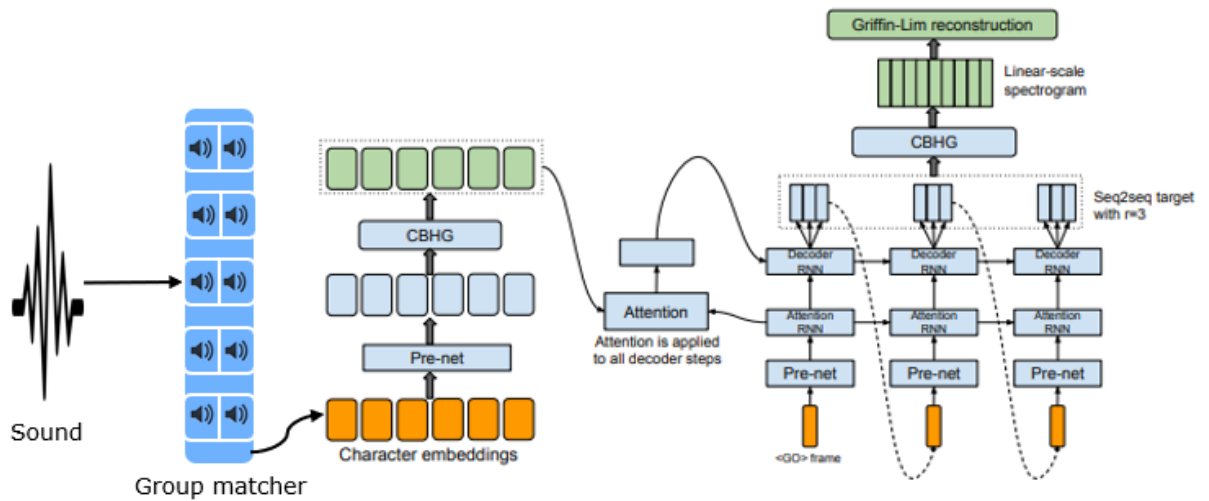
VGG16 기반 딥러닝 모델



3) TEXT 에 원하는 음성 데이터를 합성하는 기능 연구



계획서		
프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
팀 명		
Confidential Restricted	Version 1.0	2019-MAR-12



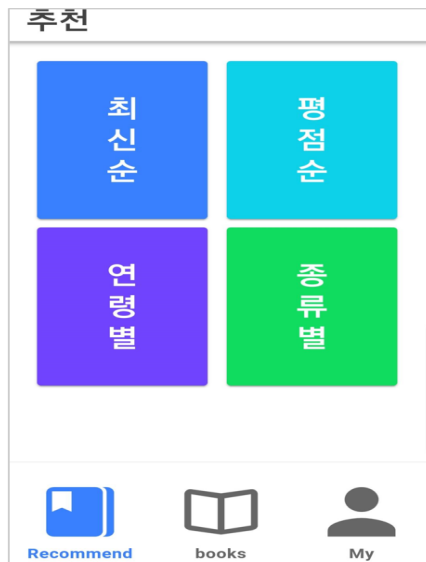
음성 합성(speech synthesis)은 말소리의 음파를 기계가 자동으로 만들어 내는 기술로, 간단히 말하면 모델로 선정된 한 사람의 말소리를 녹음하여 일정한 음성 단위로 분할한 다음, 부호를 붙여 합성기에 입력하였다가 지시에 따라 필요한 음성 단위만을 다시 합쳐 말소리를 인위로 만들어내는 기술이다.

TTS(=Text-to-Speech)라고도 한다.

4) 기타 독서에 도움을 줄 수 있는 편의기능 구현

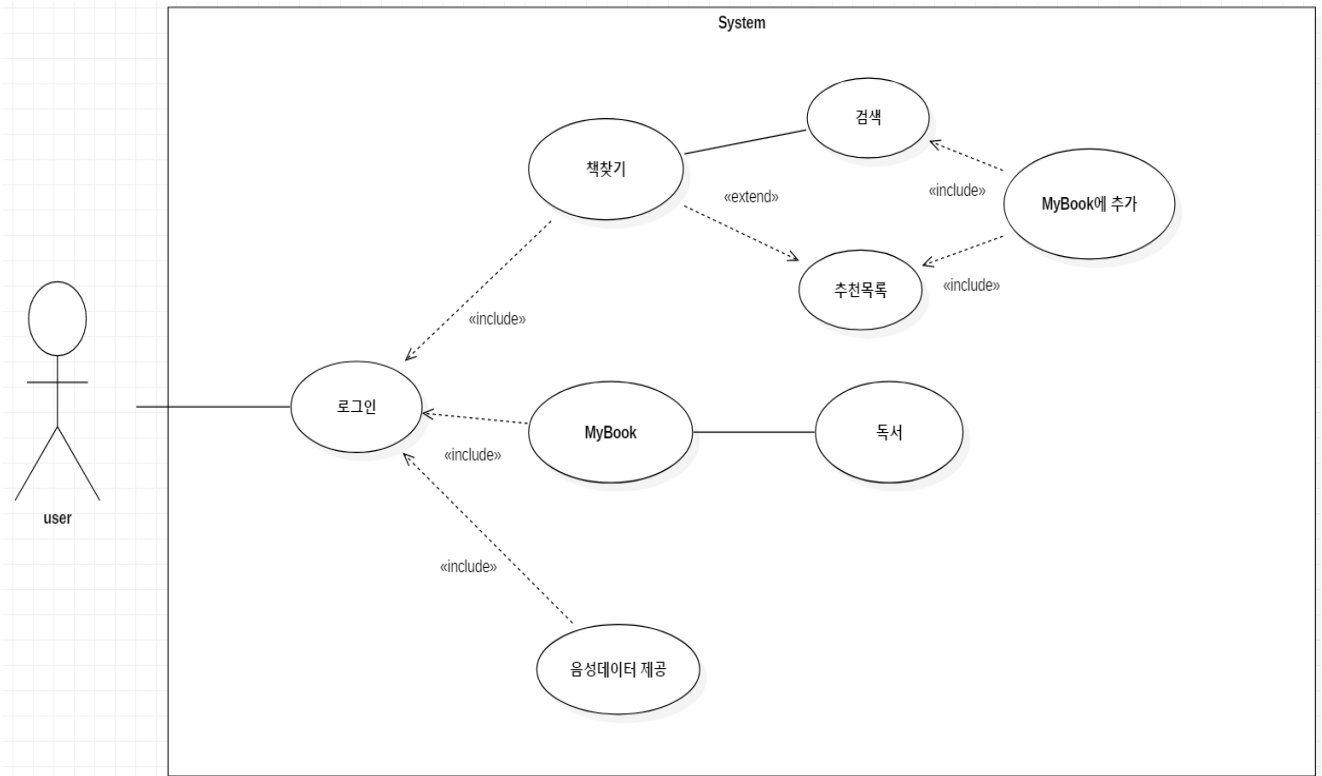
 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

각 기준 으로 책을 추천하고 추후 음성 또한 평점,성별등으로 나누어 읽어줄수 있게 할 예정이다.



2.3 결과

2.3.1 시스템 기능 요구사항

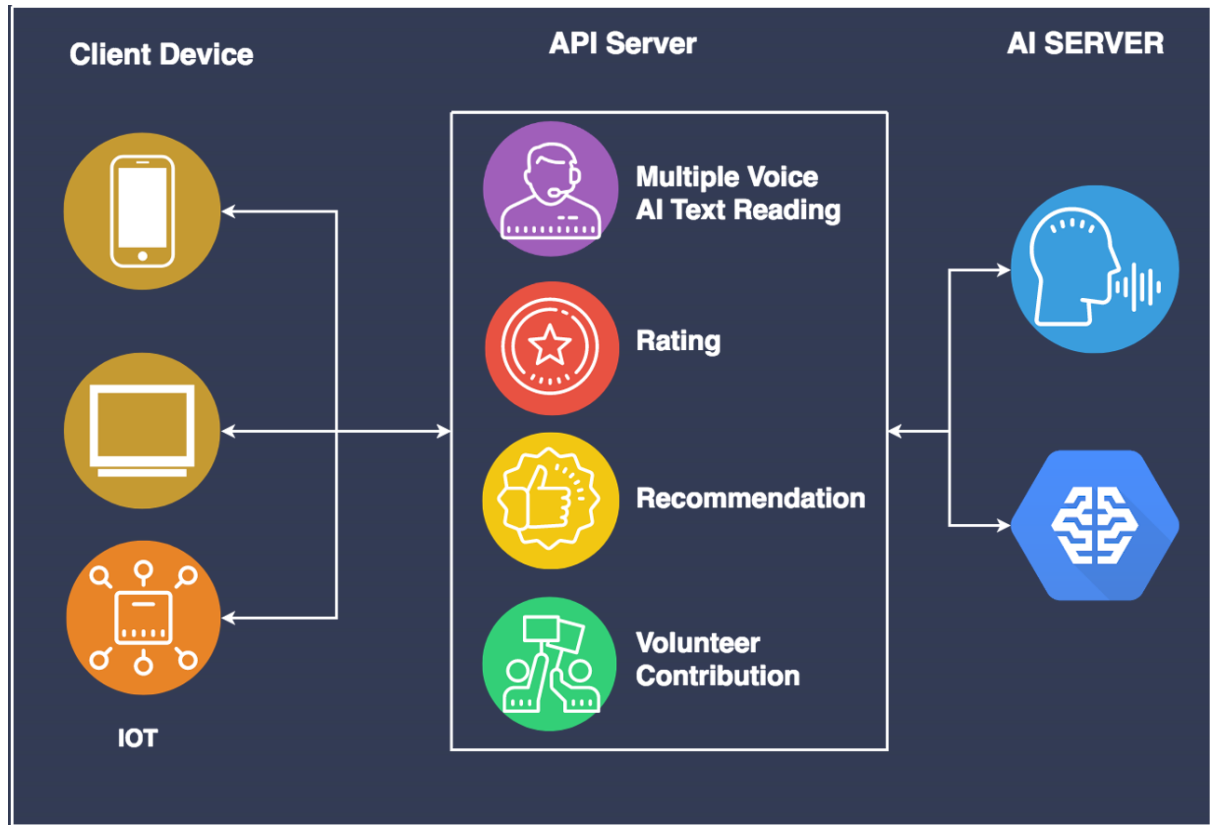


기능적 요구사항	내용
FR1 로그인	시스템을 이용하기 위해 사용자 인증을 합니다.
FR2 책찾기	사용자가 읽고 싶은 책을 찾기 위해 검색 및 추천목록의 기능을 제공합니다.
FR3 검색	책의 제목을 통해서 사용자가 읽고싶은 책을 찾아준다.
FR4 추천목록	다른 유저의 평점, 도서의 출간일등의 기준으로 책의 목록을 보여준다.
FR5 MyBook 에 추가	FR3 또는 FR4 를 이용해 찾은 도서를 자신의 MyBook (책 보관함)에 넣습니다.
FR6 MyBook	사용자가 담은 도서를 보여주고 각 도서의 진행상태(읽은 양)을 표시해 줍니다.
FR7 독서	도서의 내용을 음성으로 출력해 줍니다.
FR8 음성데이터제공	음성학습및 음성합성에 필요한 (텍스트 - 음성)를 사용자가 제공한다.

2.3.2 시스템 비기능 요구사항(미정)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

2.3.3 시스템 구조



- 웹 기반이기 때문에 어느 곳이든 접근 가능
(폰갭 사용하여 Web 을 App 으로 변환 가능)
- API server에서는 모듈 분할해서 어디든(핸드폰, 노트북, IOT 등) 쉽게 적용 가능
- A.I server에서는 음성 인식이나 음성 합성(tacotron) 구현

2.3.4 결과물 목록 및 상세 사양

대분류	소분류	기능	형식	비고
-----	-----	----	----	----

	국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
		프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
		팀 명		
		Confidential Restricted	Version 1.0	2019-MAR-12

파일	파일 저장	현재 열린 파일을 저장한다.	DLL/함수	
		다른 이름으로 파일을 저장한다	DLL/함수	
	파일 열기	문서 파일을 연다.	모듈	
출력	PDF			
	프린터			
모양	글꼴	정렬		
		폰트 바꾸기		

2.4 기대효과 및 활용방안

시각장애인이 쉽게 독서와 가까워질 수 있고, 사용자가 원하는 다양한 음성을 통해 더욱 더 편리하게 독서를 즐길 수 있다. 시각장애인을 위한 도서관 및 여러 장애인 복지단체에 프로그램을 제공하여 더 많은 시각장애인들이 도서관을 이용하여 독서를 쉽게 접할 수 있다. 언제 어디서든 핸드폰을 통해 음성 인식 기능을 통하여 시각장애로 인해 독서를 하는데 있어서의 불편함을 줄이고 독서에 대한 접근성을 높인다.

3 배경 기술

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

3.1 기술적 요구사항

구조	운영체제	개발 언어	라이브러리	기타
API server	Ubuntu	nodejs	트위터 형태소분석기	navigating
A.I server	Ubuntu	Python	Tensorflow (tacotron)	음성 인식 및 합성
Client(U.I)	*	Javascript	jquery	폰갭사용 (Web->App)

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

3.2 현실적 제한 요소 및 그 해결 방안

3.2.1 하드웨어

시각장애인들이 이 프로젝트에서 완성한 시스템을 이용할 수 있는 핸드폰이나 데스크탑이 있는지 알 수 없기 때문에 상용화할 시에 시스템을 사용함에 있어서 문제가 발생할 수 있다. 이를 해결하기 위해 시각장애인들이 다니는 도서관이나 낭독봉사자들의 핸드폰이나 데스크탑에 시스템을 제공하여 접근성을 높여 시각장애인들에게 서비스를 제공할 수가 있다.

3.2.2 소프트웨어

방대한 양의 음성 데이터를 학습시키는데에 있어서 CPU 만으로 상당한 과부하가 걸린다. 간단한 몇 문장을 학습시키는 것도 10~20 시간이 걸릴 수가 있기 때문에 GPU 도 함께 사용하여 학습시킬 필요성이 있다. 이때 필요한 GPU 를 개발비품으로 학교에 신청하여 마련하여 사용하여 해결 하려고 한다. 음성합성뿐만 아니라 음성 인식부분에서의 소프트웨어를 직접 개발하기에는 무리가 있어 구글이나 네이버에서 제공하는 오픈 API 를 사용하여 개발해야 한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

4 프로젝트 팀 구성 및 역할 분담

이름	역할
김승환	- Software Project Leader - 음성의 형태소 분석 및 네비게이팅
홍일권	- API Server 담당 및 개발
김병찬	- A.I Server 담당 및 개발
김태훈	- 음성 데이터 수집 및 전처리
김영준	- User Interface 담당 및 개발

5 프로젝트 비용

항목	예산치 (MD)
아이디어 구상 및 회의	10
API SERVER	20
A.I SERVER	30
음성의 형태소 분석 및 네비게이팅	20
음성 데이터 수집 및 전처리	20
User Interface	20
합	120

	국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
		프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
		팀 명		
		Confidential Restricted	Version 1.0	2019-MAR-12

6 개발 일정 및 자원 관리


6.1 개발 일정

항목	세부내용	3 월	4 월	5 월	6 월	비고
요구사항분석	아이디어 회의 및 구상					
	자료 수집					
관련분야연구	TTS 및 타코트론 연구					
	A.I 및 API, A.I Server 연구					
설계	시스템 설계					
구현	코딩 및 모듈 테스트					
테스트	시스템 테스트					

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

6.2 일정별 주요 산출물

마일스톤	개요	시작일	종료일
계획서 발표	프로젝트 구체화 역할 분담 및 구체화 프로젝트 설계 개발 환경 완성 (GCC 설치, 기본 응용 작성 및 테스트 완료) 산출물 : 1. 프로젝트 수행 계획서 2. 프로젝트 기능 일람표	2019-02-25	2019-03-15
설계 완료	시스템 설계 완료 산출물 : 1. 시스템 설계 사양서	2019-02-25	2019-03-20
1 차 중간 보고	데이터 학습 및 음성 학습 Back End Server 구현 완료 User Interface 구현 산출물 : 1. 프로젝트 1 차 중간 보고서 2. 프로젝트 진도 점검표 3. 1 차분 구현 소스 코드	2019-03-21	2019-04-19
구현 완료	User Interface 구현 완료 서비스 제공 시스템 구현 완료 산출물: 최종 구현 소스 코드	2019-04-20	2019-05-18
테스트	시스템 통합 테스트 산출물: A.I Text Reading 산출물	2019-05-19	2019-05-21
최종 보고서	최종 보고 산출물: 최종 보고서	2019-05-22	2019-05-31


	국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
		프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
		팀 명		
		Confidential Restricted	Version 1.0	2019-MAR-12

6.3 인력자원 투입계획

이름	개발항목	시작일	종료일	총개발일 (MD)
김승환	음성의 형태소 분석 및 네비게이팅	19-02-25	19-05-30	20
홍일권	API Server 개발	19-02-25	19-05-30	20
김병찬	A.I Server 개발	19-02-25	19-05-30	30
김태훈	음성데이터 수집 및 전처리	19-02-25	19-05-30	20
김영준	User Interface 개발	19-02-25	19-05-30	20

6.4 비 인적자원 투입계획

항목	Provider	시작일	종료일	Required Options
GPU	NVIDIA	19-04-01	19-05-29	GTX1080ti
Anaconda	Anaconda	19-02-25	19-05-29	
PyCharm Community Edition	JetBrains	19-02-25	19-05-29	
개발용 Laptop 5대	개인소유 Laptop	19-03-01	19-05-29	

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	시각장애인을 위한 음성선택형 AI Text Reading 시스템	
	팀 명		
	Confidential Restricted	Version 1.0	2019-MAR-12

7 참고 문헌

번호	종류	제목	출처	발행년 도	저자	기타
	서적					
	기사					

참고 웹사이트

http://hellogohn.com/post_one295 (음성합성 엔진)