



국민대학교
전자정보통신대학
컴퓨터공학부


캡스톤 디자인 I

종합설계 프로젝트

프로젝트 명	실시간 번역 시스템
팀 명	LTE(Language Translation Easily)
문서 제목	계획서

Version	1.0
Date	2019-MAR-10

팀원	박 지훈 (조장)
	최 창규
	공 병민
	고 준규
	박 귀환
	오 석현

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14


CONFIDENTIALITY/SECURITY WARNING

이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인Ⅰ 수강 학생 중 프로젝트 "실시간 번역 시스템"을 수행하는 팀 "LTE"의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 "LTE"의 팀원들의 서면 허락 없이 사용되거나, 재가공 될 수 없습니다.

문서 정보 / 수정 내역


Filename	계획서-실시간 번역 시스템.doc
원안작성자	박지훈, 고준규
수정작업자	최창규 공병민 박귀환 오석현

수정날짜	대표수정자	Revision	추가/수정 항목	내 용
2019-03-10	팀원 전원	1.0	최초 작성	초안 작성
2019-03-11				
2019-03-10				
2019-03-10				
2019-03-10				
2019-03-10				
2019-03-10				

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

목 차

1	개요	4
1.1	프로젝트 개요	4
1.2	추진 배경 및 필요성	4
1.2.1	실시간 번역 시스템의 추진 배경	4
1.2.2	기술의 시장 현황 및 현재 개발된 시스템 현황	4
1.2.3	현재 개발된 시스템의 문제점	5
2	개발 목표 및 내용	6
2.1	목표	6
2.2	연구/개발 내용	7
2.3	개발 결과	8
2.3.1	시스템 기능 요구사항	8
2.3.2	시스템 비 기능(품질) 요구사항	9
2.3.3	시스템 구조	9
2.3.4	결과물 목록 및 상세 사양	10
2.4	기대효과	10
3	배경 기술	11
3.1	기술적 요구사항	11
3.2	현실적 제한 요소 및 그 해결 방안	12
3.2.1	하드웨어	12
3.2.2	소프트웨어	12
3.2.3	기타	12
4	프로젝트 팀 구성 및 역할 분담	13
5	프로젝트 비용	13
6	개발 일정 및 자원 관리	14
6.1	개발 일정	14
6.2	일정별 주요 산출물	15
6.3	인력자원 투입계획	16
6.4	비 인적자원 투입계획	17
7	참고 문헌	18

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

1 개요

1.1 프로젝트 개요

사용자의 영상에서 음성을 텍스트로 바꾼 뒤 이 텍스트를 각 국가의 언어들에 맞게 번역을 한다. 그 후 번역된 텍스트를 반대쪽의 사용자들에게 실시간으로 번역된 텍스트를 자막으로 제공한다.

1.2 추진 배경 및 필요성

1.2.1 실시간 번역 시스템의 추진 배경

실시간 번역 시스템의 추진 배경으로는 글로벌 시대로 나아가고 있는 요즘 국가, 언어, 인종에 상관없이 다양한 인간관계를 형성하고 다국적 기업들의 성장이 두드러지는 현대 사회에 맞춰 서로의 언어가 달라도 의사소통을 하는데 어렵지 않게 도움이 되는 것의 필요성을 느꼈고 이로 인해 실시간 번역 시스템을 추진하게 되었다

1.2.2 기술의 시장 현황 및 현재 개발된 시스템 현황

Skype Translator

Skype Translator는 배운 적 없는 외국어도 사용할 수 있게 해줍니다. 다른 언어를 사용하는 누군가에게 Skype 영상 또는 음성 통화를 걸고 말하기만 하면 됩니다. 영상통화, 음성통화, 채팅시 사용할 수 있습니다.

번역기를 사용하면 :

- ✓ 실시간에 가깝게 대화가 다른 언어로 번역됩니다.
- ✓ 상대방이 말한 내용이 내가 사용하는 언어로 번역됩니다.
- ✓ 통화 시 화면에 번역문이 표시됩니다.
- ✓ 음성,영상 통화는 7개, 채팅은 50개 언어가 제공됩니다.

대표적인 실시간 번역 음성, 화상 채팅 서비스 중에는 스카이프가 있다. 현재까지 개발된 스카이프의 영상통화의 경우 1:1 영상통화, 그룹 영상통화로 나뉘어 있다. 또한 배우지 않은 외국어도 언어의 장벽 없이 실시간으로 대화를 나눌 수 있도록 skype translator를 지원하고 있다. Skype translator의 경우 영상 통화에서 7개의 언어를 실시간으로 번역하고, 채팅의 경우 50개의 언어번역이 제공된다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

1.2.3 현재 개발된 시스템의 문제점

대표적인 실시간 번역 음성, 화상 채팅 서비스 중에는 스카이프가 있다. 하지만 이 경우에 단점이 몇 가지 존재한다.


첫 번째로 사용자가 설정한 언어로 번역된 스크립트가 사용자에게 보여지는 시간이 충분하지 않아 사용자가 설정한 언어로 번역된 스크립트를 다 보고 읽기 전에 스크립트가 사라져 서로 의사소통에 차질을 갖게 되는 단점이 있다.

두 번째로 사용자가 설정한 언어로 번역된 스크립트의 기록을 볼 수 없다는 단점이 있다. 만약 사용자가 상대방의 말을 듣지 못하였거나 상대방이 의도한 말을 인지하지 못했을 때 번역된 스크립트를 다시 봐서 상대방과의 의사소통을 이어서 나아가야 하지만 한번 번역된 채팅의 기록을 확인을 할 수 없어서 의사소통에 차질이 생길 여지가 있다.

세 번째로 1:1 영상통화가 아닌 3명 이상의 인원이 통화를 하게 될 경우 실시간 번역 지원이 되지 않는다. 1:1로 통화를 하게 되는 경우에는 상관이 없지만 3명이상이 의사소통이 필요하게 될 경우가 있는데 예를 들어 논문을 발표하는 학회나 글로벌한 기업에서의 화상 회의같은 상황에서는 사용될 수가 없고 그 단체에서 통용되는 언어를 학습해야 참석하고 토의를 할 수 있다.

네 번째로 skype translator 실시간 번역 서비스를 사용자가 사용하기 위해 설정하는 단계가 복잡하여 서비스를 처음 이용하는 사용자들에게 어려운 단점이 있다. 이러한 경우에 사용자들이 설정을 하는 단계에서 포기를 하거나 다른 플랫폼의 서비스를 이용할 여지가 있다.

다섯 번째로 skype translator에서 한국어를 실시간 번역 서비스를 지원이 되지 않는 단점이 있다. 한국의 사용자들의 경우에 이 skype translator의 서비스를 이용할 사람들의 대부분은 어떤 외국어도 한국어로 번역되는 것을 원하지만 skype translator의 경우에 제공이 되지 않아서 사실상 한국의 사용자들에게는 큰 이목을 끌기가 어렵다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

2 개발 목표 및 내용

2.1 목표


글로벌 시대에 맞춰 학회 같은 발표에 참석한 사람들을 위한 실시간으로 각 국가의 언어로 번역이 되는 플랫폼을 개발한다. 나아가 어플리케이션 개발을 목표로 한다.

- 3명 이상의 화상 통화 서비스를 제공한다.
- 영상에서 사용언어를 다국어(중국어, 일본어, 한국어, 영어)로 번역하여 제공한다.
- 기존의 서비스와 다르게 설정 단계를 어렵지 않게 하여 접근성을 높인다.
- 사용자에게 보여지는 번역 스크립트의 기록을 일정량 저장하여 번역된 스크립트가 보여지는 시간이 이전의 서비스보다 길게 설정한다. 이를 통해 바로 의사소통을 하는데 어려움을 줄인다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

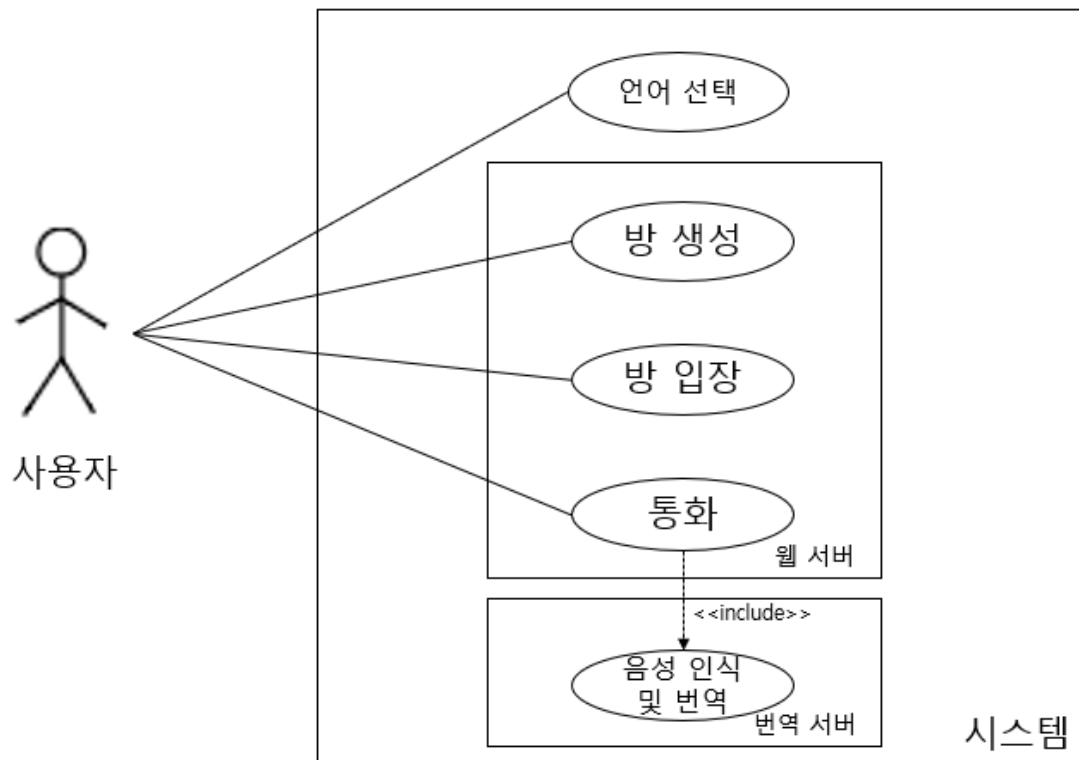
2.2 연구/개발 내용

사용자들은 React VideoCall을 통해 하나의 웹 서버에 등록된다. 사용자들이 자신의 언어를 선택 하면, 상호간의 번역될 언어가 번역 서버에 등록한다. 서버에 등록된 사용자들은 유저 인터페이스에 등록되고, 이들에게 각자 google speech-to-api를 부여하여 실시간 음성인식 기능을 확립한다. 웹에서 실행하는 사용자들은 내장 마이크에 말을 하고, 앱에서 실행하는 사용자들은 핸드폰에 대고 말을 한다. 실시간으로 인식된 문장은 각 사용자 큐에 쌓이고, 번역 서버에서 한 명씩 돌아가며 인식된 문장을 번역한다. 이 때, 문장이 인식되지 않더라도, 빈 문자열을 사용자 큐에 쌓아두고, 번역 서버에 요청된 문자열이 빈 문자열이면, 번역을 하지 않고 넘어간다. 이렇게 하면 인식된 문장이 번역 서버에 올라가지 않는 경우를 차단할 수 있다. 각각 번역된 문자열들은 웹 서버에 올라가게 되고, 프론트에서 사용자들에게 띄워주게 된다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

2.3 개발 결과

2.3.1 시스템 기능 요구사항



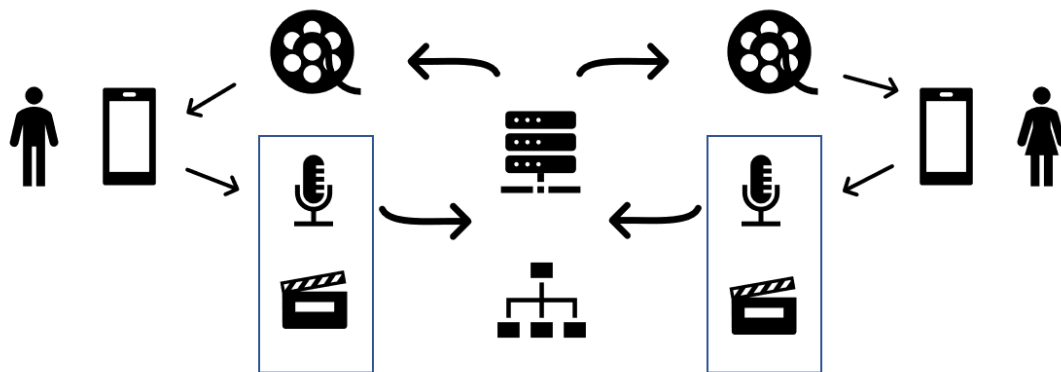
- 언어 선택:** 사용자는 웹페이지에 접근하여 본인이 사용하는 언어를 선택한다.
- 방 생성:** 사용자는 통화가 가능한 방을 생성한다.
- 방 입장:** 사용자는 이미 생성되어 있는 방에 입장한다.
- 통화:** 하나의 방에 입장된 사람이 2 인 이상인 경우 통화가 시작된다.
- 음성 인식 및 번역 :** 사용자의 음성이 웹서버에 도착하면, 번역 서버로 해당 음성을 전송하고, 번역 서버에서 음성을 인식하고 번역하여 번역된 언어를 웹서버로 전송한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14


2.3.2 시스템 비 기능(품질) 요구사항

1. **효율성(Efficiency):** 실시간 화상 통화에서 음성을 인식하고 상대방의 언어에 맞게 번역해서 보여준다.
2. **사용자 편리성(Usability):** 사용자는 자신이 사용하는 언어를 입력하고 방을 생성하거나, 방에 입장하는 방법으로 기존에 존재하는 시스템들에 비해 쉽게 번역 기능을 이용할 수 있도록 한다.
3. **사용자 편리성(Usability):** 사용자는 자신이 사용하는 언어를 입력하고 방을 생성하거나, 방에 입장하는 방법으로 기존에 존재하는 시스템들에 비해 쉽게 번역 기능을 이용할 수 있도록 한다.

2.3.3 시스템 구조



사용자는 웹페이지를 통해 접속한다. 컴퓨터에 내장된 마이크와 웹 캠을 통해 영상과 오디오 파일을 서버로 전송한다. 오디오 파일은 서버내 실시간 번역기를 통해 번역 자막을 생성해준다. 번역 자막과 영상을 다른 사용자에게 송출한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

2.3.4 결과물 목록 및 상세 사양

대분류	소분류	기능	형식	비고
파일	음성파일 열기	스크립트로 변환하고 싶은 음성 파일을 넣는다.	M4a, flac	
	텍스트파일 저장	변환된 스크립트를 다른 언어로 번역하고 저장한다.	txt	
입력	마이크 입력	마이크에서 음성입력을 받는다.	M4a, flac	
	영상 입력	캠에서 영상 입력을 받는다.	MP4	
출력	영상 출력	사용자로부터 받은 영상을 사용자에게 보여준다.	MP4	
	번역 텍스트출력	번역 서버로부터 전달받은 텍스트를 자막으로 보여준다.	txt	
서버	웹서버	사용자로부터 받은 영상 및 음성을 다른 사용자에게 전달한다.	Node.js	
	번역서버	<ul style="list-style-type: none"> - 사용자로부터 받은 음성을 텍스트로 변환한다. - 텍스트를 다른 사용자의 언어로 번역한다. - 번역된 텍스트를 다른 사용자에게 전달한다. 	Ubuntu instance	

2.4 기대효과


(1) 기대효과

이 프로젝트를 진행함으로써 기대하는 효과로는 크게 4 가지가 있다.

첫 번째로 3명 이상의 화상 통화 번역 서비스를 이용할 수 있다. 여러 명의 외국인 친구들과 과의 음성, 화상통화나, 다국적 기업의 회의 또는 입사 화상면접에 대해서 좋은 대안이 될 것이다.

두 번째로 생방송으로 진행되는 공식 발표 영상에서 이를 다국어로 번역하여 제공함으로써 많은 사람이 보다 쉽게 접근할 수 있는 장점이 있다.

세 번째로 기존의 스카이프 서비스의 설정단계의 복잡함을 넘어 사용자가 자신의 언어를 선택하면 그에 맞게 번역이 되는 서비스를 제공할 수 있게 되고 설정 단계를 어렵지 않게 하여 처음

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

사용하는 사용자들에게 편리함을 줄 수 있는 장점이 있다.

네 번째로 사용자에게 보여지는 번역 스크립트의 기록을 일정량을 저장하여 서로의 의사소통에 어려움을 주지 않고 번역된 스크립트가 보여지는 시간이 이전의 서비스보다 길게 설정하여 바로 의사소통을 하는데 어려움을 주지 않는 장점이 있다.

3 배경 기술

3.1 기술적 요구사항

(1) 개발 환경

서버 환경

- **Google Cloud platform:** 원하는 가상 서버를 구축, 웹페이지를 구동할 수 있게 웹서버와 번역서비스를 실행시킬 API용 서버를 생성한다.

개발 환경

- **front end** : React, webrtc
 - **React** : 웹 프론트 엔드를 개발하기 위해서 React를 사용한다.
 - **WebRtc** : 웹 브라우저 간에 화상통화를 구현하기 위해서 웹 브라우저 간에 플러그인의 도움 없이 서로 통신 할 수 있도록 설계된 API인 WebRtc를 사용할 예정이다.
- **Back-end** : Node.js, Express, Socket.io
 - **Node.js** : 서버는 Non-blocking I/O와 단일 스레드 이벤트 루프를 통한 높은 처리 성능을 가지는 Node.js를 사용한다
 - **Socket.io** : 서버와 클라이언트 간의 통신환경을 구현하기 위해서 Socket IO를 사용한다.

사용 기술: 여러 나라의 음성을 텍스트로 변환하는 기능과 해당 텍스트를 유저의 언어로 번역하는 기능은 Google 클라우드 API를 사용한다.

- **Google STT:** 통화 중 음성을 인식해 text 형식의 파일로 만들어주는 API
- **Google translate:** 읽어드린 text를 사용자 요청의 언어로 번역해주는 API

(2) 프로젝트 결과물을 확인할 수 있는 환경

사용자 환경

- 플랫폼 웹페이지에 접속 가능한 랩 탑: 웹 캠과 내장 마이크가 들어있는 랩 탑
- 플랫폼 웹페이지에 접속 가능한 안드로이드 휴대폰: 셀프 카메라와 마이크 내장된 안드로이드 폰

서버 환경

- 서버가 구축된 데스크톱: 결과물을 확인할 수 있는 소프트웨어가 설치되었는지 구현상에 문제가 없는지를 서버 자체에서 확인한다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

3.2 현실적 제한 요소 및 그 해결 방안

3.2.1 하드웨어

사용자 간 영상과 오디오를 서로 주고받는데 마이크와 웹 캠이 따로 없는 데스크탑의 경우 영상 송출과 음성 인식을 위한 장비가 필요하다는 단점이 존재합니다. 이를 보완하기 위해 데스크탑은 이어폰을 통해 통화할 수 있게끔 영상 통화뿐만 아니라 음성통화도 지원하려고 합니다.

3.2.2 소프트웨어

사용자간 번역서비스의 속도 저하를 막기 위해 각자에게 translator 모듈이 부여됩니다. 하지만 모듈의 비용은 인식이 필요 없는 무음의 상태에서도 빠져나가므로 회의나 통화 시 발생하는 침묵에 대해서도 비용이 발생합니다. 이를 보완하기 위해 장시간 침묵 시 통화 퇴장 기능을 두어 translator 모듈의 발생 비용을 줄이려고 합니다.

3.2.3 기타

사용자의 음성을 번역하는 서비스로서 주변소음에 민감하다는 단점이 존재합니다. 이런 단점을 보완하기 위해서는 조용한 환경에서 통화를 하거나, 이어폰을 사용해 하울링을 줄이는 방법을 써야합니다.

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

4 프로젝트 팀 구성 및 역할 분담

이름	역할
박지훈	- 웹 개발 및 네트워크 구축
최창규	- 화상 통화 서버 개발
공병민	- Speech to text 환경 구현
고준규	- Translate 환경 구현
박귀환	- 번역 서버 구축 및 최적화
오석현	- 웹 서버 개발 및 유저 인터페이스 구현

5 프로젝트 비용

항목	예산치 (MD)
아이디어 구상	3
자료 조사 및 개발환경계획	5
웹 프론트엔드 UI 개발	20
웹서버 환경 구현	40
Speech to text 기능 구현	30
Translate 기능 구현	20
음성 쿼리 스케줄링 시스템 구현	40
서비스 테스트 및 보완	10
보고서 작성 및 발표 준비	15
합	183

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

6 개발 일정 및 자원 관리

6.1 개발 일정

항목	세부내용	2 월	3 월	4 월	5 월	비고
요구사항분석	요구 분석					
	SRS 작성					
관련분야연구	주요 기술 연구					
	관련 시스템 분석					
설계	시스템 설계					
구현	코딩 및 모듈 테스트					
테스트	시스템 테스트					

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

6.2 일정별 주요 산출물

마일스톤	개요	시작일	종료일
계획서 발표	개발 환경 구축 개발 시 필요한 기술 정보 수집 산출물 : 1. 프로젝트 수행 계획서 2. 프로젝트 발표 ppt	2019-03-10	2019-03-15
설계 완료	시스템 설계 완료 산출물 : 1. 시스템 설계 사양서 2. 개발 환경	2019-03-16	2019-03-29
1 차 중간 보고	웹서버 및 번역 서버 구축 산출물 : 1. 프로젝트 1 차 중간 보고서 2. 프로젝트 진도 점검표 3. 1 차분 구현 소스 코드	2019-03-16	2019-04-19
구현 완료	음성 쿼리 스케줄링 시스템 구현 산출물: 1. 각 기능 소스 코드 2. 서버 및 웹 페이지	2019-04-20	2019-05-31
테스트	시스템 통합 테스트 산출물: 데모 영상 또는 시현	2019-05-01	2019-05-30
최종 보고서	최종 보고 산출물: 최종 보고서	2019-05-31	

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

6.3 인력자원 투입계획

이름	개발항목	시작일	종료일	총개발일(MD)
박지훈	웹 개발 및 네트워크 구축	2019-03-04	2019-05-30	20
최창규	화상 통화 서버 개발	2019-03-04	2019-05-30	20
공병민	Speech to text 환경 구현	2019-03-04	2019-05-30	20
고준규	Translate 환경 구현	2019-03-04	2019-05-30	20
박귀환	번역 서버 구축 및 최적화	2019-03-04	2019-05-30	20
오석현	웹 서버 개발 및 유저 인터페이스 구현	2019-03-04	2019-05-30	20

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

6.4 비 인적자원 투입계획

항목	Provider	시작일	종료일	Required Options
개인 노트북 4 대	Apple,Lg,Lenovo	2009-03-04	2009-05-30	
Google cloud platform API	Google	2009-03-04	2009-05-30	

 국민대학교 컴퓨터공학부 캡스톤 디자인 I	계획서		
	프로젝트 명	실시간 화상채팅 번역	
	팀 명	LTE (Language Translation Easily)	
	Confidential Restricted	Version 1.0	2019-MAR-14

7 참고 문헌

번호	종류	제목	출처	발행년도	저자	기타
	웹	Cloud Speech-to-Text 문서	https://cloud.google.com/speech-to-text/docs/			
	웹	Cloud Translate 문서	https://cloud.google.com/translate/docs/			
	웹	Skype translator 소개	http://skype.dae-sung.com/service_intro/serviceTranslator01.asp			