**캡스톤 디자인 I**

**종합설계 프로젝트**

|  |  |
| --- | --- |
| 프로젝트 명 | *LTE* |
| 팀 명 | *LTE* |
| 문서 제목 | 결과보고서 |

|  |  |
| --- | --- |
| **Version** | 0 |
| **Date** | 2019-MAY-28 |

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

|  |
| --- |
| **CONFIDENTIALITY/SECURITY WARNING**  이 문서에 포함되어 있는 정보는 국민대학교 전자정보통신대학 컴퓨터공학부 및 컴퓨터공학부 개설 교과목 캡스톤 디자인I 수강 학생 중 프로젝트 “LTE”를 수행하는 팀 “LTE”의 팀원들의 자산입니다. 국민대학교 컴퓨터공학부 및 팀 “LTE”의 팀원들의 서면 허락없이 사용되거나, 재가공 될 수 없습니다. |

**문서 정보 / 수정 내역**

|  |  |
| --- | --- |
| **Filename** | 결과보고서-LTE.doc |
| **원안작성자** | 박지훈, 공병민, 최창규, 고준규, 박귀환, 오석현 |
| **수정작업자** | 박지훈, 최창규 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 수정날짜 | 대표수정자 | Revision | 추가/수정 항목 | 내 용 |
| 2019-05-26 | 박지훈 | 1.0 | 최초 작성 |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**목 차**

[1 개요 4](#_Toc9764442)

[1.1 프로젝트 개요 4](#_Toc9764443)

[1.2 추진 배경 및 필요성 4](#_Toc9764444)

[1.2.1 실시간 번역 시스템의 추진 배경 4](#_Toc9764445)

[1.2.2 기술 시장 현황 및 현재 개발된 시스템 현황 4](#_Toc9764446)

[1.2.3 현재 개발된 시스템의 문제점 4](#_Toc9764447)

[2 개발 내용 및 결과물 6](#_Toc9764448)

[2.1 목표 6](#_Toc9764449)

[2.2 연구/개발 내용 및 결과물 6](#_Toc9764450)

[2.2.1 연구/개발 내용 6](#_Toc9764451)

[2.2.2 시스템 기능 요구사항 8](#_Toc9764452)

[2.2.3 시스템 비기능(품질) 요구사항 9](#_Toc9764453)

[2.2.4 시스템 구조 및 설계도 9](#_Toc9764454)

[2.2.5 활용/개발된 기술 10](#_Toc9764455)

[2.2.6 현실적 제한 요소 및 그 해결 방안 10](#_Toc9764456)

[2.2.7 결과물 목록 10](#_Toc9764457)

[2.3 기대효과 및 활용방안 10](#_Toc9764458)

[3 자기평가 11](#_Toc9764459)

[4 참고 문헌 12](#_Toc9764460)

[5 부록 13](#_Toc9764461)

[5.1 사용자 매뉴얼 13](#_Toc9764462)

[5.2 테스트 케이스 16](#_Toc9764463)

# 개요

## 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.

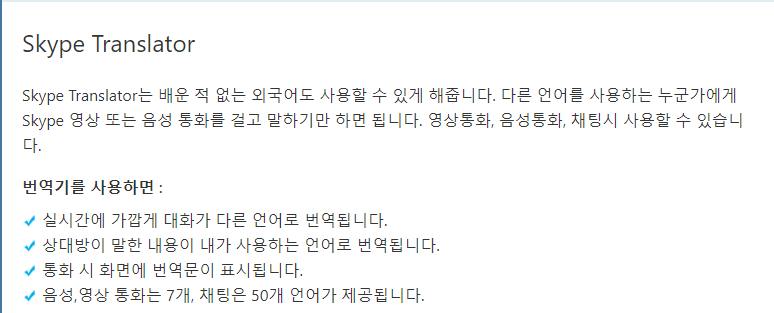
## 추진 배경 및 필요성

### 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

### 기술 시장 현황 및 현재 개발된 시스템 현황

* + 대표적인 실시간 번역 음성, 화상 채팅 서비스 중에는 스카이프가 있다. 현재까지 개발된 스카이프의 영상통화의 경우 1:1 영상통화, 그룹 영상통화로 나뉘어 있다. 또한 배우 지 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.



### 현재 개발된 시스템의 문제점

* + 1.2.2에서 소개한 실시간 번역 화상 채팅 서비스 스카이프에는 아래와 같은 문제점이 있다
  + 첫 번째로 사용자가 설정한 언어로 번역된 스크립트가 사용자에게 보여지는 시간이 충분하지 않아 사용자가 설정한 언어로 번역된 스크립트를 다 보고 읽기 전에 스크립트가 사라져 서로 의사사통에 차질을 갖게 되는 단점이 있다.
  + 두번째로 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 번역지원이 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.

# 개발 내용 및 결과물

## 목표

본 프로젝트는 다국적 간의 영상동화에서 언어의 차이로 인한 의사소통의 불편함 해소를 목적으로 한다. 사용자가 말한 음성을 텍스트로 변환한다. 변환된 텍스트는 각각의 사용자들이 설정한 언어에 맞게 번역되어 화면에 표시해줌으로서 서로 다른 언어에 대한 의사소통을 원활하게 한다.

* 3명 이상의 화상통화서비스를 제공한다.
* 여러가지 언어에 대해 제공한다.
* 기존의 서비스와 다르게 간편한 채팅 방 생성, 로그인 간소화 등 직관적이고 쉽운 서비스를 제공한다.
* 채팅 기록들을 DB에 저장하여 나중에 들어온 사용자가 이전 대화를 확인 할 수 있고 이를 다운로드하는 기능을 제공한다.

## 연구/개발 내용 및 결과물

### 연구/개발 내용

**채팅 방 생성 단계**

1. 사용자들은 웹 사이트에 접속을 하여 자신의 닉네임과 사용 언어를 지정하여 로그인한다.
2. 채팅방의 제목, 허용 인원, 비밀번호를 설정하여 채팅방을 생성한다.

**음성 번역 단계**

1. 웹 브라우저 기반 STT엔진을 활용하여 클라이언트 상에서 음성을 텍스트로 변환한다.
2. 변환된 텍스트는 서버로 보내서 서버에 있는 Google Translate API를 이용하여 번역을 수행한다.
3. 번역된 텍스트는 같은 방에 접속해 있는 모든 클라이언트들에게 전송된다.
4. Mongo DB를 이용하여 웹페이지에 보여지는 채팅방의 Script를 저장하여 후에 사용자가 늦게 채팅방에 들어오거나 채팅방을 나갔다가 들어와도 이전의 내용을 저장하여 사용자에게 제공한다.

**영상 통화 단계**

1. 사용자가 방에 참가할 시 다른 모든 사용자에게 broadcast를 하여 알린다.

2. 사용자들은 새로 들어온 참가자에 대한 WebRTC PeerConection을 생성한다.

3. 각 참가자들은 새로 들어온 참가자에게 offer를 보내면 해당 PeerConnection에 응하 여 서로 영상을 공유 가능한 서버가 연결된다.

**상세 개발 내용**

서버에서 구현 과정은 실시간 음성 채팅 서비스와, 화상 서비스 두 부분으로 나누었다. 실시간 음성 채팅 서비스는 사용자들에게 편의성을 제공하기 위하여 음성 뿐아니라, 기존의 입출력 장치를 이용한 실시간 채팅도 가능하도록 하였다. 음성을 통한 채팅과 입출력 장치를 통한 서비스를 같이 구현하는 것은 다소 복잡할 것으로 판단되어, 먼저 입출력 장치를 통한 서비스를 먼저 개발한 뒤에 음성 서비스를 이에 맞게 부착하는 형식으로 개발하였다.

간단히 구현 내용을 소개하자면 각 사용자들은 채팅을 입력하고 입력된 채팅을 ajax요청으로 서버로 보낸다. 서버에서는 현재 그 사용자가 속한 방에 있는 모든 언어를 번역한 뒤에 그것을 저장하고, 이 데이터를 다시 클라이언트로 보내어 클라이언트는 자신이 필요한 언어로 된 데이터만 추출하여 화면에 내보내도록 하였다. 서버에서 현재 방에 없는 사용자의 언어까지 모조리 번역하면 처리 시간이 느려 지기 때문에 사용자가 입장하면 그 사용자의 언어가 방에 추가되고, 퇴장하면 그 사용자의 언어도 삭제하도록 하여 지연시간을 감소시켰다.

각 채팅은 모두 데이터베이스를 통하여 저장되도록 하였는데, 이는 사용자들이 언제든지 현재까지의 채팅을 다운로드 할 수 있도록 하기 위함과, 새로 입장하는 사용자들이 현재까지 방에 기록된 채팅을 자신의 언어에 맞게 볼 수 있도록 하기 위함이다. 물론 방에 추가되지 않은, 새로운 언어를 이용하는 사용자가 참가하면 기록되었던 모든 채팅을 다시 번역하게 되어 입장할 때에 시간이 지연되는 단점이 있다. 이러한 단점은 각 방에 필요 없는 언어까지 모두 번역하여, 데이터 베이스에 저장하면 없앨 수 있지만, 실시간 채팅에서 지연시간이 상당히 길어 지기 때문에 이렇게 남겨 두었다.

현재까지의 구현을 바탕으로 한 음성인식 서비스는 html5 음성인식 기능을 이용하였는데, 원래는 구글 STT api를 이용하려고 하였으나, html 즉, 브라우저에 할당하여 각 사용자의 컴퓨터에 접근하여 마이크 권한을 가져올 수 있는 방법이 없어서 이를 선택하였다. 이것은 html 자체에서 브라우저의 오디오 스트림에 접근하기 때문에 따로 플러그인을 설치할 필요가 없고, 쉽게 오디오 스트림에 접근할 수 있으며, 무엇보다 음성인식 성능이 좋다. 자체적으로 구글 STT api와 html5 음성인식 서비스를 사용해본 결과, 성능 차이가 없었다. 또한 음성 인식 결과를 더 빠르게 산출하기 위해, 말이 끝날 때까지 기다리지 않고, 음성인식 결과가 산출되면, 바로 서버로 데이터를 보내어 처리되는 방식을 선택하였는데, http환경에서는 보안상의 이유로 결과물이 산출될 때마다 다시 마이크 접근을 요구하여 제대로 음성이 인식되지 않았다. 따라서 https 환경으로 서버를 재구성하였다.

화상 서비스는 WebRTC를 이용하였는데, 이도 마찬가지로 html5에서 제공한다. webrtc docs에서는 미디어 스트림과 오디오 스트림을 얻는 방법, 1:1 환경에서의 통신 방법 등을 간략하게 설명하여 이를 바탕으로 다대다 환경에서 이용 가능하고, 채팅 방 사이에서는 독립적이고, 사용자 세션을 바탕으로 통신을 초기화 하기 때문에 서비스 세션 관리 등을 새롭게 구현하였다. 간략하게, WebRTC 클라이언트들은 해상도와 코덱 기능 같은 미디어 정보들을 서로 교환하고 확인해야 한다. 즉, 통신할 수 있는 환경이 필요한데 이는 Signaling을 이용하여 구현하였다.

사용자들은 채팅방에 입장하면, 자신들만의 고유한 번호를 가지게 되고, 이를 바탕으로 고유의 channel을 가지게 된다. 이후, 각 사용자들의 ip주소, 포트 등의 네트워크 정보들을 가져와서 이 정보들을 교환하는데, 이를 위해서 ICE server가 필요하다. 가장 대기시간이 적은 UDP를 통해 피어들 끼리 직접 연결이 가능한지 시도한다. 이는 대표적으로 STUN 서버를 이용하여 이루어 지는데, STUN 서버는 NAT 뒤에 있는 피어 들이 연결 가능하도록 자신들의 공용(public)주소와 포트를 찾아준다. 만약 UDP가 실패하면, ICE는 HTTP상의 TCP로 시도를 하고 그 다음엔 HTTPS상에서 시도한다. 만약 NAT와 방화벽으로 인해 직접 연결이 실패하면 ICE는 중계를 위해 TURN 서버를 사용한다. 중계 서버가 없으면 직접 연결을 실패한 경우에 화상 서비스가 이루어지지 않아 TURN서버가 필수적으로 필요하였다. 하지만, TURN 서버는 무료로 제공하는 기업들이 별로 없고, 제공한다 해도 다소 불안정한 서버였지만, 개중에 가장 좋은 무료 서버를 이용하였고, 나중에 기회가 된다면 유로 서버를 이용하여 사용하고자 한다. 이로써 서로 통신이 되면 여러가지 정보들을 공유하여 다중 화상 서비스가 가능해지도록 구현하였다. 각 channel을 참가한 방에 연결시켜 방마다 독립적인 통신이 가능하도록 하였고, 다소 부하가 큰 통신 환경의 부하를 줄여서 많은 사용자가 이용하여도 화상 연결이나 채팅 서비스가 느려 지지 않도록 구현하였다.

### 시스템 기능 요구사항

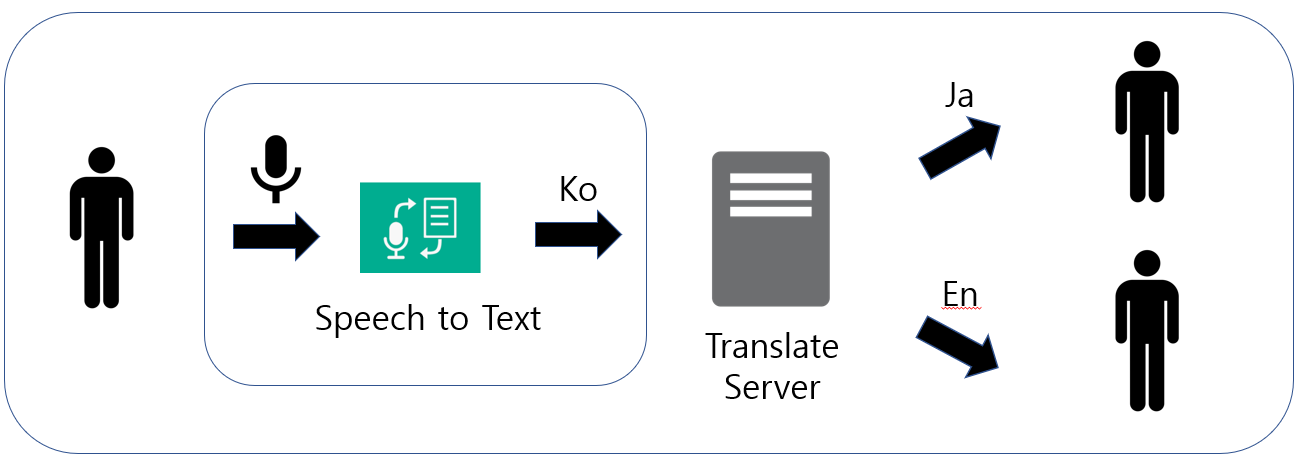
|  |  |  |
| --- | --- | --- |
| **번호** | **내용** | **완료 여부** |
| **1** | **사용자는 웹페이지에 접근하여 본인이 사용하는 언어를 선택한다.** | **완료** |
| **2.** | **사용자는 화상통화가 가능한 방을 생성한다.** | **완료** |
| **3.** | **사용자는 이미 생성되어 있는 방에 입장한다.** | **완료** |
| **4.** | **하나의 방에 입장한 사람이 2인 이상인 경우 화상통화가 시작된다.** | **완료** |
| **5.** | **사용자의 음성이 텍스트로 변환되어 다른 사용자에게 번역되어 전달된다.** | **변경사항**  **완료** |

### 시스템 비 기능(품질) 요구사항

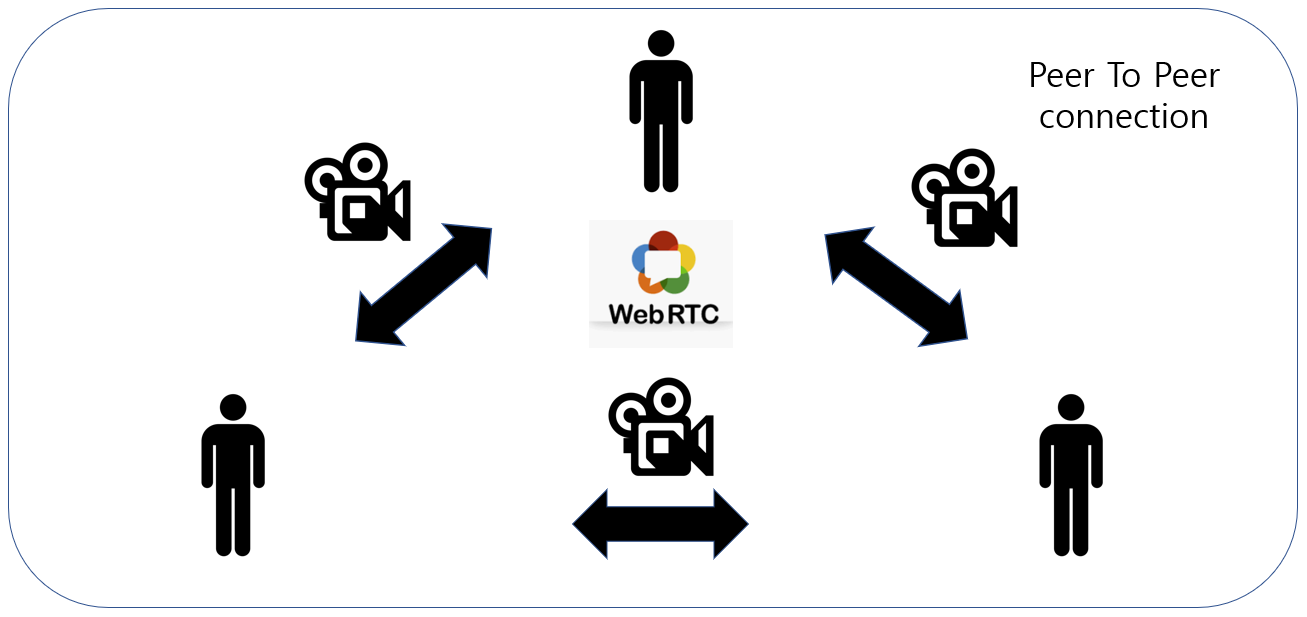
|  |  |  |
| --- | --- | --- |
| **번호** | **내용** | **완료 여부** |
| **1** | **사용자가 말한 내용이 텍스트로 변환되어 2초 이내에 번역결과가 채팅창에 게시된다.(성능)** | **달성(1초)** |
| **2.** | **사용자는 별도의 회원가입 절차 없이 쉽게 서비스를 이용할 수 있다.** | **달성** |
| **3.** | **영상통화 시 0.5초 이하의 지연율을 보장한다.** | **달성** |

### 시스템 구조 및 설계도

**Translate Server :**

****

**화상 통화 Peer To Peer 연결 구조 :**

****

### 활용/개발된 기술

**STT : 사용자의 음성을 텍스트로 변환하기 위해 웹 브라우저 기반 STT open API를 이용하여 라이센스의 제한 없이 클라이언트에서 구현**

**Google Translate API : 다른 사용자의 언어로 번역하기 위해 사용**

**WebRTC : 영상통화를 구현하기위해 클라이언트간 Peer to Peer 연결망을 사용하여 서버의**  **부하를 감소**

**MongoDB : 화상통화동안 생성된 대화 텍스트들을 데이터베이스에 저장하여 나중에 들어온 사용자가 이전 내용을 볼 수 있게 구현**

**PM2 : 웹 서비스간 예기치 않은 문제로 서버가 종료되는 문제에 대응하고 서버의 업무를 분산하기 위해 여러 프로세스를 관리할 수 있는 pm2 활용**

### 현실적 제한 요소 및 그 해결 방안

**여러 명이 대화에 참여하여 번역 업무가 많아져서 지연율이 높아지는 문제**

* **STT의 경우 비용이 크기 때문에 서버에 전달하기 전에 클라이언트에서 처리하고 생성된 텍스트를 서버에 전송해서 번역하는 방식으로 서버의 부담을 줄이고 PM2를 이용하여 여러 프로세스로 업무를 분담했다.**

### 결과물 목록

**실시간 화상통화 번역 서비스 웹 사이트**

## 

## 기대효과 및 활용방안

* 3명 이상의 화상 통화 번역 서비스를 이용할 수 있다. 이를 통해서 여러 명의 외국인 친구들 과의 화상통화, 다국적 기업의 화상회의, 화상 면접 등에 대해서 좋은 대안이 될 것이다.
* 생방송으로 진행되는 공식 발표 영상에서 다국어 자막을 제공함으로써 많은 사람이 보다 쉽게 접근할 수 있는 장점이 있다.
* 스카이프 서비스의 단점이었던 설정단계의 복잡함을 개선하여 회원가입이 필요 없는 간편한 로그인 기능을 제공하여 처음 사용하는 사용자들에게 편리함을 줄 수 있다.
* 사용자에게 대화 내용을 채팅로그 형식으로 제공하여 이전에 했던 대화내용을 볼 수 있고 이를 다운로드하여 회의록처럼 기록할 수 있다.

# 자기평가

|  |  |
| --- | --- |
| 최종결과물 | 실시간 화상통화 번역 서비스 |
| 주요 평가 기준 | 1.다른 언어를 쓰는 사용자의 번역 스크립트가 잘 보이는가? |
| 자기 평가 | 번역 자막을 표시하는 형식을 화상통화 밑 부분의 자막을 통해서가 아닌 채팅로그를 기록하는 채팅 방 형식으로 바꾸었고, 이를 통해 방에 입장한 모든 사용자들은 어떤 대화를 나누고 있었는지 파악하기 용이하다.  또한 추가적인 개발 사항으로 뒤늦게 들어온 사용자의 경우 다른 사용자들의 이전 대화 기록을 볼 수 있게 구현하여 회의 진행 사항을 쉽게 알 수 있게 구현하였다. |

|  |  |
| --- | --- |
| 최종결과물 | 실시간 화상통화 번역 서비스 |
| 주요 평가 기준 | 2.화상 통화 서비스의 경우 목표한 3명의 통화를 지원하는가? |
| 자기 평가 | 3명보다 많은 인원을 통화 가능하게 구현하였다. 테스트 결과 7명의 인원으로 통화가능 한 것을 확인하였다.  또한 삼성 핸드폰의 경우 자체적인 브라우저를 통해 접속할 경우 역시 화상통화를 지원하는 것을 확인하였다. |

|  |  |
| --- | --- |
| 최종결과물 | 실시간 화상통화 번역 서비스 |
| 주요 평가 기준 | 1. 설정하는 단계를 간소화하였는가? |
| 자기 평가 | 기존의 서비스(skype)와 달리 사용자의 신상정보를 요구하지 않고 간단한 별칭과 사용언어를 지정하여 방에 입장 가능하다.  별 다른 설정 없이 마이크와 웹 캠을 허가할 경우 바로 화상통화 번역 서비스를 이용가능 하게끔 UI를 구현하였다. |

|  |  |
| --- | --- |
| 최종결과물 | 실시간 화상통화 번역 서비스 |
| 주요 평가 기준 | 1.목표와 같이 한국어 번역을 지원하는가? |
| 자기 평가 | 인식된 언어를 Google Translate API를 통해 한국어 번역을 지원하게 구현하였다. |

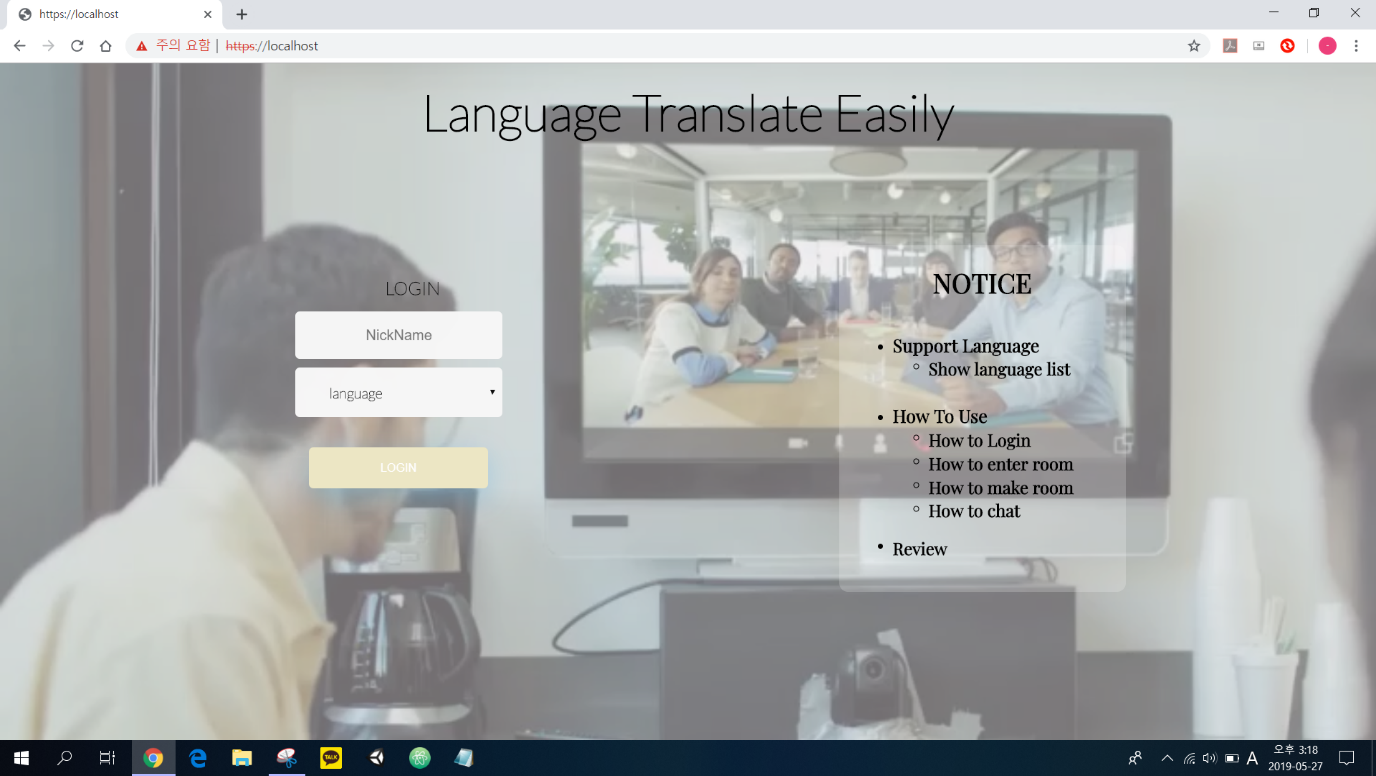
# 참고 문헌

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 번호 | 종류 | 제목 | 출처 | 발행년도 | 저자 | 기타 |
| 1 | 웹페이지 | https://colorlib.com/wp/css3-table-templates/ | colorlib | 2019 |  |  |
| 2 | 웹페이지 | <https://www.html5rocks.com/ko/tutorials/webrtc/basics/> | Html5rocks | 2012 |  |  |
| 3 | 웹페이지 | https://github.com/lucaslouca/video-conference-webrtc | Github | 2015 |  |  |
| 4 | 웹페이지 | https://jeongchul.tistory.com/539 | Tistory | 2017 |  |  |
| 5 | 웹페이지 | https://cloud.google.com/translate/?hl=ko-KRhttp | Google Cloud | 2019 |  |  |

# 부록

## 사용자 매뉴얼

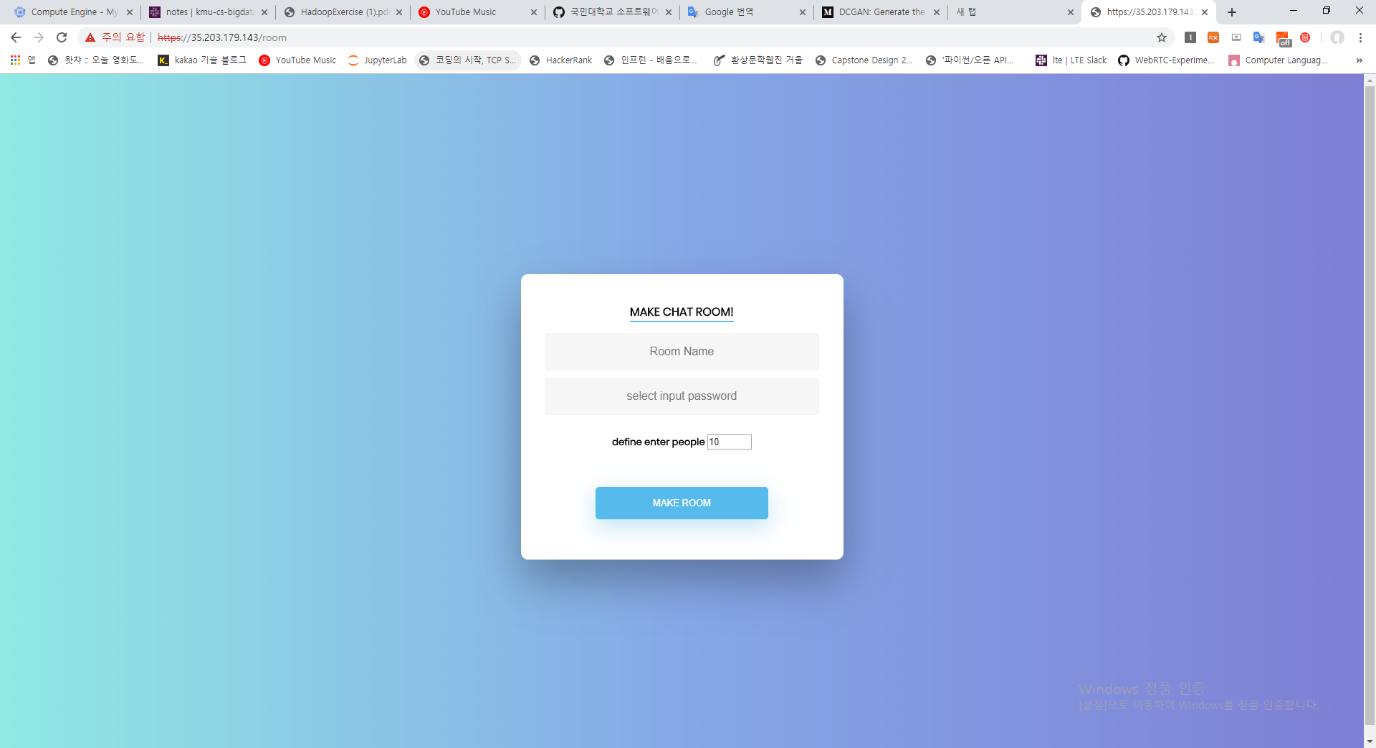
**1.초기화면**

****

웹에 접속하면 나오는 초기 화면입니다.

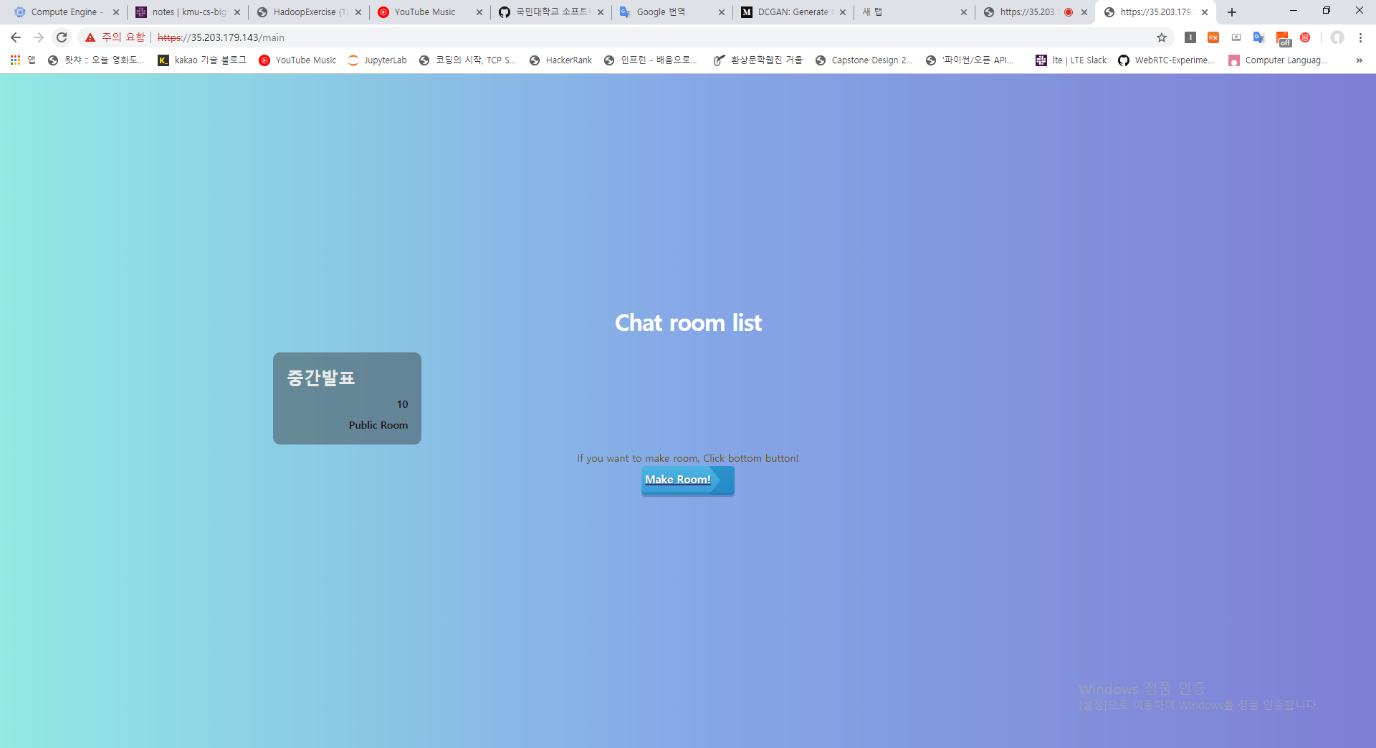
사용자는 이 화면에서 방에서 사용할 별칭(Nickname)과 사용하는 언어를 선택한 뒤 입장하게 됩니다.

**2. 방 생성 화면**

****

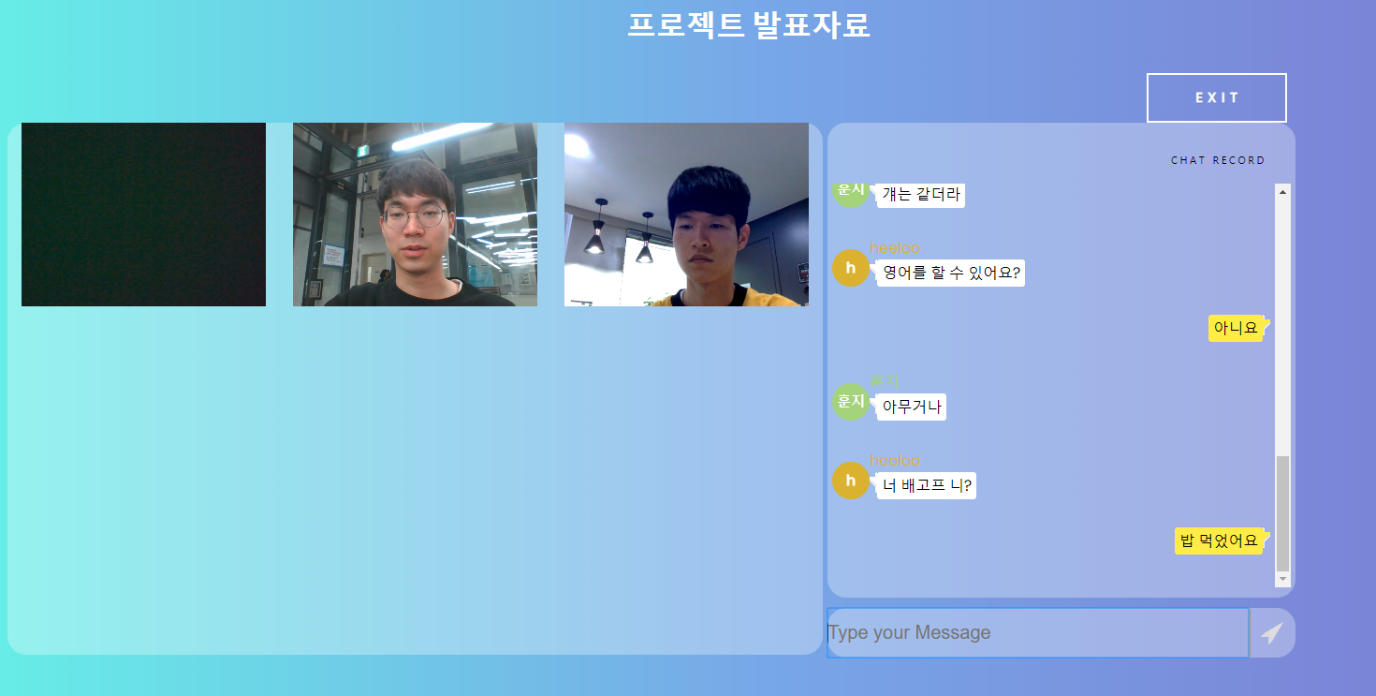
방 생성 화면입니다. 생성하려는 방의 제목과 인원 수 제한을 정할 수 있고 비밀번호 설정을 통해 이미 허가된 사람들만 입장할 수 있습니다.

**3..방 입장 화면**

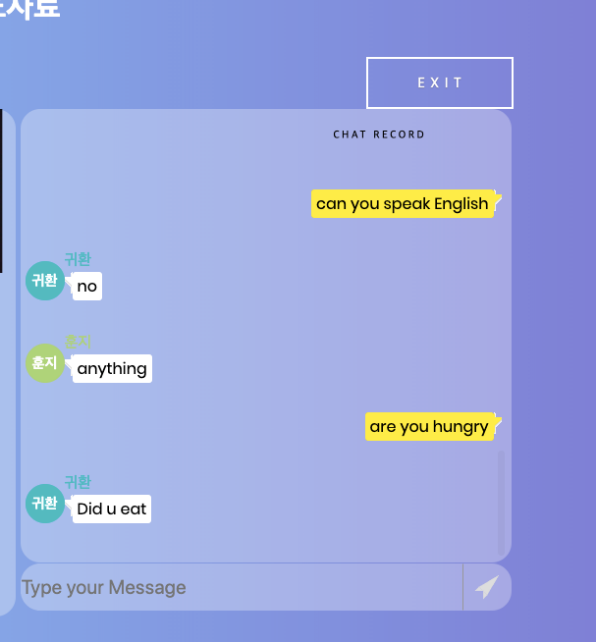
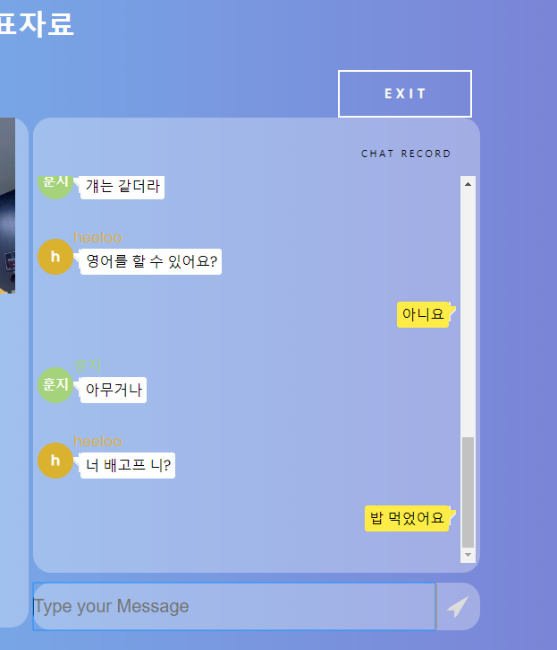
****

방 목록 화면에 이미 생성된 방의 제목이 표시됩니다. ENTER 버튼을 눌러 입장 가능합니다.

**4..채팅방 화면**

****

채팅 방 화면입니다. 화면 좌측 각자의 얼굴이 보이는 화상 통화 부분과 우측의 번역된 채팅 로그가 표시되는 채팅 부분으로 나누어져 있습니다. 각 사용자가 초기화면에서 설정한 언어로 번역되어 채팅 부분에 표시됩니다.

같은 내용이라도 사용자가 선택한 언어로만 표시되는 예시

## 테스트 케이스

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 대분류 | 소분류 | 기능 | 테스트 방법 | 기대 결과 | 테스트  결과 |
| 기능 | 다대다  화상통화 | 3명 이상의 화상통화를 지원한다. | 방에 입장 가능한 인원의 수를 센다.  또한 입장 시 화상통화가 원활히 진행되는지 눈으로 확인한다. | 3명이상 화상통화를 통해 서로 소통한다. | 성공 |
| 음성인식 | 사용자의 음성을 인식해 텍스트로 변환한다. | 마이크에 말을 해 음성인식이 잘 되는지 웹 콘솔에서 결과를 확인한다. | 인식된 결과가 웹 콘솔에 중간 결과로 표시된다. | 성공 |
| 유저의 사용언어번역 | 음성 텍스트를 해당 사용자의 언어로 변환한다. | 번역된 결과가 채팅 방 화면에 잘 표시되는 지 확인한다. | 번역된 결과가 채팅방에 잘 표시되어 언어가 다른 사용자끼리 소통 가능하다. | 성공 |