

종합설계 최종 발표

Lafuta조

201502039 김종운

201702073 정지수

CONTENTS

- 주제

머신러닝을 활용한 논문 번역기 어플리케이션

- 계기

해외 학회에 논문을 제출하려 하는 사람들이
기존의 번역기를 사용하여 작성할 때에
부적합한 단어 출력 등에 의한 불편함을 겪고 있는 것을 보고
이러한 학술 문서에 특화된 번역기를 고안하게 되었다.

- 제작 형태

웹페이지 형태로 제공

CONTENTS

디자인스프린트

- 문제정의, HMW
- 라이트닝 데모
- Crazy 8s
- 스토리보드
- 프로토타입 제작
- 피드백

SE 문서

- 요구사항명세서
- 유즈케이스 명세서
- 클래스 다이어그램
- 시퀀스 다이어그램

멘토링

- 정상근 교수님
- 아이와즈 사

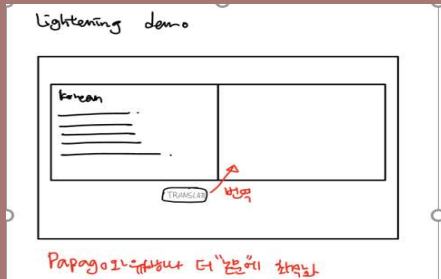
설문

- 설문
- 요청 수정사항
- 피드백

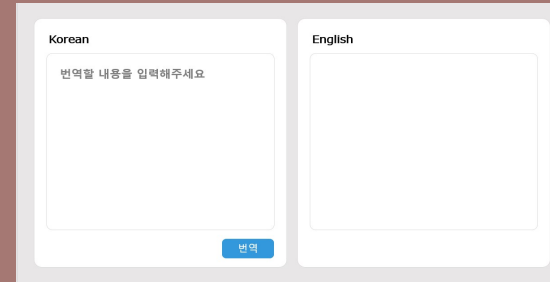
프로토타입 데모

- 데모 웹페이지

디자인 스프린트



간단한 사용자 인터페이스 데모



스토리보드를 참고하여 프론트 프로토타입 제작

문제 정의
HMW

라이트닝
데모

Craze 8s,
스토리보드
제작

프로토타입
제작

피드백

문제 정의 및 기능에 대한 map, HMW

Crazy 8s를 추합하여 스토리보드 제작

설문조사를 통해 기능에 대한 피드백

문제

- 한국어로 작성한 문장을 영문으로 옮기려고 할 때 영어를 잘 하지 못하여도 자연스러운 문장들로 번역하는 데에 어려움
- 영어를 잘 못하는 사람이 번역을 직접 할 경우 많은 시간 투자가 필요
- 일반적으로 사용하는 번역기 사용시 눈에 쓰이는 문제, 단어와 거리가 생길 수 있음

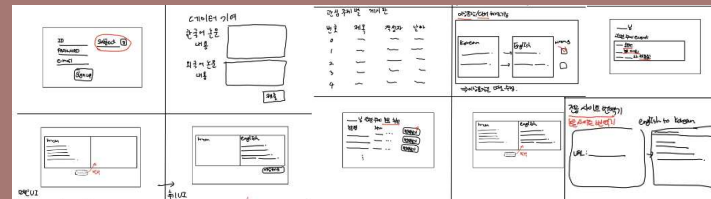
영어로 직접 map

원래는 문장일 → 영어로 번역 → 문장, 단어 검색 → 재작

HMW 병합

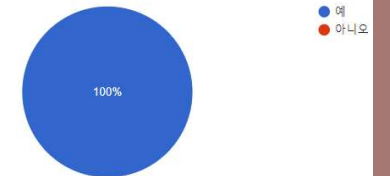
- 어떻게 하면 영어로 작성 시 적절한 단어를 선택할 수 있을까?
- 어떻게 하면 문맥을 더 자연스럽게 작성할 수 있을까?
- 어떻게 하면 더 빠르게 작성할 수 있을까?
- 어떻게 하면 사용자가 간편하게 서비스를 이용할 수 있을까?

이렇게 하면 영어로 작성 시 적절한 단어를 선택할 수 있을까?
어떻게 하면 문맥을 더 자연스럽게 작성할 수 있을까?
어떻게 하면 더 빠르게 작성할 수 있을까?
어떻게 하면 사용자가 간편하게 서비스를 이용할 수 있을까?

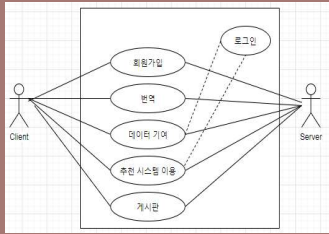


2. 위와 같은 프로토타입이 제작된다면 사용할 의향이 있나요?

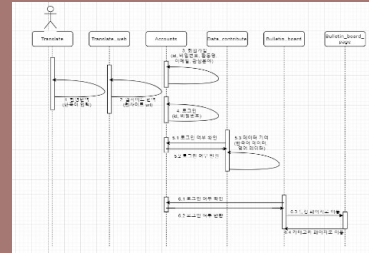
응답 5개



SE 문서



유즈케이스 다이어그램을 정의
유즈케이스 명세



유즈케이스에 대한 흐름을
시퀀스 다이어그램으로 명세

요구사항 명세서

유즈케이스
명세서

클래스
다이어그램

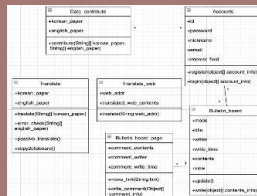
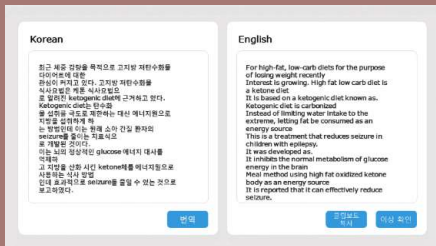
시퀀스
다이어그램

프로토타입
제작

사용자 인터페이스 정의 및 기능 요구사항,
성능 요구, 품질 속성

클래스 다이어그램을 설계
유즈케이스 시스템 모니터링
해당 클래스의 입출력, 기능 명세

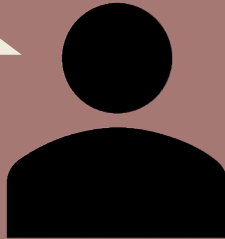
실제 프로토타입 제작



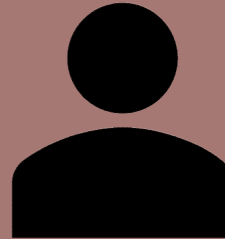
멘토링

1. 초록 번역부터 시작하여 정형적 부분부터
2. 정보과학회지에 초록이 영문과 한국어로 같이 올라오는 경우가 많음. (데이터 수집)
3. 현존하는 번역기를 이용하여 인터페이스를 먼저 생성
4. 하나의 input에 여러 개를 보여주는 형태

+ 번역된 정형적 문장이 포함된 논문을 보여주고, highlighting
+ 내가 썼던 한국어 단어들을 번역한 영어 단어를 리스팅



정상근 교수님

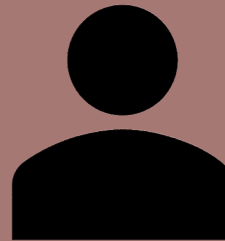
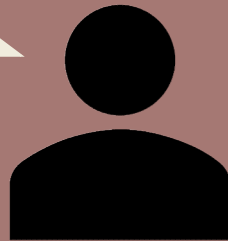


아이와즈 사

1. 머신 러닝으로 번역 엔진을 만드는 것은 난이도가 너무 어려움.
-> 관련 API들을 사용하여 애플리케이션을 제작을 추천
2. 수작업으로 데이터를 모으면 모을 수 있는 데이터양의 기대치가 많지 않을 것인데 그 정도 양으로는 좋은 결과를 보여주기가 힘들 것
-> 크롤링 오픈소스가 많으니 이를 이용하여 데이터를 수집할 것.
3. 일정 테이블을 만들어서 프로젝트를 관리

멘토링

현존 번역기 API



데이터 크롤링

BeautifulSoup



설문

요청된 수정 사항

- Input을 줄글만 받지 않고, 파일 형태로 받아 서식 유지가 되도록
- 구글 스칼라를 이용하는 현재의 모습 대신, 자주 쓰는 검색 엔진 선택
- 현재 쓰는 방식은 한 눈에 잘 들어오지 않으니, 주로 쓰는 번역기 결과에 다른 부수적 번역 결과를 선택하는 형태로
- 단어/문법이 틀리더라도 자체적으로 수정이 가능하도록

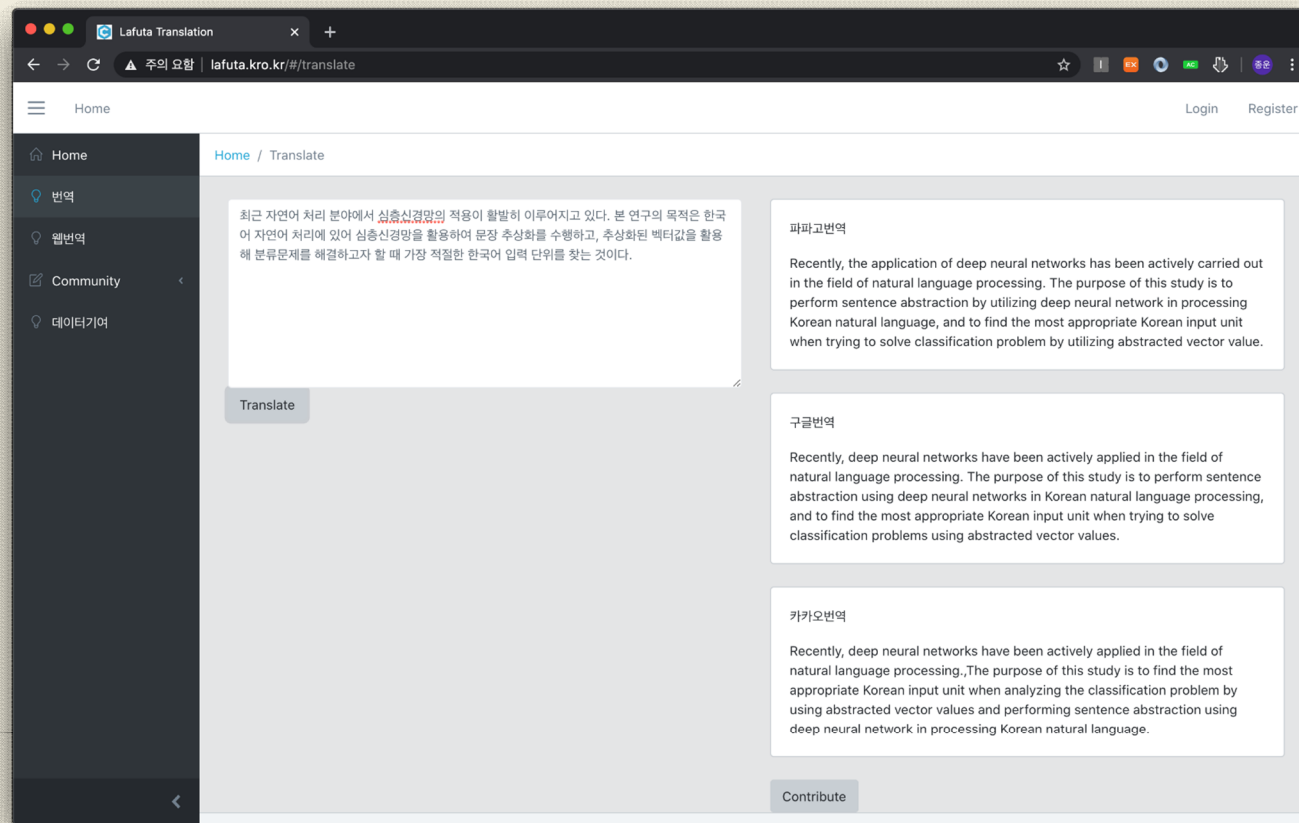


피드백

- 기능을 확장하여, Input을 파일 형태로 받아서 서식을 보존하는 형태 고려
- 검색엔진 선택 옵션도 고려
- 주로 쓰는 번역기 결과에 다른 부수적 번역 결과는 선택하는 형태는 애초에 다양한 출력 중 타당한 것을 선택할 수 있도록 하는 것과 조금 다르므로, 고려하되 기능으로서 추가는 신중
- 단어/문법 수정은 원 기능 구현 이후, 부수적 기능으로 고려하고자 함

총평: 건설적인 의견이라 대체로 반영하려 한다.
다만, 원래 구현하고자 한 기능과 조금 거리가 있는 기능의 경우
당장 반영하기보단, 부수적 기능으로서 참고하고자 한다.

프로토타입 데모



<http://lafuta.kro.kr/>

감사합니다