

문제정의서 [연구개발계획서]

프로젝트명	과학과 기술 문서 이해를 돕는 웹 브라우저 확장 프로그램 개발
-------	------------------------------------

조	02 조
지도교수	정상근 교수님 (서명)
조원	201803851 양성욱 201802074 김재윤 202104514 박규수

1. 연구개발의 필요성

- 일반적으로 전문 지식이 없는 일반 독자가 과학 문서와 기술 문서를 읽을 시, 문서를 어렵게 느끼고, 이해하지 못함
- 과학 및 기술 분야의 문서의 전문 용어나 개념을 알지 못하기 때문임
- 이를 이해하기 위해서는 해당 단어에 대한 추가적인 탐색 과정을 거쳐야 함
- 탐색 과정에는 해당 단어 관련 서적이나 문서 찾기, 온라인 검색 등이 존재하며 인터넷 환경이 준비되어있다는 가정하에 인터넷 검색이 가장 많이 사용됨
- 인터넷 검색을 사용시, 검색을 위해 검색엔진을 켜고 해당 단어를 검색하고 지식이 있는 웹 사이트에서 지식을 얻는 과정을 수행하게 됨
- 그러나 위 방식은 탐색할 단어가 많아질 수록 독자에게 번거로움
- 이를 해결하기 위해 해외 웹 브라우저 플러그인인 'Wiki-Definition' 등이 개발되었으나, 한국어를 지원하지 않고, 동음이의어가 있을시 원하는 단어와 다른 단어를 제공하는 등의 한계를 보임
- 위의 한계가 있더라도 'Wiki-definition'의 기능인 원하는 단어에 드래그를 했을 때, 그 단어의 정의를 제공하는 방식은 기존의 번거로운 검색 방식보다 단어의 정의를 빠르게 제공해 줄 수 있는바, 한국어 지원 및 동음이의어 처리 등의 한계를 극복한 새로운 플러그인의 개발을 한다면 과학 문서와 기술 문서의 독자들로 하여금 더 편하게 문서를 이해하는데 도움을 줄 수 있을 것이다.

2. 연구개발의 목표 및 내용

- 연구개발 목표

본 연구는 전문 지식의 부재로 과학과 기술 문서를 읽는데 어려움을 겪는 일반 독자들로 하여금 번거로운 검색 과정 없이 모르는 단어의 정의 및 정보를 얻을 수 있게 함으로써 검색에 들어가는 시간과 노력을 줄이고자 하며, 궁극적으로 과학과 기술 문서의 이해를 돕고자 한다.

- 연구개발 내용 및 범위

1. Web browser 플러그인 개발
2. 단어 검색 및 결과 반환 Web Crawler 개발
3. 동음이의어 처리 ai 개발

- TO-BE 계획

1. 정의를 원하는 단어에 드래그 시 해당단어의 정의가 띄워짐
2. 현재 웹페이지의 많은 단어를 한번에 해석하고 싶을 시 상단 플러그인 버튼 클릭시 모든 명사의 정의가 띄워짐
3. 동음이의어가 있더라도 문맥에 맞는 의미를 가진 단어의 정의가 반환됨
4. 정의를 얻은 단어를 클릭시 사이드바에 더 많은 정보(wiki링크, 연관 단어등)가 표시됨

3. 연구개발의 추진전략 및 방법

-
1. Web browser 플러그인

– Chrome 브라우저의 API를 사용하여 단어에 드래그를 했을 때와 플러그인 버튼을 눌렀을 때, 해당하는 단어나 페이지내의 단어를 검색하여 정의를 반환한다.
2. 단어 검색 및 결과 반환 Web Crowler

– 웹 크롤러를 개발하여 단어를 검색하고 정의를 추출해 내도록 한다.
3. 동음이의어 처리 ai

– 정의를 원하는 단어와 주변 단어를 같이 입력으로 주어 LLM을 이용하여 해당 단어가 어떤 뜻을 가지고 있는지 파악하도록 한다.

4. 연구개발 팀의 구성 및 프로젝트 추진 일정

팀 구성	구분	성명	참여내용 및 역할			
	팀장	양정욱	동음이의어 처리 ai개발			
	팀원	김재윤	Web Crowler 개발			
	팀원	박규수	Web Browser 플러그인 개발			
추진 일정	세부내용		수행기간(월)			
			3	4	5	6
	LLM 파인 튜닝을 위한 데이터셋 수집					
	LLM 파인 튜닝					
	Web Browser 플러그인 개발					
	Web Crowler 개발					
	결과물 병합 및 디버깅					

※ 반드시 5쪽 내외의 분량으로 작성

– 참고문헌(References)

웹 크롤러란? / 웹 스파이더의 작동 원리.n.d.Cloudflare.

<https://www.cloudflare.com/ko-kr/learning/bots/what-is-a-web-crawler/>

API 참조 문서.(2024.03.20(UTC)).Chrome for Developers.

<https://developer.chrome.com/docs/extensions/reference/api?hl=ko>

파인튜닝(Fine-tuning)이란? – LLM 구축 방법.(2023.12.14).appn.

<https://kr.appn.com/blog/fine-tuning/>