

15

엑소브레인 한국어 언어분석 툴킷 v3.0 기술

기술개요

- 한국어 문장 또는 문서를 입력 받아 어휘의 품사 및 의미, 개체명의 인식, 문장의 구조, 의미적 역할을 분석하는 기술

- 한국어 문어체 텍스트에 대해서 형태소분석, 개체명인식, 구문분석, 동음이의어/다의어 분석, 의미역 인식의 언어 분석을 수행하고, 언어분석된 결과를 전달해 주는 기술

기술의 특징점

- 언어학자 수준으로 어휘의 의미, 개체명 인식, 문장의 문법 등을 복합적으로 분석 가능한 한국어 분석 시스템 개발

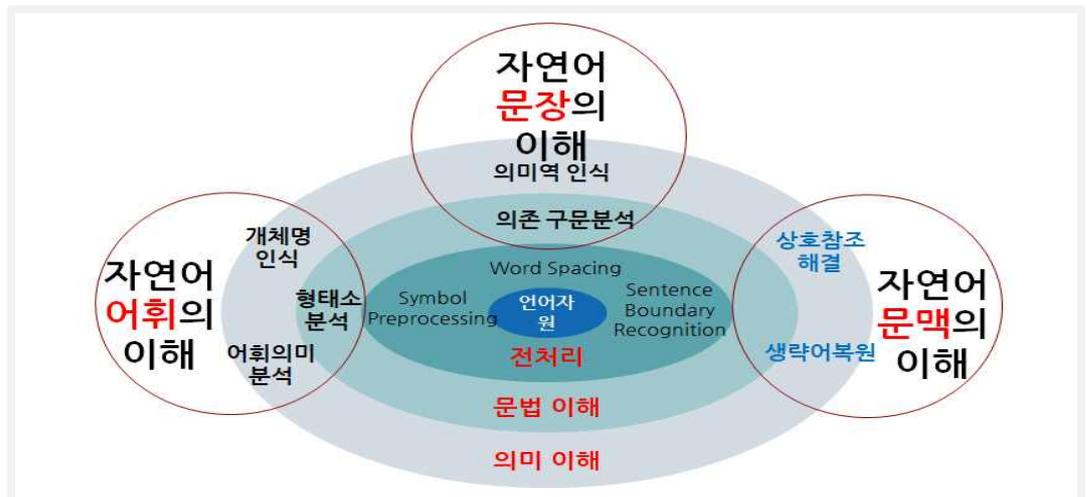
- 형태소분석 기술은 기계학습 방법에 기반해서 형태소를 자동으로 분석해 줌

- 개체명인식 기술은 sequence labeling 방법론에 기반해서 개체명 경계 및 대분류 인식

- 구문분석 기술은 세계적으로 가장 많이 연구되고 있는 의존구조에 기반하여 문장의 구조 분석

- 동음이의어 및 다의어 분석 기술은 문맥에 가장 적합한 의미를 분석

- 의미역인식 기술은 영어권을 중심으로 널리 쓰이는 Propbank 의미역 체계를 도입



적용분야

- 콜센터 상담 솔루션, VOC 분석, 빅데이터 분석, AI 챗봇, AI 플랫폼



기술완성도 (TRL)

- TRL 6단계; 파일럿 규모 시작품 제작 및 성능 평가



기술이전 내용 및 범위

- 한국어 형태소분석 기술
 - 한국어 형태소분석 엔진, 리소스
 - 언어분석 기술 사용매뉴얼, 자료구조매뉴얼
- 한국어 개체명인식 기술
 - 한국어 구문분석 엔진, 리소스
 - 언어분석 기술 사용매뉴얼, 자료구조매뉴얼
- 한국어 구문분석 기술
 - 한국어 구문분석 엔진, 리소스
 - 언어분석 기술 사용매뉴얼, 자료구조매뉴얼
- 한국어 동음이의어 및 다의어 분석 기술
 - 한국어 동음이의어 및 다의어 분석 엔진, 리소스
 - 언어분석 기술 사용매뉴얼, 자료구조매뉴얼
- 한국어 의미역인식 기술
 - 한국어 의미역인식 엔진, 리소스
 - 언어분석 기술 사용매뉴얼, 자료구조매뉴얼
- 한국어 언어분석 툴킷 전체
 - 한국어 형태소분석 엔진, 리소스
 - 한국어 개체명인식 엔진, 리소스
 - 한국어 구문분석 엔진, 리소스
 - 한국어 동음이의어 및 다의어 분석 엔진, 리소스
 - 한국어 의미역인식 엔진, 리소스
 - 언어분석 기술 사용매뉴얼, 자료구조매뉴얼

관련 지재권 현황

No.	출원번호	특허 명	상태
1	2017-0159393	채팅로봇 시스템 및 서비스방법	출원

기술이전 문의

- 연구성과확산실 (02-597-3387 / junsung.kim@etri.re.kr)