

엑소브레인 말뭉치 V2.0 (언어분석/시간/공간) 울산대학교 형태/의미 말뭉치

엑소브레인 과제 소개 드립니다.

- ✓ 인공지능 분야의 선도를 목표로 자연어 질의응답 기술을 개발하기 위한 미래창조과학부 소프트웨어 분야 국가 혁신기술 개발형 R&D 과제입니다.
- ✓ 전문분야 도메인에서 인간과 기계가 지식소통이 가능한 자연어 질의응답 시스템 개발을 목표로 합니다.
- ✓ 다양한 산업환경에서 전문가 수준의 질의응답 서비스의 제공을 위하여 ETRI, 솔트룩스, 울산대, 카이스트, 충북대 등 국내 여러 연구기관이 힘을 합쳐 연구하고 있습니다.

엑소브레인 언어분석 말뭉치(ETRI) 소개 드립니다.

- 1) **통합 언어분석 말뭉치**: 자연어 질의응답을 위한 질문과 정답 단락의 쌍으로 구성된 5개 기술(형태소분석, 세분류 개체명인식, 다의어 어휘의미분석, 의존구문분석, 의미역인식)의 태깅 가이드라인 및 말뭉치 (1,811문장, 21,355어절)
 - 2) **개체명 태깅 말뭉치**: 개체명 태깅 가이드라인 및 말뭉치 (인명/장소/조직/날짜/시간 5개 태그, 5,000 문장)
- ※ UTF-8 인코딩된 JSON 및 TEXT 파일 포맷 / 연락처 : 임준호 박사, joonho.lim@etri.re.kr

엑소브레인 Korean TimeBank 및 SpaceBank 소개 드립니다.

- ✓ Korean TimeBank (KAIST)
 - **한국어 시간 정보 주석 말뭉치**: 한국어 시간 정보 자동 추출을 위한 한국어 시간 정보 주석 가이드라인 및 말뭉치
 - 말뭉치 구성: 2,393 문서, 6,190 문장 태깅 (시간개체 3,290개, 시간관계 4,545개)
 - 연락처: 임채균, rayote@kaist.ac.kr
- ✓ Korean SpaceBank (충북대학교)
 - **한국어 공간 정보 주석 말뭉치**: 한국어 공간 정보 자동 추출을 위한 한국어 공간 정보 주석 가이드라인 및 말뭉치
 - 말뭉치 구성: 개체 태그 7종류, 관계 태그 4종류, 1,654문장 태깅
 - 연락처: 김보겸 박사 bogyum@cbnu.ac.kr, 강명윤 석사과정 myungyun@cbnu.ac.kr

울산대학교 형태/의미 말뭉치(UCorpus-MA/HG) 소개 드립니다.

- ✓ 21세기 세종계획 형태의미 말뭉치를 문장단위로 편집하고 여러 오류를 수정하고, 어깨번호가 부착되지 않은 부사,관형사,접사에 대해서도 어깨번호를 부착한 말뭉치
- ✓ 전체 942,833 문장, 11,048,780 어절 (학습/평가 말뭉치 분리 제공)
- ✓ 연락처: 옥철영 교수 okcy@ulsan.ac.kr



한국전자통신연구원
자동통역·언어지능연구부



울산대학교
한국어처리연구실



한국과학기술원
지식공학및집단지성연구실



충북대학교
언어지식공학연구실