

摘要

專有名詞辨識，或實體辨識 (Named Entity Recognition, NER)，是資訊抽取 (Information Extraction, IE) 的基本任務之一。其目標為在文字資料當中，能夠辨識出感興趣的專有名詞(包含原本資料庫不存在的新專有名詞)，並自動標記正確的分類，如人名、地名、組織名等等，是人工智慧當中理解語言的重要步驟。目前我們所開發的中文專有名詞辨識系統能辨識11類一般領域專有名詞及7類數量詞，包含：人名、團體、設施、組織、地理、地點、商品、事件、藝術品、法律、語言、日期、時間、比例、錢、數量、序數、數詞。

線上展示網址為：<https://ckip.iis.sinica.edu.tw/project/ner>，歡迎實際測試。

技術優勢

1. 能夠充分利用語法結構的深度學習技術
2. 同時考慮字與詞的訊息
3. 辨識種類多：11類一般領域專有名詞及7類數量詞
4. 相關技術發表在NLP著名國際會議 - EMNLP 2017, 是當年度的state-of-the-art (Ontonotes 5.0 English NER)

CKIP NER System Demo

Paste the text you want to process here(Chinese only):

美國參議院針對今天總統布什所提名的勞工部長趙小蘭展開認可聽證會，預料趙小蘭將會很順利通過參議院支持，成為美國有史以來第一位的華裔女性內閣成員。

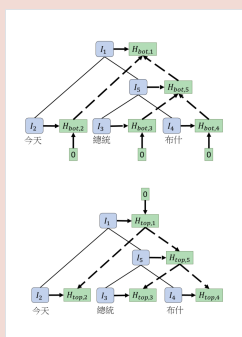
- CKIP WordSeg
- CKIP Parser

Clear

Submit

美國參議院ORG 針對今天DATE 總統布什PERSON 所提名的勞工部長ORG 趙小蘭PERSON 展開認可聽證會，預料趙小蘭PERSON 將會很順利通過參議院ORG 支持，成為美國GPE 有史以來第一ORDINAL 位的華裔NORP 女性內閣成員。

圖1. CKIP NER 系統示意



本院覽號

05T-1080409

公告日期

2024-01-19

智財權狀態

know-how

應用範圍

1. 大數據輿情分析
2. 語言理解
3. 智慧客服
4. 聊天機器人
5. 商品情報分析系統

創作人

馬偉雲