

經濟部 108 年度
《人工智慧產業關鍵技術拔尖計畫(2/4)》
合作研究計畫

《對話短句命名實體辨識研究》

建議書徵求文件

財團法人資訊工業策進會

中華民國 108 年 02 月 18 日

108年度合作研究計畫建議書徵求文件

一、簡介

對話管理可細分為對話狀態追蹤及對話策略兩部分，透過原始對話文字來建立對話策略模型。其中對話狀態追蹤的優化為重點的研發項目，利用原始對話文字以外的特徵來優化對話狀態追蹤已被驗證可行，例如運用在文字對話的命名實體辨識資訊進行對話策略模型的優化。本項目投入中文對話短句命名實體辨識之技術研究，將可有效幫助主計畫的對話管理之對話狀態追蹤模型優化。

二、計畫目標

收集對話語料與標記，研發適用於短句的實體辨識技術，運用對話短句的命名實體辨識結果來優化主計畫的對話管理之對話狀態追蹤模型，提升對話狀態追蹤的準確度。

三、計畫範圍

- 中文對話短句命名實體辨識機制
 - 建構：中文對話短句命名實體辨識機制
 - 驗證：研發特定領域中文對話短句命名實體辨識機制，可以在對話短句中偵測出命名實體的位置與類別，至少支援1個領域，如金融保險。

四、預期成果

- 完成對話短句命名實體辨識研究之雛型軟體
- 完成對話資料收集500句
- 完成對話資料之對話短句命名實體辨識格式定義與標記500句

※前述成果如有專利構想或專利申請產出時，需注意專利申請之新穎性(novelty)。因凡經公開發表之研發成果，如擬申請專利，須於公開發表後6個月內完成，前述成果如是以論文方式公開發表，將無法取得大陸與歐盟等國之專利。(※文字請保留，此括號文字請於正式版時刪除)

五、執行方式

- 合作計畫執行單位需定期舉行計畫控管會議，確保計畫的進度及執行方向一致性。
- 合作計畫執行單位需於6、12月繳交研究報告共二份。本單位要求合作計畫執行單位完成對話短句命名實體辨識研究之雛型。

六、計畫期程及預估計畫總經費

計畫執行區間：108年01月01日至108年12月15日

總經費：800,000元

七、驗收標準

- 2019/06 交付期中報告1份，內容應包括：研究動機、架構規劃、演算法設計與分析、使用技術、文獻搜尋
- 2019/12 交付期末報告1份，內容應包括：研究動機、架構規劃、使用技術、實驗設計、實驗結果、結論、文獻搜尋
- 2019/12 交付[對話短句命名實體辨識研究之雛型軟體]、[對話資料收集500句]及[對話資料之對話短句命名實體辨識格式定義與標記500句]

八、技術能力需求(請詳述所需要之技術能力或專長)

具中文自然語言處理研究經驗與成果之技術研究人員。

附件1：契約書格式

1-1：計畫書格式

1-2：經費動支報表

1-3：成果報告撰寫須知

1-4：報告格式

1-5：論文格式

1-6：保密聲明書

1-7：委託匯款同意書