

經濟部 107 年度  
人工智慧產業關鍵技術拔尖計畫 (1/4)  
合作研究計畫

省略與同指涉現象偵測與還原之技術與研究  
建議書徵求文件

財團法人資訊工業策進會

中華民國 107 年 01 月 23 日

# 107年度合作研究計畫建議書徵求文件

## 一、 簡介

建構一個優良的目標導向對話系統必須以高準確度的自然語言理解機制為基礎，然而在欲提升中文對話的自然語言理解準確度，目前仍存在許多待解問題，如在中文對話過程中，句子所出現的省略及同指涉現象是為影響自然語言理解準確度的重要因素。根據實驗統計，中文自然語言對話中出現省略及同指涉現象大約佔25%，省略及同指涉造成句子中部份資訊流失，將導致使用者所提出內容無法有效地對應到訓練資料集的句子，自然語言理解的準確度將會因此受限。若能正確地偵測中文對話過程中所產生的省略與同指涉情況，並回復流失的資訊，預計將能提升自然語言理解的準確度。

## 二、 計畫目標

發展支援文字對話前處理，可偵測中文對話句子中出現的省略與同指涉現象，並進行回復被簡化或省略的資訊，將其還原成資訊完整的句子，回饋以提升主計畫於自然語言理解的準確度。

## 三、 計畫範圍

1. 中文對話省略現象偵測與還原機制
  - 建構：中文對話省略現象偵測與還原機制。
  - 驗證：研發特定領域中文對話省略現象偵測與還原機制，可以偵測中文對話句子中省略現象，並將其還原成資訊完整的句子，至少支援1個領域，如電信。
2. 中文對話同指涉現象偵測與還原機制
  - 建構：中文對話同指涉現象偵測與還原機制。
  - 驗證：研發特定領域中文對話同指涉現象偵測與還原機制，可以偵測中文對話句子中同指涉現象，並將其還原成資訊完整的句子，至少支援1個領域，如電信。

## 四、 預期成果

- 完成一個領域的中文對話省略現象偵測與還原雛型。
- 完成一個領域的中文對話同指涉現象偵測與還原雛型。

※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。

## 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中文對話省略與同指涉現象偵測與還原雛型。

## 六、計畫期程及預估計畫總經費

計畫執行區間：107年01月01日至107年12月15日

總經費：600,000元

## 七、驗收標準

- 完成一個領域的中文對話省略現象偵測與還原離型。
- 完成一個領域的中文對話同指涉現象偵測與還原離型。
- 2018/06 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。
- 2018/12 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。

## 八、技術能力需求

- 具中文對話省略與同指涉現象偵測與還原研究經驗與成果之學術機構與技術研究人員。

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