

經濟部 110 年度  
《人工智慧產業關鍵技術拔尖計畫 (4/4)》  
合作研究計畫

《運用深度學習知識嵌入之知識圖譜跨知識節點問答技術》  
建議書徵求文件

財團法人資訊工業策進會

中華民國 110 年 03 月 18 日

# 110年度合作研究計畫建議書徵求文件

## 一、 簡介(說明本合作研究計畫之背景、動機、目的及重要性)

有別於知識圖譜推理(Knowledge Graph Reasoning)以缺少目標實體的三元關係作為輸入，知識圖譜問答技術(Knowledge Base Question Answering)旨在提供使用者透過自然語言的方式詢問知識圖譜的內容，現階段Google已經支援智慧搜尋服務，使用者可以在搜尋引擎中輸入自然語言問句，如：歐巴馬太太出生地，會自動回覆美國伊利諾伊州芝加哥，透過推論解析找出知識圖譜中不同知識節點的答案。

因此在知識圖譜問答技術中，除了知識推理解析，還需考量如何解析自然語言問句，取出與知識圖譜相關的知識節點與關聯特徵資訊，並可以從特定領域的知識圖譜提供跨一個或多個知識節點的演算法，找到對應的答案，可提供該領域及企業問答應用，是一個很重要的研發議題。

## 二、 計畫目標(應包含本合作研究計畫預期可達成或量化的目標)

從自然語言解析使用者問句時，很難從知識圖譜單一個知識節點、關聯與屬性回答問題，加上知識圖譜通常是稀疏不完善的，並且在現實場景中，更常面對到連續性的問答對話議題，本計畫也將研究如何透過知識圖譜嵌入技術(Knowledge Graph Embedding)，將問句與知識圖譜映射到向量空間，並且考量透過注意力機制(Attention)自動捕捉新問句與前問句的關聯，藉由更新權重獲取問題的重點單詞，此外也會模擬更複雜結構的句子，建構包含邏輯推理、數量推理、比較級推理的複雜式問答，透過聯接預測解析(Link Prediction)來找到跨一個或多個(Multi-hop)知識節點或屬性來回答，同時也可以解決知識圖譜中知識節點與關係不直接連接(間接關聯)的議題，透過實驗評估現今解決該任務的方法，並設計改良出一個端到端的基於知識圖譜的問答系統。

## 三、 計畫範圍(說明本合作研究計畫所需執行之項目)

運用深度學習模型建立可適用於中文知識圖譜的問答知識推論演算法，：

- 將問句透過深度學習模型建立向量空間
- 將知識圖譜映射到向量空間
- 可從現有的知識圖譜關聯中，透過連接預測解析(Link Prediction)找出跨一個或多個知識節點或屬性的未直接相連之三元關係

工作項目範圍包含：

- (1) 研讀以深度學習相關知識圖譜問答演算法之文獻
- (2) 建構特定領域知識圖譜以及關聯該知識圖譜的問題答案對(Question-Answer pairs)資料集
- (3) 建立自然語言解析機制，根據解析結果查詢在知識圖譜的三元關係和關聯特徵資訊，建模推論答案機制並與既有技術進行比較評估
- (4) 建立完整知識圖譜問答模型，透過輸入自然語言問句，提供在知識圖譜尋找答案節點的預測機制
- (5) 撰寫研究報告，含實驗評估模型有效性

#### 四、 預期成果(說明在執行期限內應完成之工作項目/成果及交付時程)

- (1) 深度學習跨知識節點相關問答知識問答推論技術之演算法評估。
- (2) 特定領域知識圖譜及作為新增語料之目標問句及關聯。
- (3) 具名實體與關聯標記資料結構，及實體關聯資料集(含驗證與測試)。
- (4) 本案深度學習演算法與既有技術演算法之實驗比較。
- (5) 本案深度學習演算法程式雛型一份，包含跨知識節點問答機制。
- (6) 研究報告一份。

※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(包括計畫時程、計畫分工方式、執行事項，但不限於前述項目)

查核點	時間	產出物	報告架構/大綱
M1	2021/07/30	期中報告	<ul style="list-style-type: none"><li>■ 文獻探討</li><li>■ 建立知識圖譜跨知識節點問答語料</li><li>■ 知識圖譜跨知識節點問答演算法設計</li></ul>
M2	2021/11/30	研究報告 程式雛形	<ul style="list-style-type: none"><li>■ 知識圖譜跨知識節點問答期末研究報告書 (期末報告需包含雛形軟體與既有技術之實驗比較)</li><li>■ 雛形軟體成果與操作說明</li></ul>

#### 六、 計畫期程及預估計畫總經費

計畫執行區間：110年04月20日至110年12月10日

總經費：800,000元

## 七、 驗收標準(含教育訓練)(詳述驗收項目、交付形式、內容及時程等)

- (1) 知識圖譜跨知識節點問答技術實作程式(包含雛形軟體與既有技術之實驗比較，包含跨知識節點問答機制，特定領域知識圖譜及目標文本相關語料)。
- (2) 研究報告一份。

## 八、 技術能力需求(請詳述所需要之技術能力或專長)

- 熟悉深度學習演算法設計研究知識圖譜跨知識節點問答技術，使問句可不受限知識圖譜相連之結構，找到跨一個或多個知識節點的演算法。
- 熟悉知識圖譜問答建構相關技術。
- 具備機器學習和深度學習演算法設計經驗，與研究分析能力。
- 具備程式撰寫與軟體架構開發能力，實作演算法原型，並進行實驗量化驗證。