

# 自然語言處理實驗室

## Natural Language Processing Laboratory

指導老師：林川傑 (cjlin@email.ntou.edu.tw)

實驗室網址：nlp.cse.ntou.edu.tw

### 研究目標

自然語言處理一直是人工智慧研究重要領域。如何讓電腦分析與處理人類語言、讓電腦理解人類意念表達後提供服務、讓電腦與使用者能以人的語言溝通，可說是電腦科學終極目標，也是未來改變世界的重要資訊科技之一。這個領域的研究近年來極為受到重視，未來所有電腦系統都需要更友善的見面，此時自然語言處理扮演極重要的角色。

本實驗室的研究目標，即為所有與自然語言處理技術相關的課題和應用，包括：

#### ❖ 台灣本土語言處理

本實驗室研究不僅針對英文和中文，也可擴展至台灣本地所使用的各種語言，像是台語、客家語、各原住民族語言、手語等。

#### ❖ 社群網站及部落格之資訊處理

網際網路上的資訊處理是目前最熱門的課題。除了一般網站外，社群、討論區、電商平台這些大量的文字資訊也是重要資源。

#### ❖ 自動問答系統

一個能接受自然語言問句、並從大量自然語言文件或網際網路上找尋答案的系統。回應包括答案本身以及出現的網頁，提供迅速找尋答案的方法。

#### ❖ 資訊檢索系統

給定使用者心中的資訊需求描述、由大量自然語言文件中選取主題上高度相關文件之工作。

這些技術的發展出可提供學生研究訓練之外，未來成果均可實際應用在各種資訊系統以及網際網路線上系統之上。

### 研究計畫 (部份)

- ❖ 使用詞綴資訊之泰雅語群與中文神經網路機器翻譯 (NSTC 112-2221-E-019 -049 -)
- ❖ 臺灣南島語與中文之神經網路機器翻譯 (MOST 111-2221-E-019 -070 -)
- ❖ 賽德克語-中文機器翻譯之研究 (MOST 110-2221-E-019-068 -)
- ❖ 賽德克語構詞結構之自動解析及標記工作 (MOST 109-2221-E-019-053 -)
- ❖ 中文歧視性文字之自動偵測與分類 (MOST 108-2221-E-019-045 -)
- ❖ 自動偵測並改寫中文不雅文字 (MOST 107-2221-E-019-038 -)

### 論文發表 (部份)

- Chuan-Jie Lin, Li-May Sung, Jing-Sheng You, Wei Wang, Cheng-Hsun Lee, and Zih-Cyuan Liao (2020) "Analyzing the Morphological Structures in Seediq Words," *International Journal of Computational Linguistics and Chinese Language Processing*, Volume 25, Number 2, December, 2020, pp. 1-20. [THCI Core]
- Chuan-Jie Lin, Chien-Wei Pao, Yen-Heng Chen, Chi-Ting Liu, and Hui-Huang Hsu (2016) "Ellipsis and Coreference Resolution in a Computerized Virtual Patient Dialogue System," *Journal of Medical Systems*, Volume 40, Issue 9, September, 2016, pp.1-15.
- Chuan-Jie Lin, Cheng-Wei Lee, Cheng-Wei Shih, and Wen-Lian Hsu (2015) "Rank Correlation Analysis of RITE Datasets and Evaluation Metrics – An Observation on NTCIR-10 RITE Chinese Subtasks," *Web Intelligence*, Vol. 13, No. 2, July, 2015, pp. 123-133.
- Chuan-Jie Lin and Wei-Cheng Chu (2015) "A Study on Chinese Spelling Check Using Confusion Sets and N-gram Statistics," *International Journal of Computational Linguistics and Chinese Language Processing*, Vol. 20, No. 1, June, 2015, pp. 23-48.

### 研究設備

GPU電腦 3台  
個人電腦 10台  
印表機 2台