台灣大學資訊工程學系 陳信希教授:自然語言處理

1. 自然語言處理主要應用
2. 機器翻譯：不同語言間的翻譯(EX.英文翻譯成中文)

困境：兩語言間的語法結構不同與詞彙的一字多意會使翻譯上有所差異

應用：圖像文字或語音整合的翻譯、與地理位置整合的方言翻譯(EX.北京話、四川話)

1. 問答系統：提問與回答問題

挑戰：分析問題(問什麼)、分析內容(正確答案)、支持或反駁資訊的信心度

應用技術:自然語言處理、資料檢索、機器學習、知識表示和推理、大規模平行計算

1. 文本分析:對不同種類來源的數位資料，進行文件探勘，針對某議題做追蹤與分析

應用：市場分析、多標的與情分析

1. 自然語言處理的典型問題
2. 容錯力：打錯字、語法錯誤、未知詞(EX.網美)
3. 強健性：領域改變、網路語言(EX.諧音 耗子尾汁=好自為之) 、表情符號
4. 分類：掌握共通性
5. 詞性類別：標記詞性EX.普通名詞(Na)、地方詞(Nc)、不及物動詞(VH)
6. 語意類別：語意關係標記EX.悲憤、悲痛、悲慟、悲戚、悲慘、辛酸、沉痛
7. 句法類別：剖析句法EX.介系詞片語、動名詞片語、不定詞片語…
8. 相依類別：相依剖析，詞彙間的分類
9. 言談類別：言談剖析EX.：時序、因果、轉折、推展
10. 意見類別：意見探勘，分類正面與負面的詞彙
11. 情感類別：情感分析 EX.憤怒、厭惡、恐懼、喜悅、悲傷、驚訝…
12. 立場類別:辨別立場 EX.贊成或反對