(1)背景應用:1.機器翻譯、2.問答系統、3.意見探勘

1.Translate 牽涉到的議題:詞彙面相、結構面向

2.問答系統的問題:文件檢索、分析問題、答案選擇、代名詞指涉、自然語

言處理、機器學習、平行計算。

方法:必須先了解問題，在作檢索的動作，然後再作答案選擇，選擇的

答案必須避免代名詞指涉。

3. 意見探勘: 文本分析，對多個不同資料型態進行分析定義，定義資料就

是用文字呈現出來的。

例:產品評論、市場分析、社群網路分析、熱門議題分析等

(2)自然語言處理問題:

歧異解釋:詞彙層次、語法分析、語義層次

用符號組合代表概念就會有歧異性。

容錯能力:語法錯誤、打錯字、未知性

強健性:領域改變、網路語言、表情符號

掌握規律性:包括詞性類別，如動詞、名詞、副詞等，

語意類別，如夕陽、斜陽(同義詞)，

句法類別，名詞片語、介係詞片語、句子等，

相依類別，詞語詞的關係，

言談類別，時序句、因果句、轉折句、推展句，

意見類別，針對議題給出正面、負面，

情感類別，驚訝、悲傷、喜悅，

立場類別，對議題贊成或反對，

以上都是所謂的符號計算。

(3)將詞彙轉成向量

利用向量計算，計算詞的相關性，詞彙在向量中是以高維度所表示，利用consine，來計算兩個詞的相關性，夾角越小，關係越緊密，高維計算複雜

可以進行降維(word2vec)。