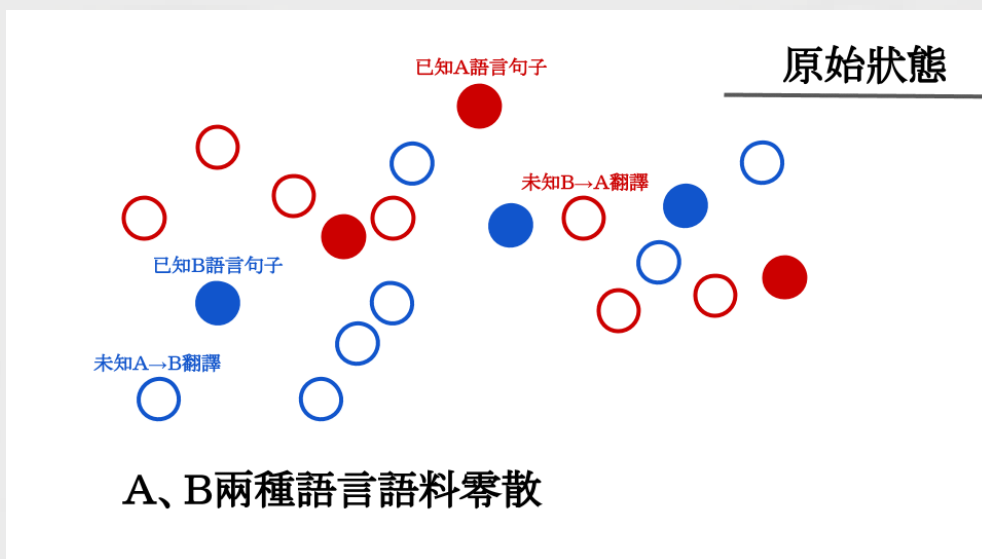


機器翻譯技術的發達，使字典的重要性日漸式微。（圖片來源 / 呂奕廷攝）

除此之外，一些應用像是社群網站的內容過濾與審查機制，也常常會配合機器翻譯的技術，盡可能地減少人工審查。不過以現階段的技術而言，由於缺乏某些語種的語料，全自動化審查在某些地區仍無法落實。也就是說，礙於語種較小、語料庫不全的機器翻譯窘境，在非監督式機器翻譯的出現與發展下，能獲得一絲希望的曙光。那麼，非監督式機器翻譯到底是如何運作的呢？

非監督式機器翻譯 讓機器自己學習

Guillaume Lample 與其研究團隊提出，要達成非監督式機器翻譯，必須有三個原則：「雙語逐字對照 (word-by-word initialization)」透過匯入雙語詞典，並結合詞嵌法 (word embeddings) 技術將每個辭彙與辭彙之間做向量標示，計算出不同語言內單詞之間彼此的關係；「語言建模 (language modeling)」是將語言內所有字符、單詞做一個以「機率分布」為導向的模型，使機器能夠分辨哪些語詞之間的搭配是合理的，盡可能地造出如正常人說話般符合邏輯的句子；「回譯 (back translation)」透過目標語言反向翻譯回原本語言，可以檢查原本的翻譯是否有誤，進而調整參數，改善翻譯品質，是非監督式機器翻譯訓練上十分重要的一環。



「非監督式」機器翻譯技術三大原則運作模式示意動圖。（圖片來源 / 呂奕廷重製）資料來源：[〈基於短語與非監督式機器翻譯，2018〉](#)

除了這三大原則以外，該技術也會結合目前神經機器翻譯 (Neural Machine Translation, NMT) 及詞語統計機器翻譯 (Phrase-based Statistical Machine Translation, SMT) 等技術。國立交通大學機構典藏系統版權所有 Produced by IR@NCTU

Translation, PBSMT) 兩個系統來增進翻譯效能。前者學習力強，能夠適應龐大數據通順地翻譯；後者能夠分析平行語料 (parallel data)，也就是人為雙向語言譯本，使翻譯正確率提升。結合兩個系統，將彼此優缺點互補，便能打造出既正確又通順的翻譯系統。

綜上所述，結合以上技術所打造出的「非監督式」機器翻譯，是藉由機器模型獲取大量語言資料後從中歸納出翻譯規則、統計出各用詞的出現機率與模式，也就是說不需要人為告知做出來的翻譯正不正確，也不需要人為翻譯文本，機器便可自己學習到語言之間的結構，進而優化翻譯效能。

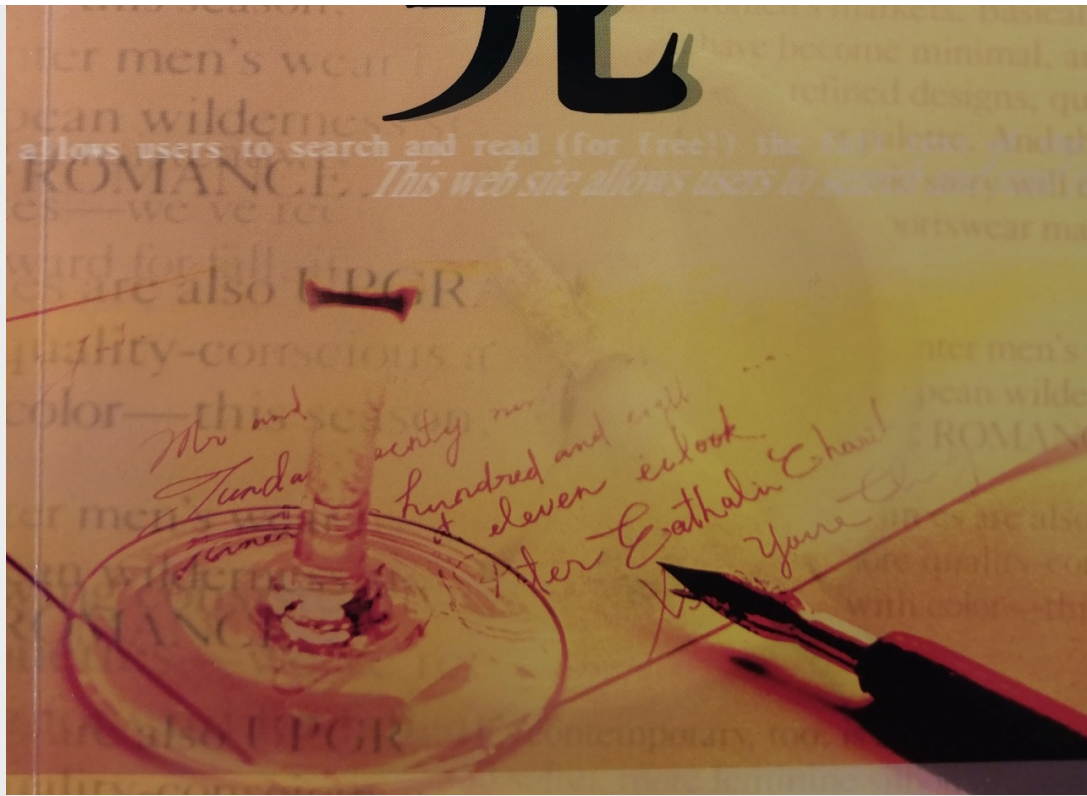
機器翻譯蓬勃發展 職業譯者的未來

機器翻譯技術的日新月異，讓我們能更方便地用自己熟悉的語言去理解外來的資訊，人與人的溝通上也更沒有語言隔閡。但是在這樣的時代背景下，科技終究影響了一批以翻譯為業的人群，在大部分的翻譯都能藉由機器達成的將來，職業譯者該怎麼順應這樣的變遷呢？

交通大學外國語文學系的兼任講師與自由譯者吳煒聲認為，現代譯者應該學習與科技共處，善用機器翻譯技術帶來的方便，配合譯者的翻譯技巧完成譯文。不過，即使機器翻譯勢不可擋，能夠處理幾乎所有以「溝通」為訴求的翻譯，他認為牽扯到藝術與美的「文學」領域，是永遠不可能被機器翻譯所接管的。

思果也在其著作《翻譯研究》中提到：「翻譯最重要的工作是思想。譯而不思，雖然譯得久也沒有用。」也就是說，譯者在進行翻譯的時候，除了必須思考原句的意思，選用目標語言中最貼切的辭彙，還要讓譯文像是目標語言中真正存在的句子。以傳遞資訊為目的的翻譯也許不必嚴格要求，但是字裡行間本身就是藝術的文學作品，以目前的科技來說仍無法透過機器翻譯達成。





思果所著的《翻譯研究》被許多人認為是譯者入門必讀的書目之一。(圖片來源 / 呂奕廷攝)

科技在地化 機器翻譯與台語

機器翻譯的發展不只侷限於不同國家的語言，隨著近年來本土意識的提升，機器翻譯也逐漸關注到在地語言，增強科技與在地的連結。以去年工研院研發的「台語語音轉中文」雙向翻譯系統為例，利用大量台語語料建立的文字、語音翻譯系統，能拉近年長者與社會的距離，讓他們能用自己熟悉的語言，接觸網路上的資訊，也能將自己的意見傳達給不熟悉台語的年輕族群。

工研院
文字轉語音 Web服務
ITRI TTS @ Web

認識我們 | 開發專區 | 線上語音 | 會員登錄

首頁 > 線上語音 > 測用展示

測用展示

音量大小: 100

- 中英統合 Bruce (男)
- 中英切換 Dale (男)
- 英文語音 Bob (男)
- 台語語音 A-Koan (女) 白話台

你好，我是a嬌，感謝您使用工研院文字轉語音網站服務。

試聽

若您需要更多文字輸入、或需進階的音量大小、速度快慢及韻律調整等功能、下載聲音檔，請加入免費會員，即可享有進階合成下載權益。

工研院研發測用展示的文字轉語音系統就包含台語的選項，並且可以在白
國立交通大學機構典藏系統版權所有 Produced by IR@NCTU

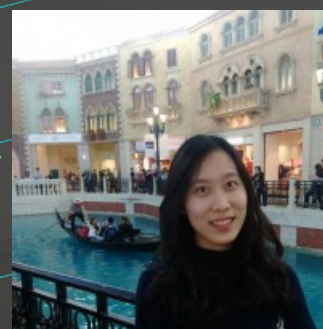
話、文讀之間做選擇。(圖片來源 / [工研院文字轉語音服務網](#))

不過針對機器翻譯對在地語言保存的貢獻，吳煒聲認為，就長期而言，方言的翻譯並不會帶來龐大的實質效益，也無助於弱勢語言的復甦。在未來，全世界的語言模式會日漸趨同，並以強勢語言為主體發展，使弱勢語言越來越銷聲匿跡，這是連語言教育、文化推廣政策都沒辦法抵擋的時代變化。畢竟，語言的流通是奠基於「使用動機」，在缺乏學習動機的情況下，即使機器翻譯朝在地化發展，仍無法扭轉語言流失的局面。機器翻譯對弱勢語言恐怕只有語料保存的功能，實際上並不能遏止語言從世界上消逝。

機器翻譯對於大量資訊的流通，以及社群網站的內容審查機制都有莫大的幫助，隨著科技的持續發展，語言越來越不是人與人交流的阻礙。但是機器翻譯畢竟不是萬靈丹，無法解決所有語言的問題，針對在地語言保存等議題，未來還需要更進一步的討論。



記者 呂奕廷



編輯 梁書瑜