

機器翻譯之誤譯探討（第 293 期 2022/3/24）

劉映秀*

一、前言

機器翻譯已廣泛運用在專利前案檢索，舉凡審查委員用來考量可專利性的引證案，或是申請人依規定提交已知悉的前案，主要的官方機構，如世界智慧財產權組織 (WIPO)、歐洲專利局 (EPO) 都有建置自己的機器翻譯系統，供各國審查委員與申請人運用。使用機器翻譯的底線是誠實，產出的文件上應載明是機器譯文，提供文件與接收文件的雙方都有共識，接受機器譯文可能有瑕疵或不夠精確。本文首先介紹一些常見的機器誤譯態樣，分析造成機器翻譯的誤譯的因素，這些誤譯反映出嚴謹撰稿的原文之重要性，在撰寫專利說明書時尤其如此。

二、常見的機器誤譯

透過機器翻譯來理解巨量的各國專利前案，已是各大專利局普遍接受的實務；機器翻譯的失誤在所難免，以中英文互譯的專利文件為例，常見的機器誤譯有以下幾種態樣：(1) 名詞被誤譯為動詞；(2) 專業術語被誤譯為普通名詞；(3) 詞彙重複；(4) 同音／近音異義的誤譯；(5) 原文筆誤所致；(6) 因口語化表達而誤譯；(7) 因無縮寫而誤譯；以及(8) 因縮寫錯誤而誤譯（詳見參考資料一）。這些機器翻譯的失誤，可以大致分析出幾種原因，最單純的是原文筆誤導致機器翻譯「忠實」呈現地將筆誤轉換成另個語言，如上述態樣(4)，「陣列控制」的簡稱「陣控」打字失誤為「真空」，以至於被機器翻譯成 vacuum，然正確應為 array control；或是機器翻譯未能正確配合上下文脈絡，僅以名詞的字面意義來解讀，如上述態樣(7)，以科學家為名的計量單位如 watt，若無提示其縮寫並留意字母大小寫問題，機器翻譯解讀時可能在人名與功率單位之間做出誤判；有些誤譯源自於未能對應特定專業領域的習慣用語，如上述態樣(2)，「後台」譯為 background，但在資訊科技的領域中，配合當時上下文的脈絡，應譯為 server。理想上，稱職敬業的真人譯者，精通該專業技術領域、語文造詣佳，搭配翻譯後的審稿與校對，可以避免上述機器翻譯的錯誤，這需要一個工作團隊，而非單靠優秀的譯者，也不易應用在專利前案檢索階段，因為這階段對閱讀文件的需求是過濾篩選巨量資料後再細讀；機器翻譯的技術發展至今，運用的範圍可逐漸被視為真人的潛在共事者。

三、嚴謹的原文有益於機器翻譯

機器翻譯的普及，凸顯出真人作者在撰寫原文時，正確且清楚明確的重要性。一些寫作上的筆誤或語法上不嚴謹，真人讀者或許可以領會，而在理解的過程中，自行校正回歸；但有失誤或語意不清的原文，由機器翻譯處理時，這些瑕疵或模糊之處可能導致天差地遠的譯文。舉例來說：有些撰稿人的寫作的風格帶有口語化傾向，在寫作中省略冠詞、介詞、代名詞，但這些字詞在機器翻譯的演算法中，常常是用來定位，以拆解句子的標記，隨興的省略可能導致機器翻譯的解讀機制做出誤判；標點符號也變得舉足輕重，若沒有完整的文法結構，思緒告一小段落便隨手來個句號，對機器翻譯的運算會產生干擾；至於俚語或反諷，試圖傳遞與字面意義相反的言外之意，透過機器翻譯後更不易重現。簡言之，在寫作階段，可將日後會遇到的機器翻譯設想為一位冷酷的審稿人：沒有幽默感、沒有同理心，無法體諒人都難免有失誤，小瑕疵一律放大檢視；撰稿時必須嚴格遵守文法規則與力求語意明確，同時減少華麗堆疊或主觀的比喻或形容，才有助於機器翻譯正確解讀譯出。

一些通用性的寫作原則，寫作修辭相關教材裡常提的：在正確性方面，必須嚴格遵守文法規則、正統的拼字與標點符號用法；為使文章清楚易理解，應該消除贅字，少用雙重

* 任職台一國際智慧財產事務所專利國內部



否定，句法上主詞與動詞避免相隔太遠，減少在主要句子的中間插入子句。近代人日常使用語言日趨隨興或喜自創慣例，但在機器翻譯的科技興起之後，這些經典的寫作圭臬在原文產出的源頭上反而變得更重要：遵守正規的拼字與文法，減少機器翻譯誤判；簡單直接的句型，讓機器翻譯更易拆解句子。

四、撰寫專利說明書，明確性更顯重要

專利說明書之撰寫，除了技術內涵的新穎性與進步性之外，明確性也常影響權利範圍的解讀；直接從寫作衍生出來的明確性議題，最常見的是元件是否具備前置基礎，此時冠詞、不定冠詞、單複數是否一致、元件名稱前面是否有附帶形容用語或量詞都可能導致前置基礎的認定不明確。形容用語的使用也需謹慎，例如在請求項中，「高」溫「低」溫都是相對性的形容用語，進一步定義溫度的具體數值範圍才可據以實施。此外術語一開始必須先寫出全稱，定義其縮寫之後，後續使用該縮寫才能避免權利範圍產生不明確之疑慮，即使有些縮寫在該領域已普及而全名反而較少使用，但在撰寫請求項時，如果沒有先定義 light emitting diode 並載明其縮寫為 LED，而是直接使用縮寫 LED，被以定義不明確為由而成為不予專利的事由，在實務上並非罕見。專利說明書在撰寫階段，寫作形式上已較其他性質的寫作更一板一眼，出發點是為了避免審查階段與後續權利行使時不明確的疑慮，而在機器翻譯已廣泛運用於專利前案檢索的情況下，這些嚴謹的撰稿習慣，亦有助於減少日後機器翻譯誤譯的機會，更彰顯出優秀的原文寫作的重要性。

參考資料：

1. Cheng Ying, Yue Shuyu, Li Jing, Deng Lin, and Quan Qi. "Errors of Machine Translation of Terminology in the Patent Text from English into Chinese." *ASP Transactions on Computers*. Vol. 1. No.1, May 2021.
2. Ken Clark. "Elements of Style for Machine Translation." *Writing for Translation. MultiLingual*. October/November 2009.