出國報告（出國類別：進修）

明德大學蒙特瑞國際研究學院
口筆譯短修班心得報告

服務機關：國防部心戰大隊
姓名職稱：少校副中隊長 黃瑞青
派赴國家/地區：美國
出國期間：2019. 6. 24 ~ 2020. 5. 15
報告日期：2020. 6. 15

摘要

本人於 2019 年 6 月 24 日至 2020 年 5 月 15 日期間，赴美國加州明德大學蒙特瑞國際研究學院（Middlebury Institute of International Studies at Monterey）參與口、筆譯短修班，本報告內容含括進修目的、進修過程、學習心得與建議事項，另於報告後面附上個人在學期間的筆譯專案作品，期能藉分享課程內容與重點、新冠病毒疫情期間的教學概況，以達進修資源與成果分享之目的，提供對蒙特瑞口、筆譯短修班課程有興趣的同仁參考。

目次

壹、 進修目的

有鑒於近年華美交流及互動日益頻繁，且單位任務亟需新聞翻譯人才，因有感於自身專業學能尚不足以應付工作所需，因此報名口筆譯短修班隊，期能在口譯及筆譯等翻譯領域有所進步，並藉在美國進修之際，瞭解美式生活與價值觀及文化，文化與歷史背景是所有語言的知識底蘊，期望透過為期 11 個月的全英語環境改善自身英語不足之處，並能在回國後貢獻所學，滿足單位與外軍交流之需要。

貳、 課程簡介及過程

明德大學蒙特瑞國際研究學院（Middlebury Institute of International Studies at Monterey, MIIS）位於加州中部的蒙特瑞郡（Monterey County），與美國海軍研究院（Naval Postgraduate School, NPS）、國防語文中心比鄰（Defense Language Institute, DLI），該校以口、筆譯相關科系於國際享譽盛名，為聯合國及各國政府、跨國企業等各領域培育不少口筆譯人才。

MIIS 除了提供口、筆譯碩士學程外，另亦廣設語言培訓短期班，本人所參與的班隊為國防部與該校合作的口筆譯短修班（Custom Academic-Year Program），該班學員共 5 名臺灣軍官，訓期自 2019 年 6 月 24 日起，至 2020 年 5 月 15 日止，為期將近一年，課程主軸以中英雙向的口譯、筆譯及英語能力為主，課程選材多半為國防、政治、外交、國際關係等相關議題，搭配經濟、文化、時事內容為輔。除了上、下學期的口、筆譯教程之外，在上學期開學前七週，校方亦安排學前英語準備課程，協助國際學生在進入美國高等教育體制前預做暖身。上述課程詳細介紹如下：

1. **研究所學前英語準備課程(English Preparation for Graduate Studies, EPGS)**：

   課程長達七週， 授課對象為國際學生，全班學生約莫 20 名，來自臺灣，中國大陸、日本、韓國、越南、烏克蘭等國家，該課程目的在讓國際學生於研究所正式學程前加強英語聽、說、讀、寫技巧，以應研究所課程之需。課程包含「學術研究寫作（Academic Writing）」、「聽力策略（Active Listening Skills）」、「閱讀策略（Reading Strategies）」、「簡報技巧（Presentation Skills）」等課目，除前述既定課程外，每週五另安排主題工作坊，內容涵括學校政策及當地文化、壓力調適、假訊息防範等。課程專為國際學生設計，在英文聽、說、讀、寫及文化適應上助益甚大。

2. **基礎口、筆譯：**一學年的基礎口筆譯課程，細分為「逐步口譯」、「視譯」、「筆譯」，以上均為中進英、英進中分開授課，授課師資應國防部要求均為臺灣籍教師，課程主旨在讓學生理解口筆譯的基礎概念與技巧，另因應國防部專案短修班需求，課程內容選材以國防、政治、外交為主，經濟、文化等相關時事為輔，並透過課堂上教師的安排與設計、課後練習與討論等方式，將理論、技巧實際運用操作。
3. **發音：**美籍老師授課，從美式自然發音為基礎，矯正臺灣學生在發音上慣有的錯誤，另在沒有音標的輔助下，能夠依循拼字的基本原則，正確發音，課程目的力求學生在口譯、專報或演說時口音清晰，明確傳達口語表達內容。
4. **英文技巧：**美籍老師授課，除分析單詞與文章結構、口語英語表達、英文簡報技巧、文法解構、俚語運用之外，還藉由時事專題（川普彈劾案、新冠肺炎等）深入解說美國歷史文化對語言的影響，此門課引領學生從「美國人」的視角重新審視如何在口筆譯工作上更精確的用字遣詞。

參、 學習心得

一、 會英文不等於擅長口譯

還未接受口譯正規訓練前，偶爾會因美方交流的案子被指派擔任口譯員，當時對口譯的認知僅止於兩種語言轉換的橋樑，因此只要說得出話就好了，就算不精確也無妨，別讓場子冷了就好！我並非唯一有如此想法的人，大多數長官與同僚都是如此看待口譯工作，外事交流任務只要找個會英文、講話別結結巴巴的就好，譯文品質如何自然已不是首要考量了。最常看到的狀況是：口譯員與講者各說各話、口譯僅傳達講者的一半內容、講者一段話只有 1 分鐘，但口譯花了 5 分鐘解釋、聽不懂口譯員的中文表達。

一年的口筆譯短修班顛覆了我的粗淺愚見。口譯需要的不只是語言能力，還需要翻譯技巧、分析判斷、專業知識等，這四樣能力缺一不可，語言能力是一切基礎，對 A、B 語言的掌握程度越高，對口譯或筆譯當然助益越多，然而中文與英文兩者的句構大不相同，中文強調意合，詞語之間的關聯即可形成意義；而英文卻重視形合，詞語之間的位置決定其意義，因此兩個語言的轉換需要靠翻譯技巧與分析判斷，才能在時間壓力下準確地處理訊息，當以上條件都具備後，專業知識就是最後那一筆畫龍點睛了，口譯員需要面對的主題包羅萬象，軍中的口譯員也不例外，軍事領域專業分工細緻，專業術語與知識隔行如隔山，口譯員往往是會議中最無知的人，但是卻擔負著處理最龐大訊息的任務，因此若沒有事前的充分準備，是無法僅靠本身的語言能力或是臨場反應補足專業知識的空缺。

二、 理解原文而非照字翻譯

中文就如同落葉堆，訊息是由不同種類詞語間的關係所產生意義，因此主詞的形式相當豐富、句構鬆散但調動彈性大；反之，英文則像是長在

枝頭的葉子，訊息是由樹枝的生長方向與架構來決定詞語間的意義，主詞往往會決定後面動詞跟整個句子的發展，句構嚴謹。這兩種語言之間的差異性如此之大，增加了口、筆譯的挑戰性，尤其是口譯，時間壓力的催促下，很容易被字面上的意思所綁住，就容易照著中文句構字對字翻譯，英文母語人士很難理解產出的譯文。

因此「理解」訊息是處理訊息的第一步，也是最重要的一個步驟，要聽的是訊息而不是語言文字，經過大腦分析訊息之間的關聯、判斷訊息間的主從與重要性，再來才是記憶訊息，轉換成 B 語言的語序及結構，接著產生譯文，最後還要監聽自己所產生的譯文，咬字是否清晰，前後文是否符合邏輯，台風是否穩健。整個過程必須在彈指間就在大腦中全部走完一遍，相當考驗譯者對中、英文的理解力與反應，絕對不是只有英文好就可以辦到的事。

三、 事前準備避免當場出糗

「口譯員往往是全場專業知識最欠缺的一人」每個授課教師都這麼說。老師們都在業界有過豐富實戰經驗，他們往往做一場 3 小時的口譯，事前的準備要好幾天，若涉及科技、醫療、資訊等相關產業或不熟悉的主題，往往準備更加耗時。試想，講者光用中文發表醫學新發現，演講中必然會出現許多醫療專業用語，對中文母語人士來說就已是艱澀難懂了，更遑論要譯成另一個語言，因為第一步驟「理解」就做不到了。

事前大量閱讀與會議主題相關的平行文本，蒐集專業術語的字彙與表達方式，也同時對該主題與週邊相關議題有基本認識；研究講者的背景，包括說話習慣、演講模式、特殊口音，甚至是講者的成長背景、工作經驗等，這些都有助於瞭解與認識講者，能夠幫助譯者預判講者的演說內容，減少大腦在處理新資訊所需耗費的能量，大腦才能騰出空間處理更複雜的訊息。

四、 保持好奇心與獨立思考

老師在課堂中的教材內容多半與國際時事有關，多虧平時有閱讀國際新聞的習慣，的確在口譯時能幫助我快速進入主題，更能快速分析講者或作者的思維脈絡。譯者無法左右講者的話語及思維，為了應對講者可能的天外飛來一筆，除了事前針對主題多做準備之外，還需仰賴平時涉足多元領域，雖無法樣樣精通，但廣泛吸取的新知都會成為未來工作時備用的資料庫，為了應對未知，所以要不斷地求取新知，因為永遠無法預知未來會接到什麼任務，只能多聽、多看、多接觸各領域的知識，指不定哪天就能用上了！

五、 視訊遠距教學利弊參半

下學期的課程因為受新冠病毒疫情的影響，校方為避免校園群聚感染暫時關閉校園，自三月起到學期結束的課程全部改為線上的遠距教學。學校採用的視訊會議軟體為 Zoom，該軟體因中方技術與資金一度遭各界質疑其資訊安全性，撇除其資訊安全性不談，Zoom 很適合多人線上會議與教學使用，螢幕畫面可同步分享所有與會者，操作介面簡單，容易上手，的確在疫情初期為教育界解了燃眉之急。

然而遠距教學的缺點也隨著疫情持續延燒，慢慢浮現。首先是網路頻寬，除了教育界，許多企業也因應疫情改為在家辦公，因此白天網路因多人使用而流量龐大，且絕大多數學校及公司都使用同一個線上視訊會議軟體，造成網路大塞車，視訊畫面不斷延遲，甚至斷線，嚴重影響教學品質，逐步口譯課程影響甚鉅。

再者，眼神交流是雙向或多向對話不可或缺的要素。受限於電腦視訊鏡頭無法追蹤人類的眼球轉動而跟著移動，視訊畫面上發言者的眼神往往無法與接收者對焦，少了眼神的交會，確實在接收語言訊息上需要耗費更多的精力在專注聆聽上，往往一堂 2 小時的課卻有連續上了 4 小時的疲憊

感，且上課期間很容易走神。課程後期我曾嘗試不看視訊畫面，僅專注聆聽上課內容，發現比同時看畫面又要聽內容，反而更容易接受訊息，這不禁讓我聯想到做會議口譯時，同時要兼顧簡報畫面又要聆聽講者訊息，還要即時產出譯文，多工處理的確對大腦來說是個負擔，尤其在簡報畫面凌亂無法提供有用訊息時，只會徒增困擾。

視訊教學亦會降低學生在課堂上的互動次數。由於缺少眼神交流與實際互動，比起實體教室上課，視訊教學給予學生更多分神的機會，無法專注於課程時，自然參與感降低；另一個原因是 Zoom 一次僅能讓單一個人發言，若同時兩人以上說話，僅會出現一人的聲音，因此每當我想說話，又會顧慮著其他同學是否也想發言，心中一有遲疑就往往不會那麼主動回應。

六、 校方新冠病毒應變得宜

早在新冠病毒疫情發展初期，聯邦政府與州政府尚未有任何處置之時，校方已預判三月春假期間是學生與教職員跨州、跨國移動或旅行最頻繁之際，開學返校後勢必會引發群聚傳染的風險，因此在三月初春假前毅然下決定，爾後課程全部改為線上教學，且校園全面關閉，校方透過電子郵件與線上教學，詳細地指引學生及老師如何操作線上教學軟體，並由校方高層主持視訊會議，邀請學生參與且提出相關建議或問題與校方討論，另外並提供心理諮商服務，定期關心學生在居家防疫期間的心緒狀況。校方相當重視每個學生受教的權益，當我提出租屋處並無提供網路服務，恐無法參與線上教學時，行政職員馬上協助我安裝無線網路，並且由校方支應網路月租費。

肆、 建議事項

一、 廣泛培育傳譯種能

鑑於國軍並未設置專責口筆譯之建制單位，但外事交流機會愈趨頻繁，每遇外軍交流場合多半臨時指派或是從其他單位調借人員支援，然而支援傳譯任務之人員多半是臨危授命，準備時間不充分、對受支援單位任務及裝備不瞭解，這些都是對傳譯人員極大的負擔，影響其擔任外事溝通橋樑的表現與成效，若能維持或增加國內、外的口筆譯短修班的進修員額，將傳譯的種能厚植於各司令部與各聯兵旅級單位，將有助於各層級遂行外事交流任務。

二、 落實為用而訓理念

外國語言一旦太久不接觸、不使用，很快就會生疏，為了使國家資源以及學員個人所投注在語言學習上的精力與時間不會船過水無痕，建議結訓後能妥善規劃派職，讓所學能發揮效用。囿於國軍大多數單位均無編設翻譯相關職缺，建議管制並編組語言專才人員，建立常態性支援模式，於外賓來訪、演訓觀摩交流、會議研討等時機，妥善運用國軍傳譯資源。

伍、 筆譯專案作品

TRANSLATION PROJECT

May 2020

Jui-Ching Huang

Table of Content

Introduction

Source Text:

1. Andrew Lohn, **What Chess Can Teach Us about The Future of AI and War**, War on The Rocks, January 3, 2022. (about 1,900 words) (https://warontherocks.com/2020/01/what-chess-can-teach-us-about-the-future-of-ai-and-war/)
2. John Villasenor, **Artificial Intelligence, Geopolitics, and Information Integrity**, The Brookings Institution, November 2019. (about 3,600 words) (Links: https://www.brookings.edu/research/artificial-intelligence-geopolitics-and-information-integrity/)

Purpose of the Translation Project:

Artificial Intelligence has become a popular and advanced technology both in areas of commercial and military. In defense technology, it has been applied into radar, UAV, command and control system, reconnaissance, etc. Besides these developments, it is also widely seen in cyberwarfare. States or non-state actors utilize AI to amplify their influence on the target audience and achieve the desired results. 2016 US presidential election is the best evident. AI produced information becomes a serious issue among governments and military forces. My goal of this translation project is to provide opinions from experts in US well-known think tanks, the Brookings Institution and Rand.

Weekly Plan:

Week	Date	Activities	Note
3	Tue, Feb 11	Submit project proposal	
4	Mon, Feb 17	1st part translation (800 words/ week)	
5	Mon, Feb 24	2nd part translation	
6	Mon, Mar 2	3rd part translation	
7	Mon, Mar 9	4th part translation	
8, 9	Mar 13- 22	Spring break	
10	Mon, Mar 30	5th part	
11	Mon, Apr 6	6th part	
12	Mon, Apr 13	7th part	
13	Mon, Apr 20	Proofreading & Editing	
14	Mon, Apr 27	Readers' reviews & comments	
15	Mon, May 4	Self-commentary	
16	Tue, May 12	Submit the portfolio	

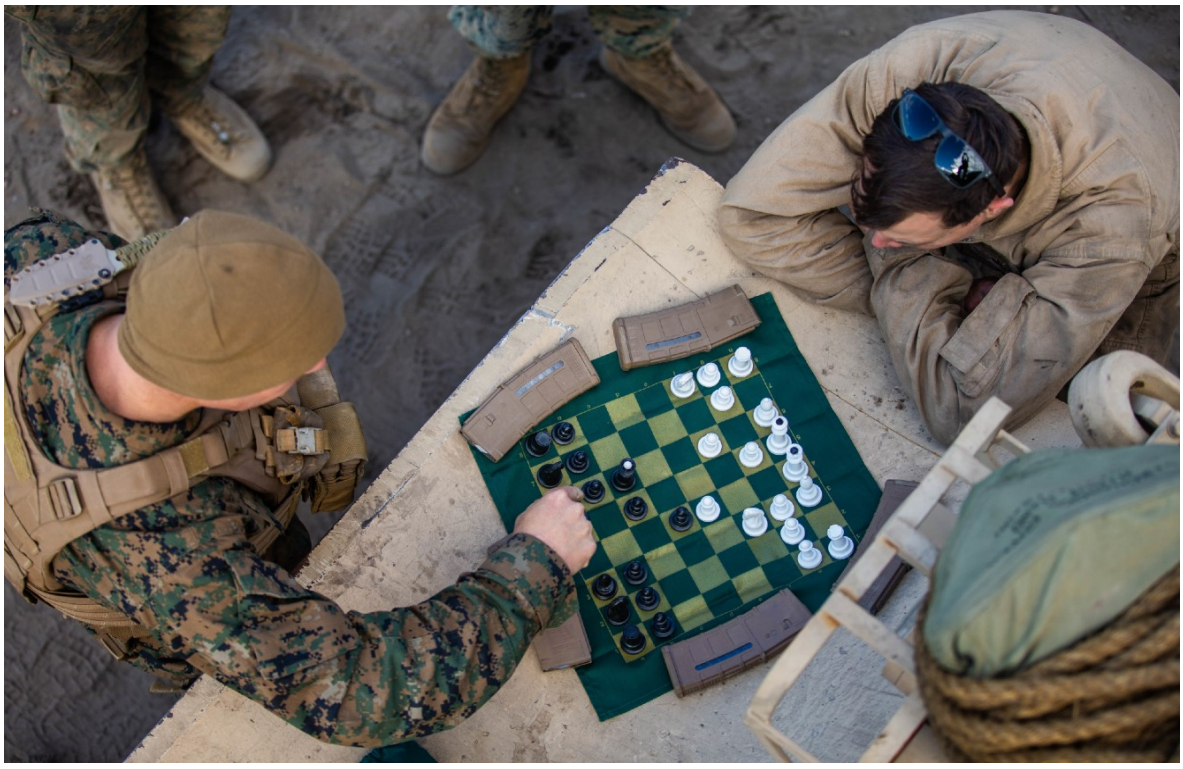
Source Text 1 : What Chess Can Teach Us about The Future of AI And War

WAR ON THE ROCKS

WHAT CHESS CAN TEACH US ABOUT THE FUTURE OF AI AND WAR

ANDREW LOHN JANUARY 3, 2020

SPECIAL SERIES - AI AND NATIONAL SECURITY



This article was submitted in response to the call for ideas issued by the co-chairs of the National Security Commission on Artificial Intelligence, Eric Schmidt and Robert Work. It addresses the first question (part a.), which asks how will artificial intelligence affect the character and/or the nature of war.

Will artificial intelligence (AI) change warfare? It's hard to say. AI itself is not new — the first AI neural network was designed in 1943. But AI as a critical factor in competitions is relatively novel and, as a result, there's not much data to draw from. However, the data that does exist is striking. Perhaps the most interesting examples

are in the world of chess. The game has been teaching military strategists the ways of war for hundreds of years and has been a testbed for AI development for decades.

Military officials have been paying attention. Deputy Defense Secretary Robert Work famously used freestyle (or Centaur) chess to promote the third offset strategy, where humans and computers work together, combining human strategy and computer speed to eliminate blunders while allowing humans to focus on the big picture. Since then, AI and supercomputers have continued to reshape how chess is played. Technology has helped to level the playing field — the side with the weaker starting position is no longer at such a disadvantage. Likewise, intimidation from the threat of superhuman computers has occasionally led to some unorthodox behaviors, even in human-only matches.

The experience of AI in the chess world should be instructive for defense strategists. As AI enters combat, it will first be used just in training and in identifying mistakes before they are made. Next, improvements will make it a legitimate teammate, and — if it advances to superhuman ability in even narrow domains of warfighting, as it has in chess — then it could steer combat in directions that are unpredictable for both humans and machines.

What Does Chess Say About AI-Human Interaction?

Will AI replace soldiers in war? The experience of using AI and machine learning in chess suggests not. Even though the best chess today is played by computers alone, humans remain the focus of the chess world. The world computer chess championship at the International Conference on Machine Learning in Stockholm attracted a crowd of only three when I strolled by last year. In contrast, the human championship was streamed around the globe to millions. In human-only chess though, AI features heavily in the planning process, the results of which are called “prep.” Militaries are anticipating a similar planning role for AI, and even automated systems without humans rely on a planning process to provide “prep” for the machines. The shift toward AI for that process will affect how wars are fought.

To start, computers are likely to have an equalizing effect on combat as they have had in chess. The difference in ability among the top competitors in chess has grown smaller, and the advantage of moving first has become less advantageous. That was evident in last year’s human-only chess championship where competitors had the closest ratings ever in a championship, and the best-of-12 match had 12 straight draws for the first time. There have been more draws than wins in every championship since 2005, and though it is not exactly known why, many believe it is due to the influence of superhuman computers aiding underdogs, teaching defensive play, or simply perfecting the game.

AI is likely to level the military playing field because progress is being driven by commercial industry and academia — which will likely disseminate their developments more widely than militaries. That does not guarantee all militaries will benefit equally. Perhaps some countries could have better computers or will be able to pay for more of them, or have superior data to train with. But the open nature of computing resources makes cutting-edge technology available to all, even if that is not the only reason for equalization.

AI Favors the Underdog and Increases Uncertainty

AI seems to confer a distinct benefit to the underdog. In chess, black goes second and is at a significant disadvantage as a result. Fabiano Caruana, a well-known American chess player, claimed that computers are benefiting black. He added that computer analysis helps reveal many playable variations and moves that were once considered dubious or unplayable. In a military context, the ways to exert an advantage can be relatively obvious, but AI planning tools could be adept at searching and evaluating the large space of possible courses of action for the weaker side. This would be an unwelcome change for the United States, which has benefited from many years of military superiority.

Other theories exist for explaining the underdog's improvement in chess. It may be that computers are simply driving chess toward its optimum outcome, which some argue is a tie. In war it could instead be that perfect play leads to victory rather than a draw. Unlike chess, the competitors are not constrained to the same pieces or set of moves. Then again, in a limited war where mass destruction is o# the table, both sides aim to impose their will while restricting their own pieces and moves. If perfect play in managing escalation does lead to stalemate, then AI-enhanced planning or decision-making could drive toward that outcome.

However, superhuman computers do not always drive humans toward perfect play and can in fact drive them away from it. This happened in a bizarre turn in last year's chess world championship, held in London. The "Queen's Gambit Declined," one of the most famous openings that players memorize, was used to kick o# the second of the 12 games in the London match, but on the tenth move, the challenger, Caruana, playing as black, didn't choose either of the standard next moves in the progression. During planning, his computers helped him find a move that past centuries had all but ignored. When the champion Magnus Carlsen, who is now the highest-rated player in history, was asked how he felt upon seeing the move, he recounted being so worried that his actual response can't be reproduced here.

It is not so much that Caruana had found a new move that was stronger than the standard options. In fact, it may have even been weaker. But it rattled Carlsen because, as he said, "The difference now is that I'm facing not only the analytical team of Fabiano himself and his helpers but also his computer help. That makes the situation quite a bit different." Carlsen suddenly found himself in a theater without the aid of electrical devices, having only his analytical might against what had become essentially a superhuman computer opponent.

His response might presage things to come in warfare. The strongest moves available to Carlsen were ones that the computer would have certainly analyzed and his challenger would have prepared for. Therefore, Carlsen's best options were either ones that were certainly safe or ones that were strange enough that they would not have been studied by the computer.

When asked afterward if he had considered a relatively obvious option that he didn't chose seven moves later in the game, Carlsen joked that "Yeah, I have some instincts … I figured that [Caruana] was still in prep and that was the perfect combination." Fear of the computer drove the champion, arguably history's best chess player, to forego a move that appeared to be the perfect combination in favor of a safer defensive position, a wise move if Caruana was in fact still in prep.

In war, there will be many options for avoiding the superhuman computing abilities of an adversary. A combatant without the aid of advanced technology may choose to

withdraw or retreat upon observing the adversary doing something unexpected. Alternatively, the outcomputed combatant might drive the conflict toward unforeseen situations where data is limited or does not exist, so as to nullify the role of the computer. That increases uncertainty for everyone involved.

How Will the U.S. Military Fare in a Future AI World?

The advantage may not always go the competitor with the most conventional capabilities or even the one that has made the most computing investment. Imagine the United States fighting against an adversary that can jam or otherwise interfere with communications to those supercomputers. Warfighters may !nd themselves, like Carlsen, in a theater without the aid of their powerful AI, up against the full analytical might of the adversary and their team of computers. Any unexpected action taken by the adversary at that point (e.g., repositioning their ground troops or launching missile strikes against unlikely locations) would be cause for panic. The natural assumption would be that adversary computers found a superior course of action that had accounted for the most likely American responses many moves into the future. The best options then, from the U.S. perspective, become those that are either extremely cautious, or those that are so unpredictable that they would not have been accounted for by either side.

AI-enabled computers might be an equalizer to help underdogs find new playable options. However, this isn't the only lesson that chess can teach us about the impact of AI-enabled supercomputers and war. For now, while humans still dominate strategy, there will still be times where the computer provides advantages in speed or in avoiding blunders. When the computer overmatch becomes significant and apparent, though, strange behaviors should be expected from the humans.

Ideally, humans deprived of their computer assistants would retreat or switch to safe and conservative decisions only. But the rules of war are not as strict as the rules of chess. If an enemy turns out to be someone aided by feckless computers, instead of superhuman computers aided by feckless humans, it may be wise to anticipate more inventive — perhaps even reckless — human behavior.

*Andrew Lohn is a senior information scientist at the nonpro!t, nonpartisan RAND Corporation. His research topics have included military applications of AI and machine learning. He is also co-author of "*How Might Arti!cial Intelligence A#ect the Risk of Nuclear War?*" (RAND, 2018).*

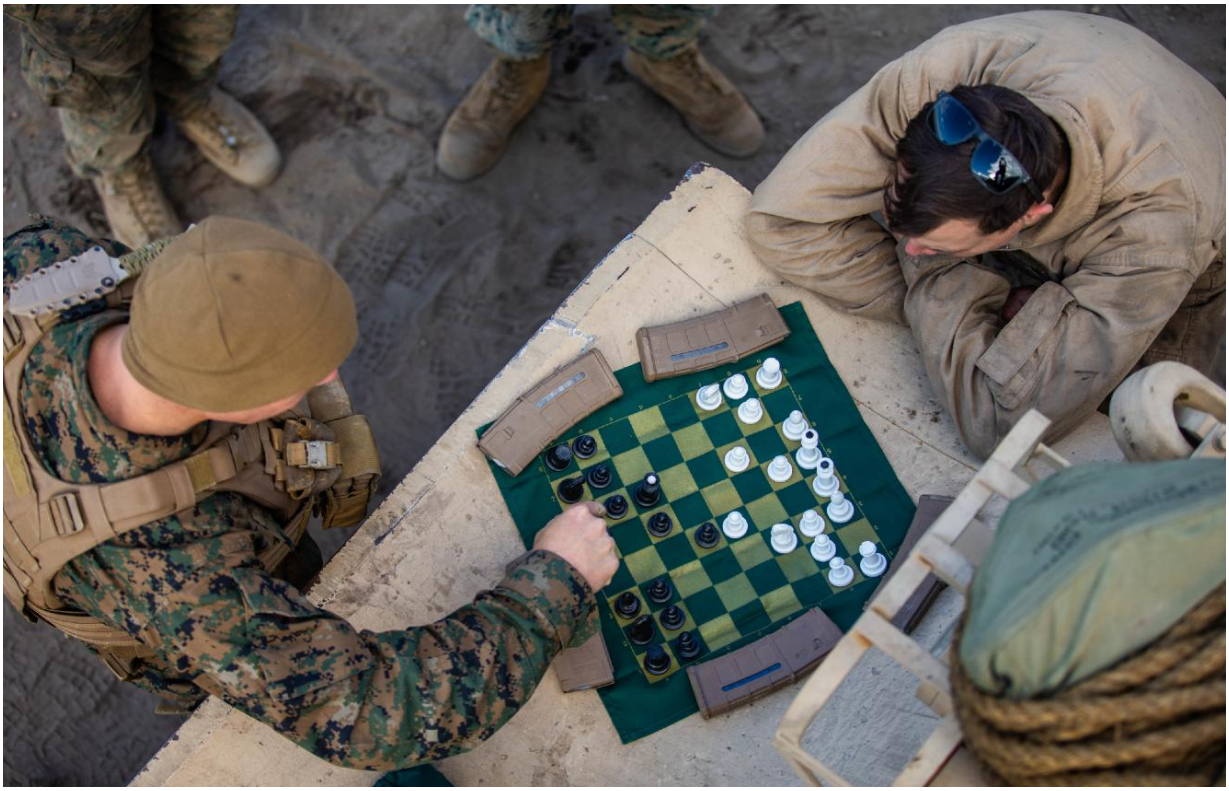
TRANSLATION: 從西洋棋看人工智慧與戰爭的未來趨勢

WAR ON THE ROCKS

從西洋棋看人工智慧與戰爭的未來趨勢

ANDREW LOHN

2020 年 1 月 3 日



本篇文章是應國家安全委員會人工智慧的共同主席艾瑞克·施密特（Eric Schmidt）和羅伯特·沃克（Robert Work）的要求而撰寫。本文第一個問題即是：人工智慧將如何影響戰爭的特性及（或）本質。

* * *

人工智慧（AI）會改變戰爭型態嗎？很難說。AI 本身並非新創科技，首個 AI 神經網路設計始於 1943 年，然而，AI 在競賽的場合中扮演關鍵角色依舊是相對新穎的趨勢，因此目前雖沒有太多的文獻可以借鑒，但是現有的文獻資料著實令人大開眼界。最有意思的例子應屬西洋棋，數百年來，西洋棋一直是軍事戰略家師法的對象；最近數十年來，西洋棋也一直是人工智慧發展的試驗場。

軍方官員相當關注人工智慧發展。國防部副部長羅伯特·沃克利用自由式（或混合式）西洋棋來提倡第三次抵銷戰略，即人類和電腦協同合作，結合人類策略和電腦速度來排除誤失，同時讓人類將專注力放在大局上。自此後，AI 和超級電腦改寫了西洋棋的世界。科技的輔助讓棋賽打成了和局——起步位置較差的一方不再居於劣勢。同樣，面對超級電腦的威脅恐嚇時，人類偶爾會出現一些異常的行為，甚至在沒有超級電腦參與的棋賽中也會出現異常行為。

AI 在西洋棋界的經驗應該對軍事戰略家具有啟發性。當 AI 運用在戰鬥領域時，初期僅用於訓練，以及在出錯前及時揪出問題。後續接著提升人工智慧性能，將它納入戰鬥隊伍，假若人工智慧能在戰爭如此狹小的領域（如同西洋棋）中表現超出常人的能力，那麼未來的戰爭走向，人類和機器都無法預測。

從西洋棋看人工智慧與人類的互動

AI 會取代戰場上的士兵嗎？從西洋棋中使用 AI 和機器學習的經驗可回答此問題：無法取而代之。儘管當今最高段的棋手是電腦單獨對弈，但人類棋手仍然是西洋棋界的焦點。去年，在斯德哥爾摩的國際機器學習大會上舉行了一場世界電腦西洋棋錦標賽，只吸引了三位觀眾。相比之下，人類對弈的錦標賽卻風靡全球數百萬人。其實，在人類對弈的西洋棋賽事中，AI 在計劃準備過程中佔有一席之地，稱之為「準備（prep）」。軍隊也對 AI 類似計劃的能力抱予期望，甚至無人自動化系統也需靠計劃程序，為機器提供「準備」。計劃準備的任務由人類手中過渡給 AI 執行，將會影響未來戰爭型態。

首先，電腦在棋局上所發揮的均化效應可能也適用於戰場上。西洋棋頂尖棋手能力不相

上下，以往開局時先手的優勢也愈不顯著了。這在去年的人類西洋棋錦標賽上已顯而易見，參賽選手的等級分差距是有史以來最小的，首次在 12 強賽出現 12 場平局。自 2005 年以來，每屆錦標賽的平局都比勝局多，雖並不十分清楚原因，但許多人認為，這是因爲有超級電腦輔助弱者並傳授守勢策略或讓棋局更完美無缺。

AI 之所以能讓軍事競賽中的參賽者都能在同一起跑線上公平競爭，歸因於商界和學術界正推著 AI 不斷向上發展，AI 在軍事以外的領域應有更廣泛的發揮。但這並不保證所有國家的軍隊都能享受到 AI 帶來的優勢，也許僅某些國家的電腦性能較優越，亦或者有財力購買更多的電腦，或是擁有高品質的資料供予 AI 學習。但是，電腦計算資源的本質即是開放性，因此人人都能取得尖端科技，即使這並非是造成均化的唯一原因。

AI 偏愛弱者，增加不確定性

AI 似乎為位居下風者帶來明顯的好處。在西洋棋中，執黑的一方須待白方開局走第一步後才能行棋，因此執黑者大大地處于劣勢。美國著名棋手法比亞諾·卡魯阿納聲稱，電腦有利於黑方扭轉劣勢。他補充說，電腦分析有助於揭示許多棋路，包括以往被視為勝算不高或死棋的棋路。在軍事領域，對上風者來說，發揮優勢的方法相對清楚可見，然而，人工智慧精於搜尋和評估，可為劣勢者提供大量的行動方案。這將是美國不樂見的情勢，美國已經在軍事領域佔上風好幾年了。

尚有其他理論可解釋劣勢者在棋局中如何迎頭趕上，電腦可能僅是將棋局導向最佳結果，意即部分人所認為的雙方平手。在戰場上的完美發揮，帶來的是勝利而非平局。戰場與西洋棋不同的地方在於參賽者不須受限於棋盤規則。在不涉及大規模毀滅性武器的有限戰爭中，雙方皆欲展現其必勝的決心，但同時也須步步為營，若掌控戰場情勢升溫必然會導致陷入僵局，那麼在 AI 加持下的軍事計畫與決策過程，恐怕兩軍交戰結果也是平局。

然而，超級電腦並非總能規劃出無懈可擊的棋局，去年在倫敦舉行的西洋棋世界錦標賽就上演了這麼一場戲劇性的轉折。在倫敦的這場賽事中，棋手們記得第二盤棋用了著名的開局方式之一『拒后翼棄兵（Queen’s Gambit Declined）』，但執黑的挑戰者卡魯阿納在第十步並未按牌理出牌，他的電腦在計畫準備期間幫他找到已存在幾世紀卻被忽視的棋路。史上積分最高的棋王馬格努斯・卡爾森（Magnus Carlsen）被問到看到對手這步棋時

當下的感受，他回憶道，現在已無法重現當時的憂心忡忡了。

與其說是卡魯阿納發現了一個超乎常規且更強大的新招，其實這步棋也許沒什麼過人之處，但是它讓卡爾森亂了陣腳，因為正如他所說：「如今我面對的不僅是卡魯阿納的分析團隊和他的助手，還要面對他的電腦，這使得情況相當不同於以往。」 卡爾森頓時發現自己身處在沒有電子設備輔助的戰場，只能靠他的分析能力來對付本質上已成了超人類電腦的對手。

他的回應可能預示著未來戰爭的型態。電腦肯定會分析出卡爾森最強的棋路與招式，而他的挑戰者也是有備而來。因此，卡爾森若不是選擇打安全牌，就只能出奇招，讓電腦摸不清他的底。

當卡爾森事後被問到是否考慮過另一個相對明顯的選項，但他當時並未選擇走那七步棋，卡爾森開玩笑說：「是啊，我的直覺告訴我...卡魯阿納還在準備階段，這真是絕佳組合。」由於對電腦的畏懼驅使這位冠軍，公認為史上最好的西洋棋手，放棄看似完美的一步棋，轉而採取較安全的守勢位置，若當時卡魯阿納確實還在準備中，卡爾森無疑是明智之舉。

在戰爭中，有許多方式可閃避敵方的超級電腦運算能力。沒有高科技輔助的戰鬥人員一旦觀察到對手有一些意料外的舉動時，他可能選擇撤退；又或者，沒有電腦輔助的戰鬥員會將衝突引導至數據有限或根本不存在的未知方向，從而抵銷敵方電腦的算計，在這種情況下，會為所有涉事者增添不確定性。

在未來的 AI 世界中，美軍將如何發展?

戰場優勢並不總是掌握在擁有最多傳統戰力的競爭者手上，也不見得是掌握在投資最多電腦運算的競爭者手上。試想一下，美國交戰的對手具備干擾與這些超級電腦通聯的能力，作戰人員可能會覺得自己的處境與卡爾森雷同，身處在沒有強大 AI 協助的戰場，隻身對抗敵方的全面分析及其電腦團隊。敵手當下採取的任何出奇之舉（例如：重新部署地面部隊，或對不太可能的地點發動飛彈攻擊）都會引起恐慌，作戰人員自然而然地會推測是敵方的電腦找到了優越的行動方案，這也被視作是美國未來最有可能做出的反應。從美國的角度來看，最佳的選項若非極其謹慎小心，就是讓人難以預測，難以捉摸的程

度到敵我雙方都猜不透。

具備人工智慧的電腦也許是均化器，可提供劣勢者新的行動選項。但西洋棋並非僅能教我們超級電腦如何影響戰爭，現今人類雖仍主導戰略策劃，但電腦亦可在速度或避免失誤方面提供優勢。然而，若電腦與人類實力過度懸殊時，應可預期人類會出現奇特的行為反應。

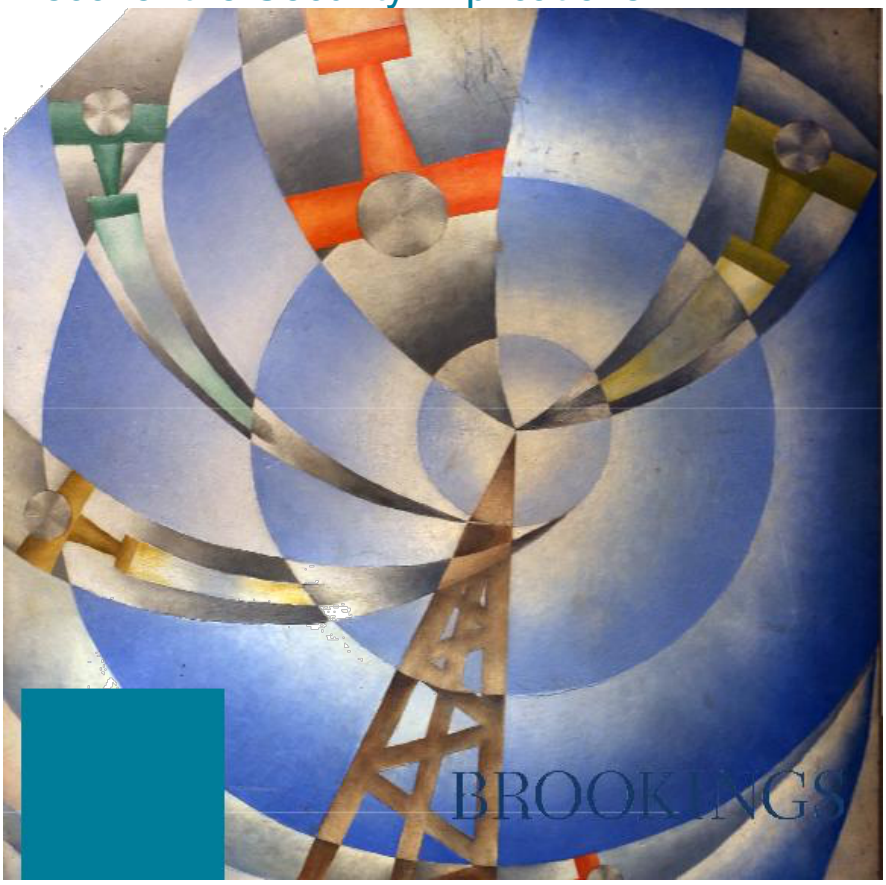
理想情況下，失去電腦輔助的人會撤退或轉向安全和保守的決策。但是戰場規則不像西洋棋的規則那麼嚴謹，如果你剛好發現敵方的電腦效率低落，而非超級電腦，那麼最好能聰明點預判，敵手可能會出現突發奇想、天外飛來一筆，甚或是輕率魯莽的行為。

安德魯·洛恩在非盈利且政治取向中立的蘭德公司擔任資深資訊科學家。他的研究主題包括人工智慧的軍事應用和機器學習。他也是《人工智慧如何影響核戰爭風險？》一書的合著者（蘭德 2018 年出版）。

Source text 2 : Artificial Intelligence, Geopolitics, and Information Integrity

THE GLOBAL RACE FOR TECHNOLOGICAL SUPERIORITY

Discover the Security Implications



edited by **Fabio Rugge**
introduction by **John R. Allen** and **Giampiero Massolo**

THE GLOBAL RACE FOR TECHNOLOGICAL SUPERIORITY

edited by Fabio Rugge

Via Alamanni, 11 – 20141 Milano – Italy
www.ledizioni.it
info@ledizioni.it
The Global Race for Technological Superiority Edited by
Fabio Rugge
First edition: November 2019
This report is published with the support of the Italian Ministry of Foreign Affairs and International Cooperation (in accordance with Article 23- bis of the Decree of the President of the Italian Republic 18/1967), within the framework of the activities of the Centre on Cybersecurity jointly promoted by ISPI and Leonardo. The opinions expressed are those of the authors.

Print ISBN 9788855261432
ePub ISBN 9788855261449
Pdf ISBN 9788855261456
DOI 10.14672/55261432
ISPI. Via Clerici, 5
20121, Milan
www.ispionline.it
Catalogue and reprints information: www.ledizioni.it
Cover image: Fulvio ranieri mariani, turbine aereo, 1938

BROOKINGS

The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public.

Table of Contents

5. Artificial Intelligence, Geopolitics, and Information Integrity

John Villasenor

Much has been written, and rightly so, about the potential that artificial intelligence (AI) can be used to create and promote misinformation. But there is a less well-recognized but equally important application for AI in helping to detect misinforma-tion and limit its spread. This dual role will be particularly im-portant in geopolitics, which is closely tied to how governments shape and react to public opinion both within and beyond their borders. And it is important for another reason as well: While nation-state interest in information is certainly not new, the incorporation of AI into the information ecosystem is set to accelerate as machine learning and related technologies experi-ence continued advances.

The present article explores the intersection of AI and information integrity in the specific context of geopolitics. Before addressing that topic further, it is important to underscore that the geopolitical implications of AI go far beyond information. AI will reshape defense, manufacturing, trade, and many other geopolitically-relevant sectors. But information is unique because information flows

determine what people know about their own country and the events within it, as well as what they know about events occurring on a global scale. And in-formation flows are also critical inputs to government decisions regarding defense, national security, and the promotion of eco-nomic growth. Thus, a full accounting of how AI will influence geopolitics of necessity requires engaging with its application in the information ecosystem.

This chapter begins with an exploration of some of the key factors that will shape the use of AI in future digital information

technologies. It then considers how AI can be applied to both the creation and detection of misinformation. The final section addresses how AI will impact efforts by nation-states to promote – or impede – information integrity.

AI and the Information Ecosystem: Some Key Factors

Advancing AI Technologies

A combination of factors will determine how AI will impact the information ecosystem over the next decade. First, there is the technology itself. Spurred by extraordinary levels of both private and public investment, AI is advancing at far greater rates than in the past. According to CB Insights, venture capital investment in the United States in AI startups grew from $4.1 billion in 2016 to $5.4 billion in 2017 to $9.3 billion in 20181. The US government has also been ramping up its support for AI research. For example, in fall 2018 the US Department of Defense's Defense Advanced Research Projects Agency (DARPA) announced a "$2 billion campaign to develop next wave of AI technologies"2.

In China, which views AI as a central focus of its goal of becoming a technological superpower, the government has launched a wide array of multi-billion-dollar AI investment initiatives3. Israel is another key player in the global AI landscape. In 2018, "AI-related companies accounted for 17% of the total number of 6,673 active Israeli tech companies in Israel tracked by Start-Up Nation Finder" and "32% of all funding rounds and 37% of the total capital raised went to AI-related companies"4. And in Europe, the European Commission has announced a plan aimed at spurring "more than €20 billion per year from public and private investments" in AI over the 2020s.

An additional aspect of the landscape not captured by the statistics above is the enormous internal AI research and development investment being made by large companies such as Amazon, IBM, Google, and Microsoft. Collectively, the capital flowing from governments, venture investors, and corporations will spur extraordinary AI advances, greatly broadening the capacity to analyze and make effective use of data. Relatedly, continued investment will make AI better at learning, opening the door to increasingly sophisticated algorithms that combine human ingenuity with computer-driven insights.

The Growing Role of AI in the Digital Ecosystem

A second factor that will elevate the role of AI is the degree to which it will be increasingly intertwined with broader digital information ecosystem. Many of the most important information technology changes of the last quarter of a century – including the growth of the internet, advances in digital storage and computation capacity, and the introduction and mass adoption of smartphones and social media – have occurred largely (though not completely) without AI. By contrast, the future evolution of the digital information landscape will be driven in significant part by AI.

Over about the last five years, we have been experiencing the first stages of this transition, and AI is now used a wide range of commercial products and services. There is an understandable temptation to predict the future by extrapolating the past, and therefore to conclude that the next 5 or 10 years see the introduction of even more AI into the commercial ecosystem to enhance consumer services in areas such as transportation, online purchasing, and media delivery. But while that prediction is no doubt accurate, it almost certainly fails to anticipate the more profound AI-induced changes that are much harder to foresee in advance.

By analogy, consider the internet in the late 1990s. At that time, it would have been relatively easy to predict dramat-ic growth in both the number and diversity of web sites over the subsequent 10 years. But it would have been much hard-er to envision the growth and impact of social media—which we now know spurred far more significant changes than did growth in the number of websites. In the same way, it is easy today to conclude that AI will play an increasingly large role in the digital information landscape over the next decade, but far harder to anticipate its use in ways that lack clear historical antecedents.

Information Gatekeepers

Information gatekeepers, including but not limited to social media companies, constitute a third factor influencing how AI will shape the information ecosystem. For large-scale social media companies, as well as other companies (such as online retailers and providers of internet and mobile phone services) that engage with millions of individual users, the question is not whether to incorporate AI, but rather how it should be most effectively used to further goals such as offering highly

customized content to consumers and detecting fraud. As AI continues to advance, companies seeking to take advantage of the cost efficiencies it enables have incentives to deploy it more extensively in their systems. Companies will make highly consequential policy choices regarding their development and rollout of AI solutions, addressing questions such as the extent to which they should curate and/or filter content, the standards they will apply in relation to testing and monitoring algorithms to detect problems such as bias, and the level of human oversight to provide in relation algorithmic decisions and algorithmic evolution.

In authoritarian countries, an additional information gatekeeper is the government itself. All authoritarian governments will seek to use AI to monitor online traffic and detect digital content deemed problematic. But there will be variations both across and within authoritarian countries in the nature of the tools employed and the extent to which they are used to actively control (as opposed to monitor) discourse.

AI and Information Integrity

"Information integrity" as used herein is intended to describe the extent to which information is accurate, non-deceptive, and properly attributed. While accuracy is clearly a baseline requirement to achieve information integrity, accuracy alone will not always be sufficient. For information to have integrity it also has to be contextualized in a manner that avoids deception. To take a simple example, consider a politician who accompanies a family member who has struggled with drug addiction on a visit to a drug rehabilitation clinic. Suppose that the politician is photographed when leaving the clinic, and that those photographs are then distributed on social media. The photographs are accurate in the sense of depicting an event that actually oc-curred, but they are deceptive because, when distributed with-out context, they could imply that the politician is personally struggling with drug addiction.

Attribution is also important. A social media posting pur-porting to come from a voter and containing accurate, properly contextualized content still lacks integrity if in fact it was posted by a foreign government aiming to influence an election. Thus, challenges to assessing the integrity of information include not only evaluating truth or falsity, but also identifying the extent to which decontextualization may lead to misinterpretation, as well

as understanding whether the purported source is the same as the actual source.

Much of the recent public dialog regarding the role of AI in information integrity has focused on potential negative im-pacts. Deepfakes, which are videos produced with the aid of deep learning techniques that portray people doing or saying things that they never did or said, have been correctly identified as a major potential concern6. A well-constructed deepfake targeting a politician, if released onto the internet at the right time and manner, could potentially swing a close election.

AI can also be used to undermine information integrity in other ways. Consider "bots", which describe accounts on Twitter and other social media platforms that masquerade as humans but are actually software (though as of yet, not gen-erally AI-enabled software). While precise statistics on the per-centage of Twitter accounts that are bots are hard to come by (in part due to fluctuations over time as different bot detection techniques are developed and deployed, and as bot creators then react by updating their methods), it is clear that the num-ber is very high.

Bots are known to play an important role in amplifying online misinformation. A November 2018 paper published in Nature Communications reported on a study of "14 million messages spreading 400 thousand articles on Twitter during ten months in 2016 and 2017"7. The authors found "evidence that social bots played a disproportionate role in spreading articles from low-credibility sources. Bots amplify such content in the early spreading moments, before an article goes viral. They also target users with many followers through replies and mentions. Humans are vulnerable to this manipulation, resharing content posted by bots"8. As noted above, in the past, most bots have not been AI-enabled. Inevitably, this will change. Well-designed AI-powered bots could do a very effective job of impersonating humans, making them much harder to detect and more effec-tive at disseminating misinformation.

As concerning as the above examples are, it is also important to consider the other side of the ledger. Just as AI can be used to promote misinformation, it can also be used to combat it. Deepfake detection is one example. There is a very active community of researchers working to develop methods, including approaches based on AI, to automatically identify manipulated

videos. Examples include the use of deep learning to identify artifacts introduced by face-swapping software9 and the use of neural networks to identify frame-to-frame inconsistencies in deepfake videos10. As a February 2019 article in IEEE Spectrum noted, “the AI Foundation raised $10 million to build a tool that uses both human moderators and machine learning to identify deceptive malicious content such as deepfakes”11. The same article also described efforts by a Netherlands-based tech-nology startup to use adversarial machine learning “as a primary tool for detecting deepfakes”12.

AI can also be used to detect activity by bots. Bots that do not rely on AI often act in recognizable ways that can easily be detected. The authors of the Nature Communications article noted above observed that when low-credibility content goes viral, it exhibits “distinctive patterns”. The authors explained that

most articles by low-credibility sources spread through orig-inal tweets and retweets, while few are shared in replies; this is different from articles by fact-checking sources, which are shared mainly via retweets but also replies. In other words, the spreading patterns of low-credibility content are less “conver-sational”. Second, the more a story was tweeted, the more the tweets were concentrated in the hands of few accounts, who act as “super-spreaders”.

By contrast, in the future when many bots become AI-enabled, they will be more capable of emulating organic, non-coordinated viral behavior, in part by creating larger networks to spread tweets and in part by relying more on including misinformation in “replies” that might appear to have been written by a real person. The most effective way to identify and block AI-enabled bots will be to use AI in the detection algorithms. Such algorithms could monitor the evolving behavior of a bot network, and in response evolve their own templates for identi-fying likely non-human social media activity.

The examples of deepfakes and bots illustrate that while misinformation poses major challenges, the same powerful AI techniques that can be employed to produce false or deceptive content can also be applied to its detection and mitigation. A challenge is that the asymmetries involved give misinformation creators an inherent set of advantages. They can continually enhance their algorithms to stay one step ahead of the latest detection techniques. And, to have impact, misinformation creators only have to succeed some of the time. Even if only a low percentage of malicious content evades detection, that can still be enough to cause significant harms.

Governments and the Information Ecosystem

As the above discussion makes clear, over the next decade AI will experience dramatic advances and take on an increasing role in the broader digital information ecosystem. At the same time, AI-based techniques for generating misinformation will become more sophisticated, as will techniques for detecting and impeding its spread.

This will impact geopolitics in multiple important ways. In authoritarian countries, governments have always sought to exert high levels of control over information, both through propagation of state-approved content and censorship of content deemed inconsistent with the government objectives. AI offers a powerful tool for achieving these ends. To take one example, AI can make it easy for an authoritarian country to perform highly detailed inspection and censorship of social media post-ings. Postings can be examined not only individually, but also in the aggregate for an individual or group of individuals to identify broader trends that might be of interest to the govern-ment. Authoritarian governments will make use of these capa-bilities to further geopolitical (and other) goals.

Inevitably, some governments will also seek to use online misinformation to alter elections in other countries. The well-documented foreign manipulation of US social media to attempt to influence the 2016 US presidential election is, unfortunately, only a foreshadowing of what is likely to occur in future high-stakes elections. AI-powered misinformation aimed at swaying voter perceptions can be very effective. Combating it will be challenging in part because of the high degree of coordination that would be needed among multiple private and public sector entities to identify and mitigate foreign government misinformation. Yet another complicating factor is that some forms of manipulation can be subtle and therefore not easily detectable. For instance, a foreign government might use AI to create social media accounts in the target country and cause those accounts to engage in much more humanlike behavior than would be possible without AI. The accounts could be used not only to propagate outright misinformation, but also to amplify nega-tive but accurate information about a political candidate, there-by giving it more visibility among the electorate than it would have received absent the foreign influence.

A foreign government seeking to tip the scales in an election would have a long list of options for specific ways of undermining information integrity. A 2019 RAND Corporation re-port on "Hostile Social Manipulation" identifies over a dozen methods of social manipulation, including "content creation", "disinformation", "social media commenting", "direct advertising", "trolling", "behavioral redirection" and "microtargeting"15. With AI, all of these methods could be used at scale and in ways that might be difficult to mitigate, particularly given the importance of minimizing false positives, which could lead to suppression of legitimate social media content posted by real voters.

While election interference is an extremely important way in which nation-state might seek to use AI-generated misinformation to further geopolitical goals, it is not the only one. Nation-states might also use AI to disseminate information aimed at influencing a foreign government's geopolitically-relevant legislation; regulations; trade, economic, and defense policies; and decisions regarding major mergers and acquisitions. A nation state might also manipulate information to boost posi-tive consumer perceptions of companies headquartered within the nation-state, thereby boosting the global competitiveness of those companies, and by extension, the nation-state. And, AI-enabled information manipulation will be a central feature of any future large-scale military conflict. This would include not only attempts to shape public opinion, but also efforts to undermine the availability and accuracy of information relied upon by military decisionmakers and political leaders.

Conclusion

So how can societies – and in particular democracies built on the free flow of information and ideas – address AI-enabled misinformation created and/or propagated by a foreign government?
Technology, policies, and awareness can all contribute to a solution. With respect to technology, as noted above, the same advances in AI that are making it easier to generate misinformation can also be used to detect it. Many of the tradeoffs involved parallel those found in cybersecurity, where there are also complex decisions to be made regarding how to allocate resources in relation to prevention, detection, and mitigation. The experience from that sector can help inform both public and private

sector approaches to ensuring information integrity.

Governments should be both investing directly in research on improved detection as well serving as a resource for the private sector through information-sharing arrangements that can help companies better understand potential foreign manipulation of social media and other online information. The information flow can work in the other direction as well: Companies, and in particular social media companies, will be at the front lines of foreign-directed misinformation campaigns, and thus are well positioned to understand their dynamics and convey the lessons learned on to other companies and to the government.

Policy solutions can include the use of existing legal frame-works as well as new legislation. In considering the legal land-scape, it is important to keep in mind that not all approaches that undermine information integrity will involve false state-ments. A foreign government might simply seek to amplify or suppress accurate information in ways aimed at swaying pub-lic opinion. When this occurs in the context of an election, it can be addressed through statutes aimed at combating election meddling. As important as such statutes are, their effectiveness will be limited due to the time scales involved (in many cas-es, the election will be long over by the time the legal system swings into action) and due to the fact that elections repre-sent only one of the many potential targets of a misinformation campaign.

That highlights the importance of a final tool: increased awareness. In an era where deepfakes and other forms of man-ufactured or manipulated content will become more common, broader awareness can help slow (though certainly not stop) their spread. In promoting this greater understanding, it will also be important not to undermine the trust in legitimate in-formation which is at the foundation of all democratic societies.

全球科技優勢競賽

發掘安全意涵

編者 **Fabio Rugge**

導讀 **John R. Allen** 及 **Giampiero Massolo**

全球科技優勢競賽

Via Alamanni, 11 – 20141 Milano – Italy www.ledizioni.it
info@ledizioni.it

全球科技優勢競賽 Fabio Rugge 主編
一版：2019 年 11 月

BROOKINGS

布魯金斯學會是一個非營利性組織，致力於獨立研究和政策解決方案。其使命是進行高質量、獨立的研究，並植基於研究，為政府和公眾提供創新、實際可行的建議。

目錄

5. 人工智慧、地緣政治和資訊完整性

John Villasenor

已有許多文章在談論人工智慧（AI）可能可用來製造與宣傳假訊息，的確沒錯，然而另一項人工智慧的應用雖鮮為大眾所知，但卻同等重要，即人工智慧用來幫助***檢測***假訊息並限制其傳播。這種雙重角色在地緣政治中尤其重要，因為地緣政治與政府在國內、外如何塑造和應對公眾輿論息息相關。另一個也相當重要的原因，雖然民族國家對資訊感興趣已不是一天、兩天的事了，但隨著機器學習及相關科技經驗的不斷進步，將加快人工智慧融入資訊領域的速度。

本文探討在地緣政治特定背景下，AI 與資訊完整性的交集。在進一步討論主題前，必須先強調 AI 在地緣政治領域的涉入程度遠超過資訊領域，AI 可重塑國防、製造業、貿易及許多地緣政治相關的領域，但是資訊特殊的地方在於，資訊流可決定人們如何看待自己國家和國內事件，以及如何看待全球事件。再者，政府在做國防、國家安全和促進經濟成長等決策時，將資訊流納入也至關重要。因此，要全面說明 AI 如何影響地緣政治，就必然無法不探討 AI 在資訊領域中的應用。

本章首先探討在未來數位資訊科技中，決定 AI 將如何發展的一些關鍵因素；接著，思考 AI 如何應用於創造假訊息和檢測假訊息；最後一節討論 AI 將如何影響民族國家在提升或阻礙資訊完整性的投入。

AI 與資訊領域：一些關鍵因素

提升 AI 技術

AI 未來十年將如何影響資訊生態系統的因素有很多。首先，是科技本身。在私人和政府大量挹注資金的驅策下，AI 技術進步的速度較以往大幅提升許多。根據美國市調公司 CB Insights 的資料顯示，美國 AI 初創企業的創投資金，從 2016 年的 41 億美元增長到 2017 年的 54 億美元，2018 年更增長到 93 億美元。美國政府也加大了對 AI 研究的支持力度，例如：2018 年秋季，美國國防部國防高級研究計劃局(DARPA)宣佈「投入 20 億美元開發下一波 AI 技術」。

中國將 AI 視為邁向科技強權目標的重心，中國政府推出了一系列耗資數十億美元的 AI 投資計劃。以色列亦是全球 AI 發展中的要角，2018 年，「根據以色列 Start-Up Nation Finder 的統計資料顯示，以色列 AI 相關企業占全國總計 6,673 家科技公司的 17%」，「所有融資回合的 32%和募集資金總額的 37%皆流向了 AI 相關公司」。在歐洲，歐盟委員會宣佈了一項計劃，旨在激勵於 2020 年代期間「政府和私人每年在 AI 產業的投資超過 200 億歐元」。

上述統計數據未呈現的另一面向：亞馬遜、IBM、谷歌和微軟等大企業正在投入大量資金在企業內部的 AI 研發。總體而言，來自政府、創投者和企業的資金都將刺激人工智慧的超常發展，極大地提升了 AI 在分析及有效利用數據的能力。相關方面，持續挹注資金將使 AI 更精於學習，可為發展日益精細的演算法開啟一扇門，將人類的創造力與電腦驅動的洞察力相結合。

AI 在數位領域日益重要

AI 地位提升的第二個因素是，它與更廣泛的數位生態系統將更緊密地交織作用。過去 25 年來，許多相當重要的資訊技術變革，包括網路的增長、數位儲存技術和電腦計算能力的進步，以及智慧手機和社群媒體的引入和大規模使用，上述的變革都是在沒有 AI 的情況下發生（儘管並不完全是）。相反地，未來數位資訊的變革，在很大程度上將由 AI 所驅動。

過去這五年來，我們經歷了這場變革的初始階段，目前 AI 已廣泛地用於商業產品和服務。鑑往知來，可推測未來 5 年或 10 年，AI 將更大程度地融入商業圈，以提升消費者服務，例如在交通運輸、線上購物和媒體傳輸等方面。儘管對上述預判無庸置疑，然而，我們卻無法預見 AI 所帶來更長遠的改變。

打個比方，想想 90 年代末的網際網路，在當時，並不難預料十年後的網站數量和種類將蓬勃發展；但是，要想像社群媒體未來數量和影響力將增加，相對困難。比起網站數量的激增，我們現今才知曉社群媒體所引起的變化更顯著得多。同理，今日很容易得知：未來十年，AI 將在數位資訊領域扮演日益舉足輕重的角色，但我們卻很難預知其未來用途，係因缺乏前鑑可參考。

資訊守門人

資訊守門人含括社群媒體公司，但不僅限於社群媒體公司，是影響 AI 如何形塑資訊領域的第三個因素。大型社群媒體公司以及其他如線上零售商、網路和電信服務提供業者，其涉及數百萬個體用戶，他們面臨的問題不在於是否採納 AI，而在於如何發揮 AI 的最大效能，例如提供消費者高度客製化的內容，以及查察網路詐騙。隨著 AI 的不斷進步，以成本效益為考量的公司廣泛地在公司系統中使用 AI。公司在開發 AI 以及 AI 解決方案等方面做決策時，將會優先考量其後果，並提出諸如以下的問題：籌劃及（或）篩選內容的標準為何？為了挑出偏見內容等問題，測試和監視演算法的標準為何？在演算法的決策和發展方面，人類該監督到何種程度？

在威權國家，尚有一個資訊守門人即是政府本身。每個威權政府都欲使用 AI 來監控線上流量，並檢測政府認為有問題的數位內容。然而，每個威權國家所用的工具性質以及其積極掌控言論（而非僅是監視）的力度，都將有所不同。

AI 和資訊完整性

此處的「資訊完整性」旨在描述資訊準確性、無誤導和轉載適切性的程度。雖然準確性顯然是滿足資訊完整性的基本要求，但僅有準確性並不夠。資訊要具備完整性，還必須在情境上排除誤導的成分。舉一個簡單的例子，試想一個政治人物陪同有藥癮的家人去勒戒診所看病，假設該政治人物在離開診所時被拍照，然後這些照片在社群媒體上流傳，這些照片確實描繪了當時實際發生的事件，但它們會造成誤導，因為若沒有解釋來龍去脈，這些照片可能影射這位政治人物正在與毒癮奮戰。

轉載也很重要。在社群媒體上的一篇貼文聲稱其資訊轉載自某選民，即使其內容正確、圖文情境相符，然實際上卻是某外國政府所為，意圖影響該國選情，則該篇貼文內容仍然缺乏完整性。因此，評估資訊完整性的困難處不僅包括判斷其真假，而且還需判別該則訊息在沒有背景情境的說明下，其可能導致誤解的程度，此外，還需瞭解所宣稱的資訊來源是否屬實。

近期，大眾在談論有關 AI 在資訊完整性中的功用時，大多側重於其潛在的負面影響。深偽（Deepfakes），是藉助深度學習技術所製作的影片，片中人的一言一行都是他們從未做過或說過的事，已被指正為主要的潛在問題。一支針對政治人物所精心構思的深偽影片，倘若在天時地利配合下於網路上發佈，恐會動搖一場難分軒輊的選舉。

AI 還能以其他手段破壞資訊完整性。想想「機器人 bots」，它是用來形容在 Twitter 和其他社群媒體平臺上的帳號，這些帳號偽裝成人類使用者，但實際上它們是軟體（雖然就目前為止，普遍不是 AI 軟體）。雖然機器人帳號在 Twitter 上的佔比數，我們難以獲得精確統計數據（有部分是因機器人帳號的偵測技術開發和使用會與時俱進，而機器人帳號的創建者隨即會更新應對之道），但很肯定的是這個數字非常高。

眾所周知，機器人帳號是網路假訊息散播的一大幫凶。2018 年 11 月發表在《*自然通訊*》上的一篇研究報告「2016 至 2017 年的十個月期間內，在 Twitter 上有 1,400 萬條訊息散播了 40 萬篇文章」，該篇報告作者發現「社群媒體的機器人帳號在散播來源可信度低的文章方面，能起到四兩撥千斤之功效。在文章流傳的初期，機器人帳號助長了其聲勢，讓該文章能在網路瘋傳。機器人帳號還會從標記他人與留言回覆中，鎖定有許多關注者的帳號。人們容易受到此種操控，而轉貼分享機器人帳號所發佈的內容。」如前文所述，以往機器人帳號多半不具備 AI 功能，但未來趨勢必定會改變。設計良好的 AI 驅動機器人模仿人類將惟妙惟肖，更難以被識破，且能更有效地散播假訊息。

除了考量上述例子，AI 的另一面向也同等重要。正如 AI 可用於宣傳假訊息一樣，它亦可用來打擊假訊息，比如深偽檢測。有一群積極的研究人員正在開發偵測深偽的方法，其中包括以人工智慧自動識別被動過手腳的視頻。例如：利用深度學習來識別換臉軟體加工過的影像，以及使用神經網路辨別深假影片中的每個影格間是否有不連貫之處。科學工程技術雜誌《*IEEE 綜覽（IEEE Spectrum）*》在 2019 年 2 月刊登的一篇文章，文內提及「AI 基金會募集了 1,000 萬美元打造

一套工具，可同時以人工判斷及機器學習來識別惡意的詐騙內容，比如深偽」。該篇文章還提及了一家總部位於荷蘭的科技創業公司，致力於運用對抗性的機器學習「作為檢測深偽的主要工具」。

AI 還可用於檢測機器人帳號的活動。不靠 AI 輔助的機器人帳號通常其一舉一動瞭若指掌，因此易於檢測。〈*自然通訊（Nature Communications）*〉一文的作者指出：當可信度低的內容瘋傳時，會表現出「特殊的態樣」，作者解釋：

> 大多數來源可信度低的文章是透過原始推文和轉推而流傳，極少是在留言回覆中分享；這與來源有事實查核的文章不同，後者主要透過轉貼分享，但也透過留言回覆。換言之，內容可信度低的傳散模式較不那麼「對話性(conversational)」。再者，一篇文章若有愈多貼文，貼文則愈集中在少數帳號上，這些帳號即是「超級傳散者」。

相對地，當未來許多機器人帳號開始有了 AI 的置入，這些帳號將更能模仿有機的、非協調的網路瘋傳行為，一方面是藉由創造更廣闊的網絡來傳散貼文，另一方面則在「留言回覆」中夾雜更多假訊息，並讓這些假訊息看起來如真人所為。欲辨識和阻擋 AI 機器人帳號的最有效方法是在檢測系統的演算法中使用 AI。此類演算法可以監視機器人網絡不斷變化的行為，並發展可與之應對的模式，以辨識非真人所為的社群媒體活動。

深偽和機器人帳號的例子說明了，雖然假訊息構成重大挑戰，但同樣強大的 AI 技術，可用於製造虛假或欺騙性的內容，亦可以應用於檢測和減少假訊息。但具挑戰性的是，兩邊形勢上的不對稱給予假訊息製造者先天上的優勢。假訊息製造者可以不斷增強演算法，總是走在檢測技術的前頭。再者，為了發揮影響力，假訊息製造者只需要在一段時間內成功散佈訊息，即使只有極少數的惡意內容躲過檢測，仍足以造成重大危害。

政府與資訊領域

以上論述已明確表達，未來十年，人工智慧的成長將大幅躍進，並在更廣大的數位資訊領域中扮演日益重要的角色。同時，用來產製假訊息的 AI 技術將變得更加複雜，同樣地，偵測和阻止其傳播的技術也將變得益加複雜。

地緣政治將因此在多重層面大受影響。在威權體制國家，政府一直試圖藉由傳播國家批准的內容，以及審查違背政府意志的訊息內容，對訊息實行高度控制，而 AI 即是威權政府強而有力的執行工具。舉例來說，AI 可以讓威權國家輕易地對社群媒體上的發文進行相當詳細的檢查和審查，不僅可以檢視單則發文，也可以對單一個人的所有貼文或是對一群人的貼文進行審查，以辨識出政府可能感興趣的輿論趨勢。威權政府將利用這些能力進一步實現地緣政治及其他目的。

無可避免的是，某些政府也會意圖利用網上的假訊息來改變其他國家的選舉。不幸的是，外國政府藉由操縱美國社群媒體以影響 2016 年美國總統大選，已證明屬實，此事件僅是個開端，它預言了未來涉及高度利害關係之選舉的可能樣貌，利用 AI 所驅動之假訊息來動搖選民的看法可說是相當有效，欲打擊此種行為具有挑戰性，部分原因是需要大力協調許多私營和公共部門，方能查明並減少外國政府的假訊息。另一個複雜的因素是，某些形式的操縱手法細微而難以察覺，因此不容易被檢測到。比方說，外國政府可能使用 AI 在目標國家創建社群媒體帳號，並用這些帳號從事比無 AI 輔助時更人性化的行為，這些帳號不僅公然地傳播假訊息，還可以放大有關候選人真實卻負面的資訊，因此藉由提高訊息能見度，讓選民比以往沒有外國影響時更容易接收到此類訊息。

若外國政府想扭轉他國選舉結果，在破壞資訊完整性方面可能有一長串的具體方法。蘭德公司 2019 年出刊的報告「敵對社會操縱」，文中列舉出了十餘種社會操縱的手法，其中包括「創造內容」、「虛假資訊」、「社群媒體留言」、「直接廣告」、「酸民意見」、「行為重導」和「微目標」。上述的手法搭配 AI 併行，其影響規模之大，難有緩解之道，尤其減少誤判相當重要，因為誤判可能導致真實選民在社群媒體所發佈的合法內容受到壓抑。

雖然干預選舉對民族國家來說是極其重要的手段，民族國家可能試圖利用 AI 生成的假訊息來進一步實現地緣政治目的，但干預選舉並非唯一手段。民族國家也可能利用 AI 散播資訊，以左右外國政府在地緣政治相關的立法、法規、貿易、經濟和國防政策，以及影響重大合併和收購的相關決定。民族國家也可能藉由操縱資訊，提高消費者對國內本土企業的正向觀感，從而提升這些企業在全球的競爭力，進而提升民族國家的競爭力，且運用 AI 操縱資訊將會是未來所有大規模軍事衝突的核心特徵，這不僅包括試圖塑造大眾輿論，而且還包括讓軍事決策者和政治領導人無法取得所需的正確資訊。

結論

那麼社會大眾要如何應對外國政府利用 AI 製造或散佈的假訊息呢？特別是思想與資訊流自由的民主社會要如何應處？

運用科技、制定政策和提高公民對假訊息的敏感度都是解決方案。關於科技方面，誠如上述，人工智慧的進步除了讓假訊息的生成更不費力氣，同樣地也可用於檢測假訊息。在防範假訊息上需要做許多權衡與妥協，這方面與網路安全相同，都需要在防範、檢測和減低傷害相關的資源分配上做複雜的決策，網路安全領域的經驗有助於提供公共和私營部門確保資訊完整性的方法。

各國政府應將經費直接挹注於改善假訊息檢測的研究，並透過資訊共享的規劃安排，提供私營部門資源，協助企業公司更加瞭解外國政府如何操弄社群媒體和其他網路資訊。資訊流也能在其他方面發揮效用，在外國政府發起的假訊息行動中，公司企業將會處於第一線，尤其是社群媒體公司，由於身處前線，因此可相當清楚其一舉一動，並將所汲取的教訓傳授給其他公司和政府。

政策制定方面，包括利用現有法律框架以及訂定新法。在法律層面的考量，必須謹記：並非所有破壞資訊完整性的手法都涉及內容虛構。外國政府很可能僅意圖透過放大或壓抑正確訊息的方式來左右公眾輿論。當此種情況發生在選舉時，可以透由遏制外力干涉選舉的法規加以解決。儘管這些法規很重要，但法規約束的效力有限，係由於其時效性差（大多數情況下，終於要採取法律行動時，選舉早已結束），而且選舉僅是假訊息眾多的可能目標之一。

這也因此突顯出最後一項解決方案的重要性：提高公民對假訊息的敏感度。深偽和捏造訊息內容的現象將會日益普遍，身處在這樣的時代，若有更多民眾能夠覺察不實訊息的存在，將有助於減緩不實訊息的傳播（即使無法阻止其傳播）。在提高公民對不實訊息的認知時，與其同樣重要的是，不要破壞公民對正當資訊的信任，因為這是所有奠定民主社會的基石。

Self-Commentary

很感謝陳瑞清老師交派了這項期末作業，我才有機會完成人生第一次的翻譯專案，第一次獨自完成長達五千餘字的翻譯作品，姑且不論品質好與壞，都要先給自己一點掌聲，因為唯有親身經歷過了，才能體驗譯者的箇中甘苦。由於是第一次選擇學術研究性質的文章來翻譯，過程中遇到了不少糾結的難題，全賴老師與同學一起在課堂上的腦力激盪與知識分享，才得以順利完成全文翻譯，這也才瞭解筆譯不是外行人所想的如此容易被機器翻譯所取代。

此次的翻譯專案，我利用微軟 Word 內建的翻譯協助處理冗長的原文，原本以為機器翻譯可以大幅縮短時程，但發現不盡然。先說機器翻譯的優點，首先它可以在短瞬間處理大量資訊，若在上級臨時急需長篇資料的譯文時，起碼可以在第一時間利用機器翻譯產生初稿，先應付長官十萬火急的需求。再者，機器翻譯的品質雖不比人工翻譯精緻細膩，但總體而言還勉強上得了檯面，我想應歸功於中國，隨著中國國力的日趨強大，華文在世界的流通使用頻率就愈趨頻繁，有了廣大的語料庫再加上人工智慧與機器學習等科技的日益精進，逐漸提升了機器翻譯的品質，在面對大量資料時，先由機器處理冗長的原文可以減低譯者在面對長篇幅蝌蚪文的煩躁感。

凡事都有一體兩面。雖然兩岸都使用中文，但語言使用習慣大不相同，就如同英國、美國、澳洲都以英文為其官方語言，但其口音、用語卻各有其文化及地方特色。而中英翻譯的語料庫取材多來自於中國大陸及香港，很多字彙的譯文都很有中國味，在考量我的目標閱眾為台灣人的情況下，更需謹慎留意哪些字詞屬於大陸用語。拿我的專案中最常出現的字來說，social media 機器翻譯為「社交媒體」，而台灣用語卻是「社群媒體」；internet 機器翻譯為「互聯網」，而台灣習慣用語卻是「網際網路」；information 機器翻譯為「信息」，而台灣用語為「資訊／訊息」。

機器翻譯的另一項缺點是太貼著原文字面翻譯，很容易就限制了譯者在譯文修校上的靈活度與彈性。機器翻譯畢竟還是有其侷限，從我此次的翻譯專案上所觀察到的問題包括：上下文的承接與流暢度、複雜子句的解構、副詞或是形容詞等修飾性詞彙的搭配運用……等方面，與人工翻譯相較之下，還是有很大的落差，技術性的文件尚且如此，更不用提文學類作品的翻譯了。從這點也可明白人工智慧機器翻譯可以成為筆譯工作者的輔助工具，但要取而代之恐怕還言之過早。

此次的翻譯專案也讓我體會到筆譯工作者的辛苦，除了譯者本身 A 語言與 B 語言需具備高度的語文能力素養之外，長篇翻譯還考驗著譯者的耐性與毅力以及自我鞭策的自制力，我才翻譯這五千餘字就快耐不住煩，且到學期後半段心態鬆懈，落後了原本的規劃進度許多，回想起來相當慚愧！

陳老師告訴我們，翻譯也可看作是「再創作」，創作本身不是件易事，相對地，再創作亦然，因此當我看著自己的翻譯專案完稿時，真有作品誕生的歡

欣感與成就感，將這種感覺轉換為保持筆譯的熱情與動力，期許自己在這條路上繼續堅持下去！

黃瑞青 2020/5/26