

站在「人工智慧」巨人的肩膀上

人工智慧崛起看似憂喜參半，但AI不該是絆腳石，而是墊腳石。人類應當與AI攜手面對難題，借助其強大的能力，成就原來無法成就的事。

從AlphaGo打敗人類棋王後，人工智慧（AI）儼然成為新一代顯學。AI領域的科學家最常採取的研究方法是機器學習，而機器學習背後最大的推手是與日俱增的巨量資料（big data）。這本博學誌《掌握數位生活》收錄了電腦、網路、教育等領域的進展，將帶領讀者從不同角度一窺這波數位革命。

首先，在〈會思考的電腦〉一文中，作者阿布·莫斯塔法深入淺出，探討機器可以「學習」如何完成以下幾種不同的工作。一種是機器學習可達成分類任務，例如我們給電腦一群正常人與另一群癌症病患的肺部X光影像，它能自動學習如何區分兩組資料，並在日後對其他X光影像進行正確分類；另一種是機器學習也可做出數值預測，例如我們把某家公司過去20年的財務資料及股價輸入電腦，它可訓練出一個預測該公司股價的模型。還有，機器學習甚至可提出一連串決策建議，當我們給電腦過去的一些決策資料，機器學習可以學到在什麼狀況下做出什麼樣的決定，可能會獲得最佳的利益，例如AlphaGo能夠根據過去棋手下棋以及自我模擬下棋的資料，來決定當下棋子的哪個著點會有最大的贏面。

機器學習的成功已受到廣泛認可。但一個擁有高度智慧的個體除了能「學習」，還要能「發明」與「發現」。AI如何超越一般人類的智慧，讓電腦成為如愛迪生、莎士比亞或福爾摩斯這類發明家或創造者，正是目前AI領域的研究者（例如我的團隊）著手努力的方向。

雖然機器學習的前景不可限量，但是它的發展過程並非一帆風順。「莫拉維克悖論」曾指出：AI在一些需要高度智慧以及推理的任務（例如下棋、預測）可以做得很好，卻無法辦到一些連兒童都做得到的簡單任務，例如辨識人臉、快速繞過障礙物。這也限制了它的應用範圍，不過近年來發展成熟的「深度學習」技術，在「辨識」這個領域終於有了突破，讓電腦在圖像及語音

的辨識成效已經逼近人類的表現。在〈機器如何深度學習〉一文中，作者班吉歐以多層神經網路為本，闡釋深度學習以及它為何成功。相信你可以發現，深度學習除了自身演算法在辨識上的優勢，也蒙受兩個重要的外來關鍵因素助益：電腦運算速度的提升（藉由新的圖形處理器）以及巨量標記資料的出現。

然而，深度學習還是有它的局限。在〈電腦有意識嗎？〉一文中，作者柯霍與托諾尼提出：電腦即使可以辨識照片中的物品，卻還無法達到更深層的「理解」。許多小朋友常常會玩一種遊戲「這張圖有什麼不對」，他們要從一張圖中找出不合理之處（例如有個人坐在電腦前，左右手都握著滑鼠）。這樣的任務對AI而言是非常困難的，因為這需要高度的資訊整合與理解能力。兩位作者認為所謂的「意識」是能夠整合訊息的理解能力，而非很多獨立存在的辨識系統。目前的深度學習可從照片中辨識出各種物品，卻還沒有能力「整合」這些資訊，獲得更高層次的理解。主要原因是這些資訊的組含量太大，電腦尚無法透過訓練來完成目標。

深度學習讓AI露出了曙光，背後一個重要的推手就是巨量資料。巨量資料不僅對個人產生影響（例如推薦系統服務），對於公司、城市、政府施政也帶來了前所未有的轉變。在〈巨量資料驅動城市〉一文中，作者潘特蘭談到自己如何利用電腦來分析人類行為的巨量資料，創造出更多價值。他用了「數位麵包屑預測」這樣一個很生動的說法：人類各式行為留下的一些數位記錄（例如臉書上聯絡誰、去哪裡、做了什麼事），都像從嘴邊掉落的麵包屑。然而在追蹤這些麵包屑之後，電腦將可以做出許多預測。例如，根據行為判斷一個人是否生病、預測他將來會買什麼衣服、甚至有沒有能力清償貸款。潘特蘭分析了某家公司內部人與人的互動，找出提高生產力的方法；他的研究也發現，從「社群」而非




個人的角度出發，往往能夠達成更大的功效；例如要推廣一項活動，鼓勵人們「傳播」這項活動給其他人，會比只鼓勵人們參與來得有用。

巨量資料在帶給人類便利時，也衍生一些潛在問題。在〈巨量資料下誰有隱私〉一文中，作者藍尼爾就點出了Google與臉書這些網路科技巨擘對隱私的弔詭作為：他們一方面向使用者提倡資訊透明化，一方面卻把自己預測使用者行為的模型或推薦系統埋聲晦跡、不願公開。他擔心這會造成資訊不對等的危機，擁有巨量資料的企業可以利用資料去操弄人們的行為決策。例如，公司為了利益，系統在推薦商品時，可能會把不適切的商品推薦給顧客；搜尋引擎也可以刻意優先呈現有付費的內容，影響使用者的選擇。藍尼爾的團隊主張「資料付費」，希望能解決隱私可能被侵犯的問題，因為當資料被視為某種具有商業價值的東西時，商業權益原則就能夠量化隱私，而經濟市場就可以決定資料的價值；在資料需要付費時，政府也不能無止盡利用國家安全之名來行使侵犯個人隱私之實。在〈挺住巨量資料 防洩密〉一文中，潘特蘭提出利用「資料分散儲存」的概念來防止資料盜取，並利用「信賴網路」設定審查機制來界定合法與非法的資料存取程序，以達成更完整的資料保護機制。

巨量資料也同時改變了人類的行為模式以及對自我的觀感。在〈Google效應——搜尋引擎如何改變你的心智？〉一文中，作者韋格納與沃德提出，網際網路降低了人類想把剛學到的重要知識記憶在頭腦中的渴望。因為我們已經認知到，幾乎所有的知識都可以藉由搜尋網路取得，於是減少了「記憶」的必要性，空出來的心智就能用在我們的雄心壯志上。他們進一步發現，很多人已經把網際網路當成自己認知能力的一部份，也就是當人們經由搜尋引擎找到相關資訊，會歸功於自己而非搜尋演算法。這也造成了一個有趣的矛盾：網路世代是一個比前人所知更多、但是大腦中儲存知識更少的一群人。搜尋引擎帶給人類的影響，也非全然是正面的，例如搜尋引擎讓網路成為抄襲者的天堂。但是在〈論文抄襲記〉一文中，作者嘉納卻透過搜尋機制創造出來的抄襲偵測系統，找到了許多論文及計畫抄襲的實例。

回過頭來，在AI亦趨成熟的當下，許多人也開始反思一些潛在的危機。在〈人類要擔心機器人太聰

明？〉一文中，AI專家暨著名AI教科書作者的羅素，提出了自己的擔憂與可能的解決方案。他指出AI並不會像電影情節一般突然覺醒而屠殺人類，也不可能無端有邪惡思想來跟人類爭權奪利。但是這不表示AI不會對人類造成危害。最大的威脅可能來自AI為了達成被賦予的任務，在過程中做出了不符預期的負面行為。例如，某個機器人接收到「拿蛋糕過來」的命令，當它發現蛋糕鎖在櫃子裡時，可能會選擇打破櫃子再把它拿出來。他認為AI的系統必須被「賦予」三個觀念，以防AI執意完成任務而造成更重大的破壞。首先，AI必須實現人類的價值觀，於是會做出符合人類希望其達成的目的。其次，AI必須理解自己「認知」的人類價值觀不一定完全正確。所以它不會去執意執行人類賦予



我們正處於一個世代的轉捩點， 「機器學習與巨量資料」將賦予電腦 前所未見的智慧。

的「表象」任務（例如打破櫃子只為了拿到蛋糕）。最後，AI必須藉由觀察人的選擇來學習人類的價值觀。如果AI可以擁有以上的觀念，那它就比較沒有機會去做出危害人類之事。

除了消極的規範，也有一些學者研究如何創造出「友善」的AI。在〈機器人 懂你心〉一文中，作者馮雁談及自己專注於做出具有「同理心」的AI，它能夠分辨人類的情緒，並做出適切的反應。其中一個原型Zara，是結合語音辨識與情緒辨識的模組，利用機器學習的演算法來跟很多人互動，學會了解人類意圖，用具有同理心的方式與人類溝通。

當AI漸漸成熟，「人」要怎麼與AI共處、怎麼利用它的優勢並防範潛在危機，也成為重要的議題。我相信不久的將來，除了許多原本需要高度智慧的工作將由AI接手，AI也更會與人類攜手一起解決許多我們面臨的難題。這樣的革命並不代表人類失去了主宰權，而是結合另一種更強大的能力，成就原來無法成就的事。

台灣大學資訊工程系教授

林宇德