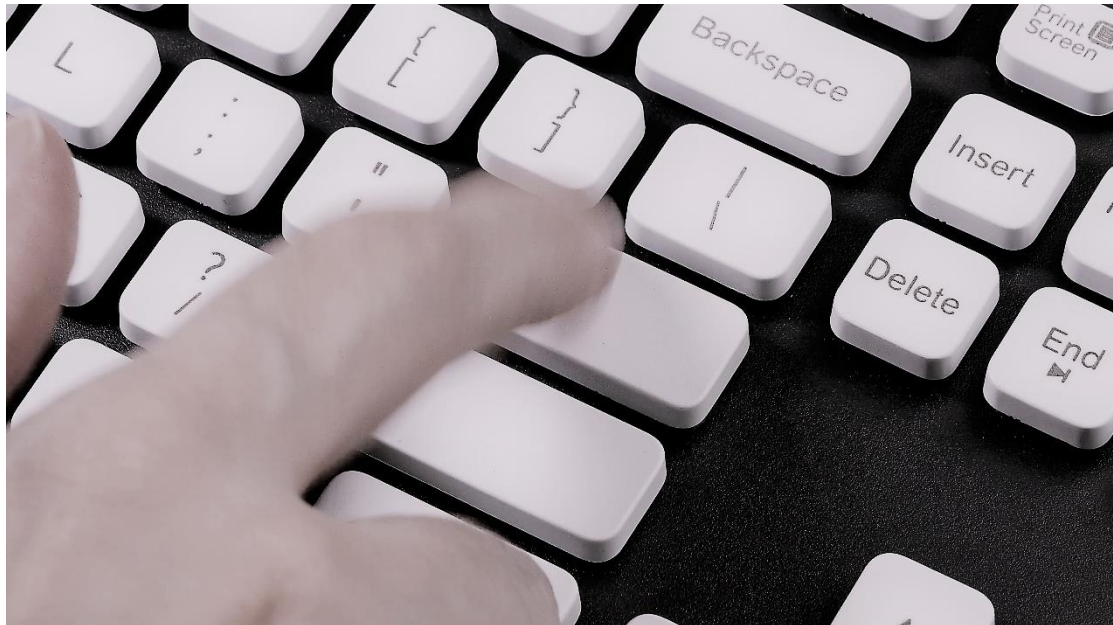


標題：AI 的新窘境：再現性危機

關鍵字：再現性、數據科學、資料挖泥



從企業決策到醫療研究，再從道路安全到預防犯罪，資料探勘的成績逐漸受到重視，但在這廣袤無邊的研究海中，難道沒有任何潛在的謬誤嗎？事實上，數不勝數。受不少粗糙的機器學習技術所害，數據科學家大量產出誤導性結果，而人們又常以這些錯誤為根基，向上堆疊出更多研究和決策，讓 AI 的不確定性加倍侵蝕產業的信心。「這真是一場潛在的風暴，」休士頓萊斯大學的 Genevera Allen 教授警告，「如果分析技術依舊停滯不前，科學家將只會不斷浪費時間和金錢。」

再現性危機

西元 2011 年，一組物理學家宣布一則駭人的發現—微中子的傳播速度竟然比光子快上六十奈秒，經過六個月的重複驗證，仍然得到這樣怪異的結果；然而，研究人員並未慶祝這項物理界的革命，反而持續相關研究以解釋這樣的異常現象。果然，他們在不久之後找出癥結，有一條光纖被錯接了。

這個例子提醒我們「再現性」對於科學的重要性，一個穩固的研究需要經過反覆實驗或多方驗證的考核，才能保證品質的可信度。畢竟，脆弱的歷史發現並無法提供足夠的前瞻性，就像在劣質的地基上蓋房子，可能將未來的努力付諸流水，需要承擔龐大的資源損失，甚至會犧牲大眾的安全與健康。

AI 產業目前遇到的麻煩，正是「再現性危機」。自動車上路想變換車道，深度強

化學習(RL)會在虛擬環境中摸索成千上萬種可能性，直到安全完成任務；但這些嘗試就像被鎖在 AI 的黑箱子中，實驗因此難以再現，不知該如何向人們交代自動車技術的可靠性。

Google 研究員 Ali Rahimi 曾將 AI 科技比為煉金術(alchemy)，不斷的試錯，不斷的「學習」，對於演算法的建構與運作缺乏清晰的解釋，又難以所謂的「再現性」支持立場。這樣，我們還能稱 AI 為科學嗎？

AI 的難處

事實上，AI「再現性危機」的發生，是有跡可循的。

在數據科學的領域中，其實沒有一個擷取、淘選及處理資料的管理標準，這就像是氣候變遷的研究中，如果沒有位置座標和氣溫單位一樣，每份報告可能會有著獨特的自訂原則，紀錄方式也因而大相逕庭。

當實驗相當費工，卻欠缺統一的紀錄模板時，「再現性」便受到挑戰。以臉部辨識系統為例，光是機器訓練、軟體的升級和檔案為演算法的修正，步驟就如千層蛋糕般複雜，必須被細緻地記錄與呈現，否則將侷限研究的可讀性，影響未來修改和再製 AI 的可能。

另外，在科學研究中，我們往往會根據觀察提出假說，再分析實驗數據來驗證推理是否成立；然而，不少 AI 實驗為了成就或利益，暴力挖掘資料之間的關係，再挑出表現較好的結果反推假說，硬將兩個毫不相干的現象兜在一起，本末倒置，違背實驗應有的初衷。

舉例來說，AI 研究人員找不出素食者和癌症之間的連結，卻意外發現紅髮的癌症患者不在少數，為了不讓機器徒勞，便將題目修改為「髮色與癌症之間的關係」，來迎合這項不合理的因果關係。這樣的作弊手段被戲稱為「資料挖泥(data dredging / fishing)」，非但是「再現性危機」的一大禍根，同時也會扭曲相關領域的發展走向。

從煉金術到真科學

不得不說，AI 確實有著翻轉社會的潛力，能為人類帶來新層次的生活型態。可是在我們能保證 AI 每項決策的品質之前，還是不能輕易託付重任給機器。

於是，為了讓 AI 躋身可靠穩健的真科學，克服「再現性危機」必為當今之務。就像其他領域的科學一樣，科技領域也需要一套明確的管理原則，才能統一繁雜的實驗記錄，提升研究的可追溯性(traceability)。首先，所有訓練資料的背景和模型參數必須被詳盡地記錄，不同資料序列的合併、分析資料的工具也要通

過標準評估；此外，還得限制自動機的各種推論，唯有剷除那些無理的弊端，才能進一步改善實驗的精準度與可信度。

Allen 教授目前正著手建造 AI 的新時代，和一群來自貝勒大學的生物醫學學者不斷錘鍊更強大的統計技術，盼望能篩查深藏在大數據中的雜訊，在降低實驗不確定性的同時，也提升其再現性。「我們或許還要投資無數的心力才能發表研究，」Allen 說著，「但這樣的成績，一定禁得起時間的考驗，並且為往後的科學家省下龐大的資金，也更能清楚掌握科學發展的未來。」

參考資料：

1. Matt Jones, "[How Do We Address The Reproducibility Crisis In Artificial Intelligence?](#)", Forbes, 26 Oct 2018
2. Pallab Ghosh, "[AAAS: Machine learning 'causing science crisis'](#)", BBC News, 16 Feb 2019
3. TED-Ed. "Is there a reproducibility crisis in science? - Matt Anticole", YouTube, 5 Dec 2016 [Video file]. [Available]: <https://www.youtube.com/watch?v=FpCrY7x5nEE>.
4. Rafi Letzter, "[Google AI Expert: Machine Learning Is No Better Than Alchemy](#)", LiveScience, 7 May 2018
5. Sam Charrington, "[The Reproducibility Crisis and Why it's Bad for AI](#)", Medium, 23 Mar 2018