

# Reading Assignment of 《AI for Science Report - chapter one》

Cheng-Che Tung 106031529

Department of Material Science and Engineering, National Tsing Hua University, 30013, Hsinchu, Taiwan

Argonne、Oak Ridge 和 Berkeley 三個美國國家實驗室在 2019 年陸續舉辦了四場有關人工智慧如何應用及協助科學發展的大型會議，吸引包含美國能源部（DOE）國家實驗室的 1000 多名科學家和工程師參加。一系列會議的主題是在下一個十年中人工智慧、大數據和高性能計算能夠提供怎樣的機會、啟發、將會面臨的挑戰和可能的解決方案。

《AI for Science Report》是這些會議的總結及精華報告，當中涵蓋了 16 個科學研究主題，諸多團隊總結最新技術、挑戰、基礎設施及未來十年利用人工智慧推進科學方法的規劃。本書第一章講述如何在化學、材料及奈米科學中利用人工智慧及高通量運算達成突破。

Artificial Intelligence | Big Data | High-Performance Computing

現今的科技發展及工業技術進程有賴於過去無數前人的努力及累積，但這樣的模式在科技日新月異的當代可能略嫌緩慢；技術的發展仰賴於作為基礎的材料不斷在性能或是設計上不斷提升，而新興的技術又可以加速材料的發展，透過這樣的相輔相成，當今快速發展的趨勢是過去前所未見的。文中簡單提到了現在的手機使了多達 75 種元素，雖然使用元素並非無窮，但是其設計空間一旦包括了配比，就會有接近無限大的可能性；由於如此的複雜度，且還要在沒有明確已知理論的情況下探索未知的化學、材料，人工智慧就非常適合應用在這樣的領域。將人工智慧作為科學研究的通用工具可以加速科學發現，包括：識別新材料及化學物質、新的合成與反應途徑<sup>1</sup>。

本章節提到以傳統方式：根據直覺設計實驗⇒發現規律⇒設計模型⇒實驗和模擬來研發新材料雖

然精確卻非常的耗時；與其如此大海撈針，科學界現在正開始利用各種機器學習方法來探勘隱藏起來的複雜信息和數據，文章中提出以機器學習加速電池開發，將 14 年的開發時程加速到不到 5 年完成；同時利用機器學習在影像識別的長處，可以提供科學家即時的影像分析及模擬並加速決策。隨著高性能計算的增長，科學家透過將高性能運算叢集跟大型數據庫串聯，已經可以在虛擬的設計空間中探索材料和化合物；根據材料基因組計畫所建置的硬體及建立的資料庫，可通過理論驅動的方法進行高通量的可能性篩選。

當然，這樣的新型研究方式有其優點也會有其挑戰，例如：如何以最有效的方式搜索選定的空間？設計空間的選定有其合理的規則嗎？機器學習的模型是否有跨空間和時間尺度推論的能力？...等問題，但最重要的是，現實世界遠比用於建立資料庫的簡化模型複雜得多，而且目前對材料系統進行仿真研究的模擬仍然非常昂貴，如何解決這些問題才能讓這類型的研究方法得到實際的進步。文章中也開展了對未來的展望，希望以合理的方式生成科學數據並產生新的物理模型和理論見解，從而推動材料和化學物質合理設計的新路徑；從製備、數據解釋、反饋進行集成以最大地減少實際上所需的實驗（以減少時間），但這樣的過程仍需科學家參與以確保結果的可行性。

本文章提出了一些非常吸引人的預期結果，如：材料的多樣性可能超出熱力學或我們目前所知設計規則的極限、實現多功能和自我修復的材料系統、對計算系統間的接口進行優化、設計動態材料、提升發現未知材料和複雜化合物的速度。若這些結果真的可以應用於材料開發中，那就不單純是技術的進步，而是整個科學典範的轉移。

<sup>1</sup> Sumpter, B. G., Vasudevan, R. K., Potok, T., Kalinin, S. V., A bridge for accelerating