

# AIGC 對於臺灣教育現場的挑戰

吳育龍

國立臺中教育大學數位內容科技學系副教授兼任教育資通創新學術研究中心主任

## 一、前言

在當今數位化時代，人工智慧生成內容（Artificial Intelligence Generated Content, AIGC）的技術日益成為世人的焦點。此技術透過人工智慧（AI）演算法產生多樣化的內容，形式橫跨文字、圖像、動畫、3D 模型、音效、音樂及語音等多媒體類型。尤其是 OpenAI 於 2022 年 11 月基於 GPT-3.5 演算法推出的 ChatGPT 產品，讓 AI 在全世界受到前所未有的關注，其應用也發展至日常生活中的各個層面。GPT (Generative Pre-trained Transformer) 屬於大型語言模型 (Large Language Model, LLM)，能夠表現出類似人類認知領域的多種能力，包括記憶、理解、應用、分析、創造與評鑑 (Masia & Bloom, 1964)。此外，OpenAI 在 2021 年發布的 DALL-E 模型，以及其後於 2024 年推出的能夠基於文字描述生成超過一分鐘動畫的 Sora 模型，AI 在各個領域不斷有新的突破發展，推動該領域進入一個競爭激烈的新時代。

然而 AI 技術的發展速度超出了人類的預期，相關的素養教育和法律規章卻未能及時跟進，導致了諸多糾紛和誤解。例如，利用 AIGC 技術生成的藝術作品在比賽中獲獎，或將 AIGC 視為免費的醫療諮詢服務等情形。教育領域的應用同樣處於摸索階段，教師如何將 AI 工具與現有教學內容有效結合，以及學生是否應該被允許使用 AI 工具完成作業或報告等問題，都是當前亟需解決的課題。

值得關注的是，LLM 特有的「AI 幻覺」(Artificial Hallucination) 現象可能導致 AIGC 產生不實的回應，給教學應用帶來困擾。這種現象突顯了在使用 AIGC 工具時，無法預測何時會產生錯誤或虛假知識的問題，需要配合媒體識讀能力的培養，以引導學生正確使用 AIGC 工具。

## 二、AIGC 發展現況

AIGC 的發展可以概分為 2010 年前的預備期、2010~2022 年的萌芽期、2022 年後的爆發期 (Wang, Y., Pan, Y., Yan, M., Su, Z., & Luan, T. H., 2023)。預備期受限於當時的硬體性能不足及計算成本過高，難以形成可行的商業模式。這個階段最具代表性的指標成果就是 1997 年 IBM 深藍 (Deep Blue) 在西洋棋打敗人類棋王卡斯帕羅夫。

萌芽期在 Goodfellow 等人 (2014) 提出了對抗生成網路 (generative adversarial network, GAN) 的演算法，能夠更有效率地找出更好的解決方案。這個階段最

具代表性的指標成果就是 2017 年 Google AlphaGo 在圍棋打敗人類棋王柯潔，自此所有棋類遊戲都由 AI 打敗當代的人類棋王。2021 年 OpenAI 推出的 DALL-E 的文字驅動生成影像的網站服務後，AIGC 在影像的應用有了突破性的發展，甚至可以產生以假亂真的影像，例如德國攝影師 Boris Eldagsen 利用 DALL-E2 產生的圖像參加 SONY 世界攝影競賽獲得優勝的爭議事件，證明 AIGC 在影像的生成品質已接近或者超過大多數人類創作者的能力。

進入 2022 年的爆發期，各種媒體形式的 AIGC 工具迅速發展，能夠生成包括文字、圖像、動畫、音效及音樂在內的多媒體內容，幾乎涵蓋大部分數位創作的表現技法。

在 Netflix 播出的《The Dog and the Boy》，在製作時使用 AIGC 工具來產生動畫的背景以及背景音樂，雖然還不是主要的製作工具，但可以看到 AIGC 產出內容的品質與精緻度已經可以達到商業產品的水準（Gao, R., 2023）。

此外，哈佛大學的 CS50 課程與 OpenAI 合作，推出了基於 ChatGPT 的 AI 程式家教 CS50 ddb (<https://cs50.ai>)。該平台不直接提供答案，而是引導學生進行思考，從而幫助學生自行完成程式碼的撰寫，展示了 AIGC 在教育領域應用的潛力與實際效益。

### 三、AIGC 在教育應用的政策支持

負責任的人工智慧（Responsible AI）是指在倫理和道德原則下發展與使用 AI 技術。AIGC 能夠落實在教育上使用，除了科技工具的開發與使用之外，還需要政府從政策面與資源面進行盤點與整合，有了政策面與基礎建設的支持，才能夠讓產學研相關資源與人才能夠放心投入發展。

LLM 的建立是 AI 運算的基礎，但自行建置 LLM 會有需要耗費大量硬體資源與收集高品質資料集的兩個高門檻，各國為了掌握 AI 世界話語權與基礎奠基，紛紛以國家的層級投入 LLM 的建置，臺灣發展也需要由政府部門介入統籌資源來完成基礎建設。

中研院、國教院與聯發科技於 2023 年 2 月共同發表以開源語言模型 BLOOM 為基礎的繁體中文 LLM-LOOM-zh（Ennen et al., 2023），此合作模式由國教院提供高品質的繁體中文語料，聯發科技提供硬體運算資源，再加上中研院長期投入中文語意的開發與研究，在跨領域的合作下完成此次專案。國科會推動的 T.A.I.D.E.計畫 (<https://taide.tw>) 在打造臺灣專屬的「可信任的人工智慧對話引擎」（Trustworthy AI Dialogue Engine，TAIDE），在臺灣文化的基礎上，使用特有的

語言、價值觀、風俗民情等資料集，打造可信任的生成式 AI 引擎基礎模型。並於 2023 年 6 月發布了第一階段成果，以供合作團隊後續開發應用。臺大研究團隊（Lin, Y., & Chen, Y., 2023）在 2023 年 11 月發表以開源語言模型 LLaMa2 為基礎的繁體中文 LLM- Taiwan LLM，強調以臺灣在地文化語觀點為特色，使用公開的百科、新聞、小說、雜誌以及社群討論內容進行訓練。至此，臺灣在地化 LLM 已有初步的成果，可以在此基礎建設下發展產業與教育應用。但目前政府資源主要投入在比較泛用的文字 LLM 技術，其他像是影像、音樂甚至是動畫的影音多媒體生成技術較少著墨，尤其是 Open AI 在 2024 年 2 月發表 Sora 後，引起高度重視，並被視為未來影視產業的殺手級應用，是值得投入資源發展的重要應用。

在政府政策面規劃與支援的部分，行政院在 2018 年 10 月提出臺灣 AI 行動計畫（2018~2021）（<https://www.ey.gov.tw/File/3FF4BEDA21AE3937>），提出了 AI 人才衝刺、AI 領航推動、建構國際 AI 創新樞紐、法規與場域開放、產業 AI 化等五大主軸的發展目標。但 AI 科技發展一日千里，行政院在 2023 年 4 月提出了臺灣 AI 行動計畫 2.0（2023~2026）（<https://digi.nstc.gov.tw/File/7C71629D702E2D89>），調整為人才優化與擴增、技術深耕與產業發展、完善運作環境、提升國際影響力、回應人文社會議題等五大主軸的發展目標。在人才優化與擴增有提到人才培育包括高等教育、國民教育及在職/就業培訓三個面向，涵蓋體制內教育與終身教育等全年齡段的資源投入。除了政策面的支持外，法制面是整體產業與應用的發展依循的準則，人工智慧發展基本法草案目前還在國科會擬定中，預計最快要 2024 年底才會送到立法院審議，過渡期需以行政指引的方式來規範。在現今全國大規模使用中的新科技，竟然沒有一個可以遵循的規範，中間的空窗期對於相關發展與應用會有很大的限制，需要加速立法的速度。

#### 四、AIGC 在教育應用的困境與挑戰

綜上所述，AIGC 發展在各種領域都有突破與新的應用，在教學與數位學習輔助方面則還有以下關鍵問題無法克服，老師在教學使用時須謹慎小心，減少受影響的程度。

1. 智財權使用與歸屬：AIGC 的訓練和運作依賴於龐大的資料集（data set），這些資料集的來源及其使用是否需獲得原作者同意，目前尚未形成統一的共識（陳家駿，2023）。這一問題觸及了知識產權的核心，尤其是關於版權的歸屬和使用權的問題。此外使用者藉由 AIGC 產出的內容是否可主張為作者而擁有著作權，在 2022 年 2 月美國著作權局 U.S. Copyright Office（2022）拒絕給予 AIGC 畫作的著作權登記後，目前大多數的國家對於 AIGC 產生的創作多認為創作人參與程度過低，無法以創作人的身分主張擁有著作權。
2. 抄襲與原作之判別：對於 AIGC 所產出的內容越來越難以與人類創作的作品

區分（樂羽嘉，2023）。不管是小到作業，或是升學審查資料，甚至是比賽作品，都會發生單純從內容無法辨識是否為學生所完成（Habibzadeh F., 2023）。若要禁止學生在課程中使用 AIGC 工具，也很難舉證學生抄襲的來源，從而對教育評價體系的公正性和有效性構成了威脅。

3. 繁體中文與在地資訊內容：AI 訓練使用的資料集是決定 AI 如何產生內容的重要因素，根據 Open AI（2020）所公布的資料，訓練 GPT-3 使用的資料依語言分類，英文資料占比 92.6%，中文僅占比 0.099%，而繁體中文更僅有 0.0197%。較晚發布的 Meta LLaMa2（Touvron et al., 2023），英文資料占比 89.7%，中文提升到占比 0.13%，但包含簡繁中文。由如此稀少的資訊來源，很容易會發生不符臺灣現況或是社會價值觀的回答，或是以簡體中文的形式呈現。
4. 不當與虛假內容：AI 是一個中性的演算法，訓練 AI 給予什麼資料集就會產出什麼內容（福田和代，2023）。目前 AIGC 對於違法、色情、暴力或偽造虛假等不適合在校園內出現的內容，現有演算法過濾阻隔的成功率還不高。引導 AI 產出內容符合法律與道德規範的人工智慧對齊技術（AI alignment）還在發展階段（IBM, 2023）。若是老師要自行設定過濾條件時間成本過高，目前只能依賴 AIGC 服務提供商從源頭進行管控。
5. 無法檢核正確性：目前 AIGC 所使用的 LLM 演算法尚無法理解生成內容，因此很容易發生 AI 幻覺（hallucination）（Athaluri et al., 2023），產生一個看似合理但實際不正確的內容。若是學生對於 AIGC 產出的內容過於信任，反而會學習到錯誤的知識。在 Arthur 的報告中（<https://www.arthur.ai/gap-articles/hallucination-experiment>）對於現今知名的 5 種 GPT 網站服務進行分析，在回答問題的時候會出現 4%~30%不等的幻覺情況，所有的網站都會發生。因此國科會在 112 年 8 月提出「行政院及所屬機關（構）使用生成式 AI 參考指引」（<https://www.ey.gov.tw/File/CAE5B756153299FD?A=C>），其中第 5 點有提到各機關不可完全信任生成式 AI 產出之資訊。

## 五、AIGC 在教育應用的未來發展

由於 AI 發展過於迅速，在教育方面使用的規範，各大學所提出的 AI 工具使用於教學的指引都僅止於宣示的大方向性質，臺灣學術倫理教育資源中心在 2023 年 8 月發布的「大學校園因應生成式 AI 之指引及教學建議」（<https://ethics.moe.edu.tw/resource/epaper/html/21/>）中指出目前關注的 5 大議題有資訊驗證、著作權、個人資料保護與資通安全、依賴科技而缺乏自我意識、學術寫作的自我揭露。教育部在 2023 年 10 月發布的中小學數位教學指引中（[https://pads.moe.edu.tw/pads\\_front/index.php](https://pads.moe.edu.tw/pads_front/index.php)）首次正式對於教學活動中使用 AIGC 工具之提出輔助說明。

AI 技術與工具的使用只會越來越普及，教育現場也必須要回應世界發展的趨勢與需求，無法用禁用的方式限制學生的使用。教學現場的老師在導入 AIGC 工具時需要更多的支持系統協助老師進行教學，這有賴政府與學術研究單位共同合作，建立 AI 教學的生態系統。

1. 現有師資的增能：目前教學現場的教師在經過「106~109 年前瞻基礎建設國民中小學校園數位建設計畫」和「110~114 年推動中小學數位學習精進方案」，這兩個由教育部推動的重點數位學習計畫後，對於數位科技融入學習、數位科技載具的使用都有基本的觀念與基礎。對於 AIGC 導入教學的方式與做法，後續再搭配主題研習課程或是教學教案觀摩，可以讓老師能夠更了解也更願意使用。此部分已有「推動中小學數位學習精進方案」中的「教育大數據分析計畫」以及「生成式 AI 協助數位學習實驗方案」兩項小規模的計畫支持，但後續如何大量推廣與普及，還需盤點教育部與國教署現有計畫方向，整合計畫之間的橫向連結，建立更完整的教師支援配套措施。
2. 未來師資的培育：在師資培育端的部分，現有課程多已融入數位學習輔助、數位科技導入等相關的主題與培育內容，AIGC 工具對於教學現場會有跳躍式與破壞式的教學改變（宋明君，2023），不管是教學法或是教學活動的設計都有很大的影響，如何讓師資生擁有足夠的 AI 科技素養來正確的使用 AIGC 工具，對於師培課程的調整是很大的挑戰。
3. 教學策略的驗證：在教學現場導入 AIGC 工具，除了常見的創意發想與作品品質會有明顯的提升，但同時間可能會帶來學習挫折與放棄學習的副作用（郭旭展、陳逸萱，2023）。當學生意識到不管再怎麼努力學習，學習結果都無法比 AI 更快更好，學習表現無法受到肯定就會降低學習動機，進而依賴 AI 放棄學習。如何讓教學活動能夠引導學生同時兼顧 AIGC 工具的優勢以及找到人類獨有的能力與價值，是老師再使用 AIGC 工具前必須要先回答的問題。
4. 教學工具的開發：受限於 AIGC 特有的 AI 幻覺、不當與虛假內容、在地內容資料及過少等先天限制，可能會產出錯誤的答案，造成教學時知識正確性判別的衝突，需要媒體識讀能力的培育（高立芸、王俊斌，2023）。如果學生過於依賴 AIGC 工具導致於學習到錯誤的知識而產生認知偏誤（Cognitive Bias），之後要再矯正必須花費更大的心力與時間，因此教學工具是否能夠降低甚至控制前述 AIGC 的使用限制，或是透過教學活動設計事先引導學生去質疑 AIGC 產生的內容而不是直接接受，才能夠讓 AIGC 對於學習有正面的幫助。

#### 參考文獻

- 宋明君（2023）。是誰在詠唱？—漫談 AI 備課、AI 教材與 AI 作業。臺灣教育評論月刊，12(10)，87-93。

- 高立芸、王俊斌（2023）。當 AI 遇到師資培育。教育研究月刊，355，53-67。
- 陳家駿（2023）。AI 自動生成內容非人類所寫—是否可享有著作權？權利保護資格篇。取自 <https://iknow.stpi.narl.org.tw/post/Read.aspx?PostID=19616>
- 郭旭展、陳逸萱（2023）。整合 AI 數位科技於創新教育課程設計與實踐。教育研究月刊，355，16-35。
- 福田和代（2023）。Deepfake 深度偽造。奇幻基地出版社。ISBN：9786267210581
- 樂羽嘉（2023）。用 AI 繪畫到底算不算自己的作品？天下雜誌。取自 <https://www.cw.com.tw/article/5124726>
- Athaluri, S. A., Manthena, S. V., Kesapragada, V. K. M., Yarlagadda, V., Dave, T., & Duddumpudi, R. T. S. (2023). Exploring the boundaries of reality: investigating the phenomenon of artificial intelligence hallucination in scientific writing through ChatGPT references. *Cureus, 15*(4).
- Ennen, P., Hsu, P., Hsu, C., Liu, C., Wu, Y., Liao, Y., Lin, C., Shiu, D., & Ma, W. (2023). Extending the Pre-Training of BLOOM for Improved Support of Traditional Chinese: Models, Methods and Results. *ArXiv, abs/2303.04715*.
- Gao, R. (2023). AIGC Technology: Reshaping the Future of the Animation Industry. Highlights in Science, *Engineering and Technology, 56*, 148-152. Retrieved from <https://doi.org/10.54097/hset.v56i.10096>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems, 27*.
- Habibzadeh F. (2023). GPTZero Performance in Identifying Artificial Intelligence-Generated Medical Texts: A Preliminary Study. *Journal of Korean medical science, 38*(38), e319. Retrieved from <https://doi.org/10.3346/jkms.2023.38.e319>.
- IBM (2023). *What is AI alignment?* Retrieved from <https://research.ibm.com/blog/what-is-alignment-ai>.

- Lin, Y., & Chen, Y. (2023). Taiwan LLM: Bridging the Linguistic Divide with a Culturally Aligned Language Model. *ArXiv, abs/2311.17487*.
  
- Touvron, H., Martin, L., Stone, K.R., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D.M., Blecher, L., Ferrer, C.C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., Fuller, B., Gao, C., Goswami, V., Goyal, N., Hartshorn, A.S., Hosseini, S., Hou, R., Inan, H., Kardas, M., Kerkez, V., Khabsa, M., Kloumann, I.M., Korenev, A.V., Koura, P.S., Lachaux, M., Lavril, T., Lee, J., Liskovich, D., Lu, Y., Mao, Y., Martinet, X., Mihaylov, T., Mishra, P., Molybog, I., Nie, Y., Poulton, A., Reizenstein, J., Rungta, R., Saladi, K., Schelten, A., Silva, R., Smith, E.M., Subramanian, R., Tan, X., Tang, B., Taylor, R., Williams, A., Kuan, J.X., Xu, P., Yan, Z., Zarov, I., Zhang, Y., Fan, A., Kambadur, M., Narang, S., Rodriguez, A., Stojnic, R., Edunov, S., & Scialom, T. (2023). Llama 2: Open Foundation and Fine-Tuned Chat Models. *ArXiv, abs/2307.09288*.
  
- U.S. Copyright Office (2022). *Second Request for Reconsideration for Refusal to Register A Recent Entrance to Paradise*. Retrieved from <https://www.copyright.gov/rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf>
  
- Wang, Y., Pan, Y., Yan, M., Su, Z., & Luan, T. H. (2023). A Survey on ChatGPT: AI-Generated Contents, Challenges, and Solutions. *arXiv preprint arXiv:2305.18339*.

