

AI、傳播與社會專題序言

客座主編 王維菁

摘要

人工智慧已默默進入人類社會生活，但許多的社會運用其使用者並未認知到他們是與 AI 系統在進行互動，而相關的運作，也並未受到社會有關透明、問責、隱私保護及社會公平性等之嚴謹監督。而在傳播社會領域，人工智慧運用案例，也逐步在改變過往傳播模式或傳播社會理論對傳播的定義、想像與規範。

AI 帶來的傳播社會變化，看似真實，卻又具想像誇大成分，也因此人類社會遲未有釐清相關變化所帶來之社會、倫理與集體秩序挑戰之具體規模行動。但 AI 發展牽涉巨量公眾數據使用，相關運用與公共利益牽連甚深，為了讓 AI 系統符合人類社會的規範準則，將人類社會珍視的價值嵌入 AI 設計與應用過程至關重要，但在確認 AI 系統所需之道德準則時，這件事由誰來做？道德內涵由誰來決定？將是複雜的社會議題。然而，將人社思考帶入 AI 發展的相關討論，使人工智慧開發運用能夠回應人性、社會需求、與公共利益福祉，將是 AI 人社研究未來重要且主要的核心精神。

◎ 本文作者王維菁國立臺灣師範大學大眾傳播研究所教授。

◎ 聯絡方式：Email：weiching@ntnu.edu.tw；通訊處：10610 台北市和平東路一段162號
臺師大大傳所。

Special Issues Preface: Artificial Intelligence, Communication, and Society

Wei-Ching Wang

Abstract

Artificial intelligence (AI) has entered the daily life of human society. However, many social application users are not aware that they are interacting with AI systems; and the relevant operations are not subject to strict social supervision in terms of transparency, accountability, privacy protection, or social fairness. In the social communication field, AI applications have gradually changed the communication model of the past as well as the definition, imagination, and norms of communication in terms of social theory.

The social changes brought about by AI appear to be real, but also feel like imaginary exaggerations. Apparently, the human society has not yet clarified the specific scale of social, ethical, and collective-order related challenges corollary to the changes. Moreover, AI development involves the use of Big Data on the public, and the relevant applications are deeply intertwined with public interests. It is essential to incorporate the values cherished by human society into AI designs and application processes to ensure AI systems conform to the norms and standards of human society. But who will ensure the code of ethics for AI systems? Who will determine the moral connotations? These are no doubt complex social issues. Bringing human and social concerns into the AI development discourse and ensuring AI developments and applications are bound by the terms of human nature, social needs, and public welfare will become the core spirit of AI related Humanities and Social Sciences research in the future.

◎ Wei-Ching Wang is a Professor in Graduate Institute of Mass Communication at National Taiwan Normal University.

◎ Corresponding author: Wei-Ching Wang, e-mail: tweiching@ntnu.edu.tw address: No.129-1, Section 1, Heping East Road, Taipei City 106, Taiwan.

人工智慧 (artificial intelligence, AI) 的發展已被認為是繼網際網路與數位化後，下一波影響人類社會生活各面向的重要科技創新。人工智慧是模擬能夠和人類一樣進行智慧運作之程式或系統，如感知、認知、決策、執行、學習、創造、溝通交流等複雜目標之能力 (Kaplan & Haenlein, 2019; McCarthy, 1998)。近年 AI 的實際運用已逐漸從想像實踐成社會事實，包括圖像識別、搜尋與推薦、知識問答系統、機器感官、自動化機器、自然語言理解對話、自動內容生成與智能機器人等相關運用，均已默默進入人類社會生活，許多的社會運用甚至其使用者並未認知到他們是與 AI 系統在進行互動，而相關的運作，也並未受到社會有關透明、問責、隱私保護以及社會公平性等之嚴謹監督。

而在傳播與社會領域，人工智慧的運用案例，如社交對話程式、新聞或廣告內容自動生成、閱聽眾接收資訊的程式排序與挑選、人工智慧系統主動偵測移除平台爭議或違規內容等，也逐步在改變過往傳播模式或傳播社會理論對傳播的定義、想像與規範。Artificial intelligence and life in 2030 (Stone et al., 2016) 描繪傳播娛樂領域的初步改變，隨著人工智慧的運用，媒體將具備為特定細分群體提供個性化的微分析和微服務的能力，也因此使其同時擁有了控制特定個體思想與體驗的能力，此外，先進設備與相關應用將大幅降低文化生產成本，但因人工智慧創造出的「新媒體民主」，也將使人們的文化消費傾向變得更難以預料，其對社會甚或政治的影響也將更難以判斷方向。而人類與程式及機器的交互性，透過人機互動也將發展出更多無法預期的面向與新挑戰，包括情感、社會關係、人類主體性、換位思考、人類環境節奏的改變等等。

然而人工智慧所將帶來的傳播與社會變化，看似真實，但又似乎具想像誇大成分，也因此人類社會遲遲未有釐清這些變化可能帶來之社會、倫理與集體秩序挑戰之具規模的具體行動。目前的先行者仍多以技術開發領域的自省為主，人文社會領域的參與至今仍大多缺席。Institute of Electrical and Electronics Engineers (IEEE) 年提出 Ethically aligned design v2 (2018)，檢視人工智慧可能涉及隱私侵害、歧視、技能喪失、負面經濟影響、關鍵基礎設施的安全風險、社會福祉的長期影響等。並認為 AI 系統必須能夠符合人類社會的道德價值與倫理原則，這些系統才能實現其益處，因此社會必須建立框架，指導人們認識這些技術可能造成的社會影響，並就此進行對話與討論。IEEE 認為，社會須建立 AI 的社會與政策指南，確保 AI 系統以人為本，並服

務於人類價值和倫理準則。人工智慧系統不應只是實現功能性目標與解決技術問題，而應造福人類社會，這是開發運用人工智慧系統的前提，而省思倫理議題有助於我們建立 AI 存在人類社會的理想生活方式。而人工智慧的設計、開發與應用應遵循以下原則：確保不侵害國際公認的人權、在設計與使用中優先考慮人類福祉目標、確保設計者與操作者負責任且可問責、確保人工智慧系統以透明的方式運行、將濫用的風險降到最低。因此，IEEE 敦促開發設計者將人類社會價值嵌入人工智慧系統。因若機器將作為準自主的主體參與人類社區生活，那麼 AI 主體也應遵守社區的社會規範與道德規範，但在將規範嵌入 AI 系統前，人類社會需先對系統所在社區與相關道德倫理價值有一清晰的描述及共識，如此亦使其成為一相當複雜之社會工程。

而歐盟 2020 年也發布了 White paper on artificial intelligence—A European approach to excellence and trust (European Union, 2020, February 19)，指出社會需了解使用 AI 同時帶來的機會與風險，且 AI 發展牽涉大規模公眾數據使用，相關運用與公共利益牽連甚深，需制定適當的開發標準原則並揭示人工智慧發展的基本倫理規範。AI 發展也將引發新的責任，對職業道德、研究道德、使用者道德，以及公共安全評估等形成重要挑戰。另外，當受 AI 影響之群體無法對系統決策進行檢驗或質疑時，將惡化各種形式的權力不平等，並讓受影響者處於權力相對缺失的位置上。從另一方面來說，AI 系統的科技至上主義可能削弱弱勢群體的質疑權力，並賦予設計者更多定義道德行為的權力，並扮演某種支配性角色，例如網路社群平台人工智慧言論查核目前所正在做的事。

為了讓 AI 的設計、開發與使用符合人類社會生活的道德、價值與規範準則，將人類社會珍視的價值嵌入 AI 的設計與應用過程至關重要，但在確認人工智慧系統需要的道德準則時，重點是，這件事由誰來做？道德準則的內涵由誰來決定？將會是複雜且重要的社會議題。

目前與傳播社會相關之人工智慧倫理道德議題可能包括如人工智慧結構化了偏見與社會不平等。打造和維護 AI 系統需要大量的計算資源與數據，但這些數據是否複製或放大現實社會中原本既有的不平等與偏見？要如何檢測甚至矯正這種符合社會現實的偏見與不平等？此外，人工智慧設計反映創建者的價值觀，在女性與少數群體參與比例偏低的狀況下，將導致 AI 技術發展缺乏多元性與包容，致使偏見無可避

免，AI 發展將無法滿足不同人群的利益，而為了解決偏見、歧視與不平等問題，人工智慧開發應該有更廣泛的參與及視角，也應有檢測與導正數據偏向的措施（Kerr, A. et al., 2020）。而除了支持研究開發衡量與評估 AI 系統準確性與公平度的方法，也應建構當 AI 系統導致錯誤與傷害時的問責制度，對科技可能造成的危害進行風險控制，對科技發展帶來不可欲知社會結果，需建立監督問責制度（Doshi-Velez, F. et al., 2017）。另外，AI 系統創建過程需符合一定的透明性，對其決策結果也應負起解釋義務（Morley, J. et al., 2020），而對於使用者，必須清楚標示他們是與 AI 系統進行互動，不應去誤導或影響人類自主性、自主權之決定與行動，管制者與使用者需要有共同監督 AI 系統的權力。最後，對 AI 發展應用應進行社會影響評估，但目前社會缺乏對 AI 可能帶來之社會經濟影響評估的嚴謹作法，使相關利害關係人的影響程度與影響範圍並不明確，令社會較難以判斷 AI 系統發展對人類社會的實際福祉，但人類也需切記人工智慧風險問題根本仍在於人類自身，因為是人決定了賦予機器這樣的能力

實際上已出現的 AI 與傳播社會倫理案例，以平台運用人工智慧程式檢查並移除仇恨性言論為例來進行初步討論。目前搜尋並移除仇恨性言論已逐漸造成社會對平台人工智慧演算法可信度、透明度以及可問責性之懷疑。比如，由平台與演算法來定義仇恨性言論是否已侵犯言論自由？訓練人工智慧時所定義的仇恨言論，是採用什麼樣的標準？由誰定義的標準？相關標準之建構是否符合民主過程與社會檢驗？而移除內容的行為與侵害言論自由的界線何在？非法言論與有害言論之間的界線又在哪裡？而相關審查是否符合比例原則要件？平台對由自動化演算程式做出的決定，是否應負有解釋責任與傷害性結果的回復措施？制定相關程序前，是否訂定並公開明確的相關規則？是否應開發模擬與檢測演算法程序的評估工具，來評估人工智慧程式決策的合理性與道德性？最後，最重要的，又該是由誰來詮釋移除仇恨性言論的最終公共利益，並形成一具共識性的人工智慧社會新倫理規範？

以仇恨性言論案例來說，以上均是 AI 言論審查所必須面對的社會質疑，而監督機構運用 AI 所做的決策，也應同步考量倫理專業標準，以及廣泛考慮其對政治、社會之影響，並主動尋求利害關係人間之對話與民主公開參與，以建立多方利害關係人之動態、有彈性的共管機制與生態系統，好確保公共利益、言論自由與商業競爭能夠維持在盡量平衡的狀態。

回到本期專刊「AI、傳播與社會」的三篇論文，也都反映了在 AI 發展初期，人文社會學者對於人工智慧的省思，主要仍聚焦在倫理議題上，而倫理也是引導 AI 發展能夠符應人類社會規範、秩序甚或福祉的重要基礎。第一篇由劉湘瑤等作者著作的「人工智能倫理的挑戰與反思：文獻分析」，搜尋國內外 AI 與倫理為關鍵詞的期刊論文，並針對 2018 年至 2019 年的文獻進行綜合剖析，彙整出四個目前最常被討論的 AI 倫理議題：通用人工智能是神話還是現實？開發具有倫理思考的自主武器？AI 在醫療上能成為人類的助手？以及 AI 衝擊勞動權益關係？論文並針對此四議題進行初步的倫理治理對策討論，包括肯定自主與智慧系統研發規範的必要性、倫理價值觀嵌入 AI 系統的可行性與需求、開發偵測與預防 AI 歧視、偏見與犯罪的機制、以及人進入 AI 人機迴圈中的干預機制，運用人機合作的方式，結合可變更的運算參數，來校正修擬 AI 演算法的決策更符合社會的價值與需求。有關 AI 與倫理的議題目前仍在不斷深化與多元化，「人工智能倫理的挑戰與反思」一文引領我們進入 AI 倫理之領域，也提醒我們注意倫理的思考若未追上科技開發腳步，科技極可能會發展出社會不可欲的後果。

第二篇論文為林思平的「人工智慧與人機傳播：傳播的擴充與弔詭」，本文旨在探討人類與人工智慧的關係，從不同的人與機器關係理論演變著手，包括早期的技術工具主義，乃是以人類為中心的技術概念，認為機器是作為工具事物，而技術是達到目的之方法。但作為一準自主科技之人工智慧其發展可能挑戰顛覆了這樣的想法，作者討論了包括能夠和人類交談的電腦機器、會交談且會自主學習的電腦機器、以及在棋弈中打敗人類的電腦機器，這些機器可能不再能被單純視為工具事物，因為機器已被人類設計成為具有獨立行為或自主決定的能力，此時機器可能得以超越人類控制，成為超越人們能夠回應或為它們負責的主體。文中因此引入人機溝通、人機傳播、電腦中介傳播與媒介理論，特別是德國媒介理論代表基德勒，拓展媒介的物質性意涵，翻轉媒介只是工具的思維，認為其亦是人類理解世界的框架與行為模式。人工智慧的發展下，技術不僅只是工具、管道或媒介，也已經成為具備互動性與能動性的傳播主體。人工智慧打破人類中心主義，進入非人類中心主義，也讓我們重新思考人類與機器的關係。

第三篇論文黃昭謀的「AI 看你玩：串流遊戲平台的演算法治理與遊戲工人勞

動」，從人工智慧與演算法在傳播領域的實際案例出發，討論人工智慧與演算法下的傳播秩序樣貌。本文以近年熱門的監控資本主義理論為基底，分析串流遊戲業者如何以演算法監控並開採遊戲工人的遊戲勞動剩餘，並發展出串流平台、遊戲開發商、與玩家之間多重交錯的監視與權力不平等的環境。最後作者提出 Fuchs 所倡議的共享社會，希望藉由數位共享、平台合作、與公共服務網絡等，來達到經濟共享、政治共享、與文化共享的社會目標。

上述三篇論文都是人工智慧在台灣傳播社會領域討論與對話的重要開始，人工智慧科技發展之路方興未艾，傳播社會研究又該如何與 AI 產生連結？首先我們可以思考運用人工智慧來協助解決傳播與社會問題，或是探究 AI 發展可能衍生的傳播社會問題，並逐步建構人工智慧研究開發所需的社會倫理框架與法律規範，當然，有朝一日，AI 也可能成為人文社會科學的一種研究方法，此乃端賴研究者各種無拘無束的想像。最後，當然最重要的依舊是，我們需要將人文社會思考帶入 AI 發展的相關討論，使人工智慧開發運用能夠回應人性、社會需求、與總體社會福祉，此亦是 AI 人社研究未來重要且主要的核心精神。

參考文獻

- Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., ... & Wood, A. (2017). Accountability of AI under the law: The role of explanation. arXiv preprint arXiv:1711.01134.
- European Union. (2020, February 19). White paper on artificial intelligence—A European approach to excellence and trust, COM (2020)65 final. Retrieved from https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
- IEEE Global Initiative. (2018). Ethically aligned design, v2. 0.
- Kerr, A., Barry, M., & Kelleher, J. D. (2020). Expectations of artificial intelligence and the performativity of ethics: Implications for communication governance. *Big Data & Society*, 7(1), 2053951720915939.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25.
- Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and engineering ethics*, 26(4), 2141-2168.
- McCarthy, J. (1998). What is artificial intelligence?
- Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., ... & Teller, A. (2016). Artificial intelligence and life in 2030: the one hundred year study on artificial intelligence.