

前言：AI 科技與生命倫理學

Introduction: AI Technologies and Bioethics

張 穎

Ellen Y. Zhang

人工智慧 (AI) 對醫學-生命倫理學的影響是當今不可忽視的話題，亦是頗有爭議的話題。自 2007 年以來，美國 IBM 的華生醫生 (Dr. Watson) 通過認知運算 (Cognitive Computing)，從事疾病的早期發現、治療方法及新藥研發等工作。近幾年，AI 科技得到突飛猛進的發展，其機器學習演算法 (machine learning algorithms) 持續對人類社會帶來前所未有的改變。AI 工具如 OpenAI 公司的 ChatGPT、Sora，微軟的“必應” (Bing)，穀歌的“巴德” (Bard) 以及醫療大模型 (PaLM 2)。高級自然語言處理模型的出現，掀起了通過人工智慧聊天機器人收集資訊、回答醫療問題的新浪潮。美國加州大學聖地亞哥分校神經生物學教授泰倫斯·塞諾夫斯基 (Terrence J. Sejnowski) 在一次有關人工智慧未來展望的會議上明確指出，“人工智慧的最新進展似乎是革命性的，但我們才剛剛

張 穎，澳門大學哲學及宗教研究系教授，中國澳門；香港浸會大學應用倫理學研究中心研究員，中國香港。

Ellen Y. Zhang, Professor, Department of Philosophy and Religious Studies, University of Macau, Macau, China; Research Fellow, Centre for Applied Ethics, Hong Kong Baptist University, Hong Kong, China.

《中外醫學哲學》XXII:1 (2024 年)：頁 1-11。

International Journal of Chinese & Comparative Philosophy of Medicine 22:1 (2024), pp. 1-11.

© Copyright 2024 by Global Scholarly Publications.

起步——未來超出了我們有限的想像範圍。” (Sejnowski, “GPT: From Brains to the Talking Dog”)。¹ 目前人工智慧在中國的發展亦是與日俱增，其未來的發展問題以及其背後的倫理思考愈來愈受到學者的關注，是否運用 AI 技術建構的人工智慧系統的問題已經轉向如何運用 AI 技術建構的人工智慧系統的問題。

2018 年 6 月，美國醫學會 (AMA) 發佈了首個關於如何在醫學領域開發、使用和調節人工智慧的指南。值得注意的是，該協會將人工智慧稱為「augmented intelligence」（增強智慧），而非我們廣泛認為的「artificial intelligence」。這一指南聲明表明了一個根本的問題，即人類對通用[超級]人工智慧 (AGI) 未來發展的擔心。如果說今天的 AI 技術還是在一組預先確定的參數內執行，那麼 AGI 開發的是人工智慧具有自主自控能力，因而有可能解決在建立它時沒教給它的設定和情境中的複雜問題。由此，我們不能不擔心，超級人工智慧是否為了維繫自身的存在，而不再扮演僅僅輔助人類工作的角色。另外，機器人在 AI 系統中如何展現感官感知和物理操作功能？還有其他一系列有關 AI 的倫理議題，如隱私問題、數據處理問題、知情權、自主權等。還有，誰來處理技術本後的道德倫理問題？還有更深層的哲學問題，如技術只是輔助人類達到謀殺目的之工具，還是改變人類認識自我與世界的方式？本期的四篇文章從不同的角度思考現代科技對醫學-生命倫理學的挑战。

陳子瑜、程國斌的文章〈“仁術”概念與“仁”的價值實現——醫用護理機器人倫理設計反思〉涉及的是護理機器人 (care robots) 的議題。這是一個近期較受熱議的話題，尤其是機器人作為老年護理的話題。譬如，德國 Liectroux 機器人可以檢測老人的健康狀況、帶領老人去洗手間、抱起老人從床上移入座位等。英國 Care-O-bot 3 機器人的目的是照顧那些行走不便以及孤單的老

(1) 參見〈人工智慧的未來就在眼前〉，美國大學新聞 (petbyus.com)，2023 年 11 月 2 日。 <https://petbyus.com/52193/>。

年人，澳大利亞 Zora 機器人可以做運動、跳舞、讀書、講笑話，並基於語音辨識技術實現與老年人的言語交流。夏普迷你型 RoBoHon 機器人，除了幫助老人通話、上網、發短信、自拍等，還能坐下站立、行走跳舞及與人交流。日本築波大學的研發的混合助力肢體 (HAL)，透過神經系統的感應，可以感知老人的肢體運動的需求。美國 Luvozo 公司的 SAM 看門機器人能夠通過自動導航、遠端監控和摔跤風險檢測系統，為住戶長期提供頻繁的查房以及非藥物類的護理服務。中國護理機器人，包括人形機器人 (humanoid machines) 的研發亦引起社會廣泛的注意，如 2016 年杭州市社會福利中心一批名為“阿鐵”的智慧型機器人，自這些護理機器人亮相以來，有關使用護理機器人的討論一直未停。據《人形機器人產業研究報告》預測，2024 年，中國人形機器人市場規模將達到約 27.6 億元；到 2029 年達到 750 億元，將佔世界總量的 32.7%，佔比位居世界第一；到 2035 年，市場規模有望達到 3000 億元。²

面對上述的問題，文章作者認為，醫用護理機器人是面向醫療護理實踐的輔助工具，其產生與發展離不開護理實踐的現實需求與機器人產業的技術進步。在機器研發和實踐的過程中，愈來愈多的學者意識到倫理治理的重要性，在此現實下，是否應將道德價值嵌入機器內部的機器人倫理設計得到了學界的關注，但與此同時，如何通過倫理設計來確保醫用護理機器人對道德價值的實現，以實現其對現實醫療護理實踐的輔助與促進作用，使得醫療護理實踐實現其初衷與價值追求。美國學者大衛·鄧肯思 (David Danks) 曾提出在科技上的“踐倫理學” (translational ethics)，即主張“倫理 AI” (ethical AI) 而不僅僅是“AI 倫理” (AI ethics)。(Danks 2021) 倫理 AI 更強調是將道德倫理價值嵌入機器內部，即在設計過程中已經考慮道德倫理價值。這也是文章作者所提出的

(2) 參見〈人形機器人不斷“進化”，我國這一產業正處於“井噴前夜”〉一文。引自新華報業網 (xhby.net)，2024 年 8 月 28 日。
<https://www.xhby.net/content/s66ce76aee4b00cab55aa8e98.html>。

“倫理設計”，也就是說，將人類的道德價值觀與倫理規範嵌入機器的系統中。這種思考也讓我們想到目前較為流行的“超級對齊”(superalignment)的說法。³但需要指出的是，這種“倫理先行”的主張也受到一定的質疑。

就儒家而言，“仁術”概念與“仁”的價值實現是儒家醫學倫理學的核心。文章指出，“仁”的概念對於長期受到儒家價值觀影響的中國社會而言，本就是一種公共道德，同時，由於“仁”是人的最高德性，並在實踐中具有不同層面的內容、意義和價值。儒家的“醫乃仁術”思想中的“術”不只是“技巧”(craft 或 skill)，而是與“心”相聯繫的“情感”。“醫乃仁術”得以成立的根本原因亦在於作為“仁術”之發端的“仁”。作為“仁術”的“醫”必然是儒家道德理想“仁”的實現方式，其終極價值也必然指向於“仁”的德性。就“仁”與“術”的關係而言，可以看成“體”與“用”的關係，仁為義理層面的體，術為實踐層面的用。這一原則可以運用於護理機器人的倫理設計，使得道德可以在一定程度上被“物化”。

然而，這裡會引出另一個問題，護理機器人是否可以成為擁有人格的主體？在這個背後還有一個更大的哲學問題，這就是人工智慧具有人格(personhood)嗎？如果回答是肯定的，我們又如何能在法律與倫理上擬制人工智慧所具有的人格權？目前法律規範中的權利主體包括自然人與法人，法人也並非人類，但卻由法律擬制具有權利能力，從而可成為適格的權利主體。光是擁有智慧並不能被認為是人。當機器人在護理過程中出現錯誤，是否要負責任？如果AI不是人，如何承擔在責任？所謂“負責人工智慧”(responsible AI)又如何理解？再者，作為演算與數據控制的決定，護理機器人的設計者是否應該被看作道德的主體(moral

(3) Open AI 為了對應 AGI 可能帶來的失控危險，提出一個“超級對齊”的方案。在 AI 領域中，對齊(alignment)意指確保人工智慧系統的行為，符合人類設定的目標和道德標準。而超級對齊(superalignment)是指研究更為前瞻的議題，即針對未來可能出現超越人類智慧的 AI 系統有所控制。

agent)？在西方規範了倫理學中，道德行為的對與錯是基於道德主體的能動行為，也就是道德的能動者。再之，護理機器人的內在價值與 (intrinsic value) 和工具價值 (instrumental value) 的關係，是否會因為機器人的能動者身分而有所改變。

本期的第二篇文章是何禹平的〈人工智慧助力抑鬱症治療：道家視角〉。文章指出，近些年，全球抑鬱症患者顯著增加，傳統線下心理治療面臨可及性、便利性和社會污名化等挑戰，已成為全球公共衛生領域的關注焦點。在此背景下，人工智慧 (AI) 聊天機器人的快速發展為其提供了新的解決方案，但 AI 聊天機器人 (woebots, chatbots) 仍面臨患者隱私信息保護、演算法偏見、人機過度依賴及責任歸屬等倫理問題。一方面，隨著 AI 技術的迅猛發展和應用，AI 聊天機器人作為治療抑鬱症患者的新興工具，有效降低了成本，提高了治療的及時性，但另一方面，聊天機器人是否能與人建立親密的關係，以真正地達到治療抑鬱症患者的目的也是備受質疑的。

毫無疑問，高科技諸如 ChatGPT 和 Google Bard 等大型語言模型的聊天機器人，展現了跨多學科領域處理知識，查詢資訊等各種能力。但聊天機器人是否可以與人建立某種健康的人際關係，這裡存在許多不確定的因素。即便是社交工具，如 Soulmate AI、Paradox、XiaoIce，以及 Kuki，也受到不少社會學家和心理學家的質疑。這裡存在著一個不可忽視的轉變，即從 HHR (human-human relationship，人與人的關係) 轉向 HMR (human-machine relationships，人與機器的關係)，或者 HCR (human-chatbot relationships，人與聊天機器人的關係)。問題是，人與聊天機器人的關係是否只是一種幻覺？文章作者指出了聊天機器人所存在的潛在的道德風險，如數據隱私、演算法偏見、責任歸屬等問題。就人機關係，作者提及過於依賴的風險，並從道家哲學的視角給予具體的分析。

道家的視角是一個有趣的話題。傳統上，道家被看成是一個“反智”和“反技術”的傳統。《老子》中“道法自然”被解釋為“自-然”與“他-然”的矛盾；《莊子·天地》中指出“有機械者必有機事，有機事者必有機心。機心存於胸中，則純白不備；純白不備，則神生不定；神生不定者，道之所不載也。”莊子的“機心說”常被引用作為道家“反技術”的論據。但正如文章作者所指出的，道家並非對技術持有絕對的否定態度，而是反對技術作為“工具”對人的異化，即“機心”陷入了一種封閉式循環，最終導致“機心”對“道”的遮蔽。實際上，道家的觀點是“道”與“技”統一。用莊子的話，我們對技術應該採取“物物而不要物於物”的態度。也就是說，我們可以使用技術，但不要被技術所束縛。作者認為，我們需要優化 AI 技術在醫學領域的解決方案，同時需要在倫理和人文關懷方面取得進步，讓 AI 技術發展真正服務於人類福祉。

本期的第三篇文章是肖巍、高然、陳珊的〈人工智慧時代醫患關係的“關懷”倫理構建〉。作者的探討著重在醫患關係以及 AI 技術對醫患關係的影響。文章指出，AI 科技的高速發展和應用對傳統醫患關係提出新挑戰，譬如過度依賴 AI 的診斷和建議、患者的隱私權對醫生的信任程度等等。作者認為，AI 時代醫患關係把以往的醫生與患者 (Physician-Patient, PP) 關係變成 AI-患者 (AI-Patient, AI-P)、醫生-AI-患者 (Physician-AI-Patient, P-AI-P) 等關係，使醫生與患者的互動方式、人與人交往方式發生前所未有的改變。當 AI 診療機器隔離或替代醫患之間面對面交流的情感和溫度時，如何闡釋醫患關係問題便不再是簡單的概念爭辯，人們都希望 AI 被用於促進以人為本的醫患關係。那麼，如何應對 AI 的挑戰？文章作者提出了“關懷倫理學” (care ethics)。關懷倫理學在西方已經流行幾個世紀，其強調關懷的兩個主要目的：防止傷害和提高幸福感，前者側重身體的層面，後者側重精神的層

面。⁴ 另外，關懷倫理學的本質是基於關係之本體」(relational ontology)，而非西方啟蒙時代所強調的自主的個體。在這一點上，關懷倫理學與儒家的倫理思想有一定的相似性。

針對目前內地的醫患關係的爭論，文章指出了學界的兩種觀點：一種觀點認為醫患關係是一種合同關係，患者可以選擇醫院和醫生；另一種觀點認為醫患關係不僅是一種合同關係，而且是一種消費關係，醫患關係是按照市場規律來操作。但也有學者持有不同的觀點，他們認為把醫患關係視為消費者和經營者的關係是不公正的，因為醫生本職是解除患者的痛苦，醫患關係不是單純的經濟利益關係。就此，作者認為，醫患關係應是一種關懷關係。文章指出，關懷倫理學把醫患關係視為關懷關係的根據在於三個方面：其一、醫患關係是圍繞著生命場域，特別是人的生命場域形成的一種倫理關係；其二、醫患關係需要以“關懷”作為媒介進行溝通；其三、“關懷”是醫生的職業美德。這三個方面，都有助於消解 AI 對醫患關係所產生的負面作用，為醫患關係的道德構建提供一條路徑。

然而，與傳統的關懷倫理學不同，文章作者不認為對“關懷”的強調意味著減低自主性 (autonomy)，認為關懷是基於關係自主性的構建。作者採用美國哲學家安樂哲 (Roger Ames) 有關“角色倫理學” (role ethics) 的觀點說明關懷與自主性並不矛盾。

首先，自主性並非是傳統意義上的患者或者醫生單方面的自主，而是以關懷為媒介的雙方的自主。由此，關係自主性是聯繫而不是分離為特徵的。再者，關係自主性不僅要考察患者的理性決策能力，更需要考察他們的境遇、感覺、體驗和情感因素，因此關係自主性有助於消解醫患關係中的二元對立。

本期的最後一篇文章是楊慶峰的〈數字復活——記憶哲學審視〉。“數字復活” (digital revival 或 digital resurrection) 及“死亡科技” (thanatechnology) 是近年較為熱門的話題。作為一種特殊的

(4) 就關懷倫理的具體的討論，可參見 (Engster 2007)。

記憶形式，數字復活透過 AI 技術跨越生死界限，改變人的記憶，甚至造成“數字永生”(digital immortality)的幻像。作者把討論的重點放在記憶倫理的議題上，提出 AI 智慧體/數位人所引發的記憶主體的困惑。文章指出，從字面上看，“復活”與“死亡”相對，意味著重新讓死者獲得生命。就數字復活而言，首先資訊是復活的基礎，其二是人格的重構，但復活並不是單純的人格重構，而是同一人格的再現。這裡，作者將復活區分為兩種：一種含義是文化意義上的，另一種含義是增強意義上的，即生者採取技術方法實現逝者的生命延續。

的確，伴隨 AI 虛擬化神技術的崛起，人們可以讓逝者得到數字化的保存，亦通過與逝者的互動，延續與已故親人的交流。虛擬的化身利用生成式 AI 技術，通過學習逝者生前的錄音、資訊、趣聞軼事等資料來生成內容。數字永生技術甚至可以基於逝者在網上留下的資訊，生成一個虛擬化身或聊天機器人，供人們與逝者交流，以這種方式“延續生命”。譬如，2018年，美國記者詹姆士·維拉豪斯 (James Vlahos) 開發了一個聊天機器人，經過訓練後，說話口吻與其因癌症去世的父親一模一樣。在父親生前，維拉豪斯對父親進行了採訪，並全程錄音。不久，維拉豪斯與他人共同創辦了一個“互動式記憶應用”平台，稱之為 HereAfter AI。⁵

值得注意的是，作者在文章中區分了數位復活的兩種含義：一是死者的復活；二是生者的復活。前者的數位復活常常是數位替代，因為死者無法真實再存在。而生者的復活意味著復活的物件是未死的自我，以及這一自我想獲得永生的願望。在一個虛擬世界中存在，AI 是主要角色。就此，文章作者提出了數位人及記憶主體的問題。從記憶哲學的角度，作者對 AI 記憶、數位記憶、

(5) 參見網站 James Vlahos of Hereafter AI Discusses Avatars That Preserve the Memories and Thoughts of Real People—The Voicebot Podcast '2024年8月15日。<https://podcasts.apple.com/ie/podcast/james-vlahos-of-hereafter-ai-discusses-avatars-that/id1264180365?i=1000491909442>.

記憶歸屬進行探討。如果說記憶有三重歸因假說，及自身、親者和他者，那麼“數字復活”中的記憶主體是誰？是自身、親者、還是另一個他者？作者之所以糾結數字記憶的歸屬問題，是因為這個問題引發了 AI 技術對人與自我、與他者（逝者或虛擬的數位人）關係的轉變。這個數位人是真實的，還是讓·布西亞 (Jean Baudrillard) 所稱的“超真實”(hyperreality)？作者認為，雖然記憶歸因於親者意味著我們負有必須記住親者的責任。但在現實生活中親者沒有責任記住數位人，數位人也無法納入家族記憶的譜系中。由此可見，真實世界與虛擬世界之間缺乏相互嵌入的維度，因為這裡涉及數字記憶的歸屬問題。文章最後的結論是，數字復活是一個非常複雜的倫理問題，因而容易引發社會爭議。

其實，作為死亡科技“數字復活”還帶出其他一系列的倫理問題，諸如親人（逝者）是否同意復活、用戶隱私、用戶被操縱的危險，以及這項技術對哀悼和追思的影響。譬如，在某些情況下，孩子可能會獲得其父母或祖父母的同意來創建人格模擬，但在其他情況下則可能沒有獲取這種同意。此處出現的問題涉及逝者的執行人提供同意的有效性，執行人可能被視為逝者的適當代理決策者，尤其使用逝者/親者留下的文字和圖像創建人格時。因此，我們需要問：“數字復活”的邊界何在？再者，作為數字虛擬化的人格，可以產生讓死人“復活”的幻覺，而這種幻覺會直接導致人們對生死的看法。眾所周知，親友的離世，尤其是由於意外事件導致突然離世，會給家人朋友帶來心理的創傷。有時，這種創傷很難一時癒合。那麼數字復活技術可以解決生死問題嗎？中國傳統“慎終追遠”的情懷又在哪裡？

本期的主旨是通過不同側面的研究，探討現代 AI 科技對醫學生命倫理學以及生死學的挑戰。人類與 AI 的關係，勢必是今後倫理學不能迴避的重要議題。2016 年，《全球自動控制和智慧系統倫理倡議》(The *IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems*) 已就人工智慧的諸多倫理問題提出建議。倡議

書的起草成員拉迦·查提拉 (Raja Chatila) 和約翰·哈文斯 (John Havens) 指出，倡議書旨在“為人類的福祉促進科技” (Advancing technology for humanity)。(Ferreira 2019) 2024年7月12日，歐盟官方公報發佈《人工智慧法》(Artificial Intelligence Act) 最終文本。該法案預計在8月2日生效。在中國國內，人工智慧與醫學倫理學也在這幾年得到快速的發展。譬如《中國醫學倫理學》雜誌就上述問題舉辦了多次專家研討會，並制定人工智慧在中國語境下的倫理和法規。

從更為廣義的角度來審視，AI科技帶出更根本的哲學問題，即人是什麼？人的定義是什麼？人應該如何生活？科技是否顛覆了道德根基，把人類變成的工具？讓人與人的關係更加疏離？科學技術在延長我們壽命的同時，是否也讓我們失去了自我、變得更為虛無？因此面對科技所引發的道德危機，我們應該如何適應未來的生活，尤其是數字化的生活，成為一個不可忽視的議題。⁶

(6) 就此議題，可參見 (Edmonds 2021)。

參考文獻 References

- Danks, David. “Digital Ethics as Translational Ethics.” *Applied Ethics in a Digital World*. (New York: Information Science Reference, 2021), 1-15.
- Edmonds, David. *Future Morality* (Oxford: Oxford University Press, 2021).
- Engster, Daniel. *The Heart of Justice: Care Ethics and Political Theory*. (Oxford: Oxford University Press, 2007).
- Ferreira, Maria I. A. et.al. (eds). *Robotics and Well-Being*, VI (Switzerland: Springer, 2019).