

誰來界定人工智能的“倫理邊界”——“人工智能”的道德超載

AI's Moral Overload: Who Should Define the Ethical Boundaries of AI?

劉博京

Liu Bojing

Abstract

An irreconcilable conflict between “humanistic rationality” and “technological rationality” is becoming increasingly evident. AI, as the representative of technological rationality today, is suffering from “moral overload.” Therefore, who should define the ethical boundaries of AI and who should solve the problem of moral overload have become the most important questions. This paper analyzes an article entitled “The Promise and Perils of AI in Medicine” by Robert Sparrow and Joshua Hatherley, sharing with you some views on AI in the medical field.

劉博京，溫州醫科大學國際教育學院講師，中國溫州，郵編：325035。

Liu Bojing, PhD, Lecturer, School of International Education, Wenzhou Medical University, Wenzhou, China, 325035.

《中外醫學哲學》XVII:2 (2019 年)：頁 139-143。

International Journal of Chinese & Comparative Philosophy of Medicine XVII:2 (2019), pp. 139-143.

© Copyright 2019 by Global Scholarly Publications.

人工智能對於人類而言已然不是一個陌生的詞彙，隨著技術全球化的革新和產業發展，人工智能已經被應用到了越來越多的領域當中。不論是電腦科學、金融貿易、基礎醫療、教育創新等等。人工智能彷彿已經逐漸成為了人們生活之中不可或缺的重要組成部分，在某些極特殊的領域，人工智能已經大有取代人類自身的趨勢。但是，人工智能給人類帶來的只是“福音”嗎？在人類不斷開發、研製更為高效、便捷的人工智能設備時，當人類不斷提升對於自身和整個自然界的認知之時，當一個又一個“不可能”的曾經成為現實時，或許另一種“災難”正在逐步向人類靠攏。誠然，人類需要一定程度上，實現對於自我的深層次認知，通過一定的外部手段在某些層面達到對於人類極限的突破。然而，在這樣不斷往復的過程之中，誰能夠來界定人工智能的倫理邊界呢？又或者說，人工智能本身是否應該遵從特定的“倫理”道德呢？如若答案是肯定的，那麼，人類又是否有資格和權利來定義“被”人類所創造之物“人工智能”的倫理呢？

當下，“人工智能”已經被人們主觀性的賦予了過多的意義和概念，似乎只要一提起“人工智能”浮現在人類面前的就是靈活的機械“手臂”，繁雜的“計算公式”，難以捉摸的“巨大”機器。任何無法通過人類自身獨立完成的過程，都成為了“人工智能”需要解決的問題，精準的醫學手術需要“人工智能”的操作；信用卡的交易需要人工智能來完成；基於大資料背景的資訊需要人工智能來收集。但是，當我們給“人工智能”自身提供這樣“工作”的平台時，實際上已經先驗的將人類的“預期情感”投入了其中，這就使得“人工智能”需要按照人類社會的道德規範、按照人類意識來實現其自身的價值所在。如史派羅和哈瑟利的論文〈人工智能醫學應用的前景與風險〉所指出的，“人類同時還需要考慮人工智能在塑造我們的目標的同時，將如何服務於人類。很多醫學都以最終目的作為導向性，也就是說以價值觀念本身為出發點和落腳點。”（Sparrow & Hatherley 2019）因而，

在此過程之中，“人工智能”本身並沒有“能力”、“意願”、“道德”來對其自身的活動進行規約，我們所言說的關於“人工智能”的一切，都存在一個“被動”的前提，這就是說看似具有“主觀能動性”的“人工智能”實際上所展現和進行的工作是“被動”的，是在一定程度上“違背”其主觀意願的，“人工智能”所體現的價值，是人類想要人工智能所要體現的價值，其內在的驅動力並非來自於人工智能本身，這就使得人工智能背負了過為沉重的“使命”，其所承載的已經不僅僅是對於人類社會現代化的內涵幫助，而是成為了人類未來發展的希望和根源所在，而在我們使用人工智能來改善人類社會生活時，沒有任何人、組織機構、團體等等，對“人工智能”與“人類活動”之間進行明顯地劃分，儘管兩者之間的關係看似處於一種“動態”的平衡之下，但實際上，這種所謂的平衡所體現出的正是目前“人工智能”領域的道德超載現象。就醫療領域而言，責任、信任、資訊、安全等等，都是擺在“人工智能”面前亟待解決的重要問題，我們一次又一次的強調這些問題的重要性，卻又再一次又一次慘痛教訓之後痛定思痛，一方面人類認識到人工智能的重要性，一方面又無法實現與人工智能之間“平等”有效的溝通。如果，我們僅僅單純的將“人工智能”劃分為技術理性之中的外化，未免對於“人工智能”本身的“定義”有失偏頗，當我們不厭其煩地無數次談論起人工智能所帶來的風險時，我們實際上擔心的是，在這個以“人類”作為主體的社會之中，“人文理性”和“技術理性”兩者之間的“衝突”，以及隱藏在這種“衝突”背後，無法調和的兩者之間的特定“關係”。如史派羅和哈瑟利所指出的，“技術哲學的一個重要觀點是，工具絕不僅僅是工具。他們從不“中立”。相反，它們塑造了我們的目標。工具具有“可視性”——它們使某些事情比其他事情更容易完成，並且憑藉這一事實，它們“構建”了問題。”（Sparrow & Hatherley 2019）當歷史進入實利時代，“技術理性”與“人文理性”之間的衝突已

經脫離了人類所能夠掌控的範疇，對於“人工智能”的“去人性化”就最能夠體現和說明這種悖離人類道德場域的現象。

或許，我們大可不必站在人類中心主義者的角度，來重新審視和看待“人文理性”與“技術理性”兩者之間的問題。但是，我們不得不思考的是，人類作為這個世界天地的精華，萬物的靈長是否有資格界定人工智能的倫理邊界所在？“人工智能”這一被人類創造物又應該以怎樣的“姿態”展現在人類面前，並更好地參與到人類社會的建構之中呢？正如史派羅和哈瑟利在文章中所提及的一樣：“事實上，我們很難想像機器能提供任何與這些問題相關的東西，因為機器不能有效的支撐和踐行他們的觀點，就像人們必須認真對待他們的道德建議一樣”。如果沒有辦法處理好“人文理性”與“技術理性”兩者之間的關係，那麼等待人類社會的只有價值觀念的“崩塌”，倫理道德的“失衡”。從“被”定義，到“自我”定義，對於“人工智能”而言，並非是一個漫長的過程，甚至在某些醫療領域，這種自我“自我診斷”並進行糾正的技術已經十分成熟。然而，這種“自我的定義”如何能夠在“人文理性”的規約下，實現對於自身道德場域的超越，真正地解決人工智能領域“道德超載”的問題，才是我們應該首先進行討論和研究的議題。人類社會存在的根本是人類本身，從這個層面而言，脫離人類談論任何他者事物都沒有任何意義，如果社會之中沒有人類的存在，那麼社會的存在對於人類的存在也毫無意義，“人文理性”需要高揚旗幟，“技術理性”同樣需要探尋正確的方向，人類的主觀意識在兩者之間不斷遊走的過程，才應該是我們所要解析的空間。因而，要解決“人工智能”的道德超載問題，我們首先應該從理性的角度出發，明確人工智能產生的內因，以此為基點，實現對於自身屬性和價值的外在“輸出”。或許只有這樣，從倫理道德的層面而言，人工智能才有進一步“發展”“前進”的可能。

參考文獻 References

- 羅伯特·史派羅、約書亞·哈瑟利：〈人工智能醫學應用的前景與風險〉，
《中外醫學哲學》，2019年，第XVII卷，第2期，頁79-109。Sparrow,
Robert and Joshua Hatherley. “The Promise and Perils of AI in
Medicine,” *International Journal of Chinese & Comparative Philosophy
of Medicine* XVII:2 (2019), pp.79-109.