

評論〈人工智能醫學應用的前景與風險〉一文中“責任歸咎”的疑慮與可能的因應

Possible Responses to the Problem of Responsibility

周琬琳

Chou Wanling

Abstract

The ultimate result of AI medicine may be the birth of “Dr. Super AI.” Dr. Super AI would not only supplement medical care, but also acquire a level of autonomy. Its capacity for reason would be much greater than that of most human doctors, but it would not possess certain intrinsic characteristics of human beings, so it would fall somewhere between “machine” and “human.” This raises an important question: if such an agent makes mistakes, who or what is responsible? This paper argues that this “responsibility gap” is not insurmountable. If the types of mistakes made by Dr. Super AI can be distinguished, we

周琬琳，福建醫科大學醫學人文研究中心助理教授，中國福州，郵編：350001。
Chou Wanling, Assistant Professor in the School of Humanities, Research Center for Medical Humanities, Fujian Medical University, China, 350001.

《中外醫學哲學》XVII:2 (2019 年)：頁 127-131。
International Journal of Chinese & Comparative Philosophy of Medicine XVII:2 (2019), pp. 127-131.
© Copyright 2019 by Global Scholarly Publications.

can find the corresponding object of responsibility. This is a viable countermeasure, at least until AI robots have developed human emotions and moral awareness.

1. 超級智能醫生

隨著達芬奇手術機器人的誕生與應用為醫療活動注入一股新血之後，IBM 沃森(Watson)系統助力輔助診療的出現也預告了醫學領域在不久的將來會掀起一股巨大的革命，雖然沃森目前還不是很成熟，甚至提供了錯誤的治療建議（王姍姍、翟曉梅 2019，973），但是，如果人工智能在醫學領域中的發展順利進行的話，“超級智能醫生”將會是人類臨床醫療活動中的一大躍進，或者說它將是人工智能在醫療產業中的結果。根據作者(Robert Sparrow)在其文章中的闡述，那些人工智能在醫學中的樂觀前景，在相關技術成熟的前提之下，我們將會製造出一“位”“超級智能醫生”，它擁有龐大且多個數據庫，能夠以極為快速的瀏覽搜尋能力檢視醫學文獻，為患者提供最佳的治療方案，它還可以精準地預測患者未來病情發展的軌跡。（Sparrow & Hatherley 2019）如果你能夠負擔得起費用，你可以擁有一位超級智慧家庭醫師，只要釋出你個人的生活方式等資訊，它也能針對你的基因測序、健康數據和生活方式等進行客製化醫學，包含預防、治療、甚至美體、非侵入性增強能力等。

超級智慧醫生原初被設計出來的原因可能是善意的，為了解決偏鄉地區醫療資源及人力的缺乏，或是提供進階版的居家照護人工智能設備等，但是最終可能會趨向一種滑坡，例如醫院經營高層為了減少醫護人員成本開銷引進“超級智能醫生”取代人類醫生，護理人員如果還有提供不能輕易被取代的人類特質，例如給予關懷溫暖等，只要在學會跟“超級智能醫生”“溝通”，協助安撫患者情緒可能還能保有工作。在焦慮醫生或護理人員是否能 AI 被取代的同時，我們還關心如果發生醫療事故，責任歸屬如何界定的問題。

2. 責任歸咎

是人都會犯錯，機器就不會犯錯了嗎？

人要為自己的行為負責，機器犯錯又該怎麼究責？

超級智能醫生的錯誤可區分為 3 種類型：

1. 系統設置有漏洞，造成自動化偏差。Ex. Therac-25。
2. 控制監管有疏失，忽視系統警訊。Ex. 疲勞現象警報。
3. 醫學的不確定性與特例。人體是複雜的有機體，除了基因與細胞的變異之外，微生物的影響力也不容小覷。再多的數據庫和醫學文獻都有不能窮盡之處，即便是超級智慧醫生也不可能達到 0% 的誤診率。

針對不同犯錯的類型，責任歸屬應該有不同的判定。人工智能產品在民事侵權責任如何歸屬尚無定論，這關係到人工智能產物是“物”還是“人”。第一種情況，在尚未開展出具有情感及道德意識功能的機器人之前，我們傾向以“物”來處理責任問題。主要可以從產品責任及動物管理人責任著手歸咎責任。產品責任規定如果因為產品本身的問題造成其他人的財產或人身損害，受害人有權可以向生產者請求賠償；動物管理人責任比照所飼養動物造成他人損害侵權當由飼主負責來處理。(游文亭 2018，71)

在第 1 種犯錯的類型當中，系統設置漏洞屬於產品設計不良，可採用產品責任的觀點來歸咎責任；在第 2 種類型當中，監管疏失類似於飼主疏失導致他人的損傷，可用動物管理人責任的觀點來進行究責。最後一種類型的究責則較為複雜。事實上，第 3 種類型的錯誤不是超級智慧醫生才會發生的，它同時發生在人類醫生的診療活動當中。嚴格來說，第三種類型不是犯錯，它是一種患者錯誤認知及過度期待的結果。在許多醫患衝突的研究中顯示，患者對於醫療行為的不當認知及醫生當成上帝的過度期待會在結果不如患者及家屬預期時產生醫患衝突（冼詠琪、譚一笑

2014, 56) , 包含對於醫療風險的認識不足、患者歸因偏差等因素。這類問題主要還是應該著重在醫患溝通部分, 超級智慧醫生如果被視為“物”, 表示本身尚不具備情感道德等自主意識, 患者與“醫師”的溝通可能只是指令的出入與治療方案的給出, 真正的溝通應當要往前推, 在患者自願選擇人類醫師看診或是超級智慧醫師看診前就需要進行告知同意的過程了。患者有權利知曉選擇超級智能醫生有什麼樣的風險, 並且應該了解不論是選擇人類醫生或是超級智能醫生, 都應當對於醫療風險有正確的認識: 不是同一種治療方法或藥品用在每個人身上都是可行的。並且信任醫生一定會盡力而為, 但是醫生不是上帝。

第二種情況, 超級智慧醫生某個意義是具有自主性的, 它不僅能歸納整理龐大的數據庫和醫學文獻, 還能自主推論演繹出最合適的治療方案, 1997 年 IBM 開發的“深藍”和 2016 年 Google 開發的 Alpha Go 圍棋不敗的戰績展示了人工智能的強大, AlphaGo Zero 以不基於人類學習經驗的自學習演算法更能表示出自主性。(唐振韜、邵坤、趙冬斌、朱圓恒 2017, 1) 如果說人類的責任歸咎預設了自由意志, 那麼自主性是不是自由意志的表現? 超級智慧醫生做為高級機器人, 是不是可以比照“人”的責任歸屬來看待? 我們應該不會認為自主性等同於自由意志, 畢竟機器人還是人類設計創造的, 況且, 即使把機器人視作人, 在造成重大人員傷亡的責任歸咎中, “處死”一個機器人對於受害人的心理似乎不會有太多彌補效用。但是如果把機器人視為“人類”來看待, 法人責任觀點或許可以參考。法人責任觀點認為人工智能可類比於法人擬制, 其侵權將由管理人承擔責任。(游文亭 2018, 70)

3. 結語

“責任鴻溝”(responsibility gap)並非不可突破。誠然, 超級智慧醫師可能無法作為人來獨力承擔責任, 但是只要根據不同犯

錯類型即可對相應的對象進行責任究責，也可促使製造者及管理者在人工智能醫學在開發及應用的過程中提高警示，強化監管機制。

參考文獻 References

- 王姍姍、翟曉梅：〈人工智慧醫學應用的倫理問題〉，《中國醫學倫理學》，2019年，第32卷，第8期，頁972-976。WANG Shanshan, ZHAI Xiaomei. “Ethical Issues of Applications of Artificial Intelligence in Medicine,” *Chinese Medical Ethics*, 2019, 32(8):972-976.
- 冼詠琪、譚一笑：〈我國醫患矛盾中患者因素分析〉，《醫學與社會》，2014年，第27卷，第9期，頁55-57。XIAN Yongqi. “Analysis on the Patient Causes in Doctor-patient Conflicts in China,” *Medicine and Society*, 2014, 27(09), pp.55-57.
- 唐振韜、邵坤、趙冬斌、朱圓恒：〈深度強化學習進展：從AlphaGo到AlphaGo Zero〉，《控制理論與應用》，2017年，第34卷，第12期，頁1529-1546。TANG Zhen-tao, et. al. “Recent Progress of Deep Reinforcement Learning: from AlphaGo to AlphaGo Zero,” *Control Theory & Applications*. 2017, 34(12):1529-1546.
- 游文亭：〈人工智慧民事侵權責任研究〉，《學術探索》，2018年，第12期，頁69-77。YOU Wen-ting. “Legal Study on Artificial Intelligence: Attribute and Tort Liability,” *Academic Exploration*, 2018(12):69-77.
- 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.