

# 人工智能批判的話語悖論

## The Discursive Paradox of AI Criticism

林 芬

Lin Fen

### Abstract

In the dominant discourse of the “human–machine relationship,” people and machines are the subjects, with a mutually shaping influence. However, this framework neglects the crux of the current critical analysis of AI. It reduces the problems with new technology to the relationship between people and machines, ignoring the re-shaping of the relationship between “people and people” in the era of new technology. This simplification may mislead policy and legal regulations for new technologies. Why would a robot killing cause more panic than a murder committed by a human? Why is a robot’s misdiagnosis more troubling than a doctor’s? Why do patients assume that machines make more accurate diagnoses than doctors? When a medical accident occurs, who is responsible for the mistakes of an intelligent medical system? In the framework of traditional professionalism, the relationship between doctors and patients, whether trusted or not, is based on the premise that doctors have

---

林 芬，香港城市大學媒體與傳播系副教授，中國香港。

Lin Fen, Associate Professor, Department of Media and Communication, City University of Hong Kong, Hong Kong.

《中外醫學哲學》XVII:2 (2019年)：頁 133-137。

*International Journal of Chinese & Comparative Philosophy of Medicine* XVII:2 (2019), pp. 133-137.

© Copyright 2019 by Global Scholarly Publications.

specialized knowledge that patients do not possess. Therefore, the authority of a doctor is the authority of knowledge. In the age of intelligence, do machines provide information or knowledge? Can this strengthen or weaken the authority of doctors? It is likely that in the age of intelligence, the professionalism, authority and trustworthiness of doctors require a new knowledge base. Therefore, the de-skilling of doctors is not an issue of individual doctors, but demands an update of the knowledge of the entire industry. Recognizing this, policy makers must not focus solely on the use of machines, but take a wider perspective, considering how to promote the development of doctors and coordinate the relationship between doctors with different levels of knowledge development. We often ask, "In the era of intelligence, what defines a human?" This philosophical thinking should be directed toward not only the difference between machines and people as individuals, but also how the relationship between human beings, i.e., the social nature of humans, evolves in different technological environments. In short, this commentary stresses that a "good" machine or an "evil" machine—beyond the sci-fi romance of such discourse—reflects the evolution of the relationships between people. In today's smart age, the critical issue is not the relationship between people and machines. It is how people adjust their relationships with other people as machines become necessary tools in life. In the era of intelligence, therefore, our legislation, policy and ethical discussion should resume their focus on evolutionary relationships between people.

人工智能在各個領域的應用一直是各界關心的話題。在推進智慧醫療的發展過程中，人工智能能帶來什麼好處，又會引發什麼樣的風險？Robert Sparrow 教授的這篇文章就此做了詳盡的回顧。一方面，在醫療大資料的基礎上，人工智能的海量計算、深度學習與自動化能力，能為醫療研究、病情診斷、醫療管理以及病人護理等方面提供前所未有的能性。但另一方面，智慧醫療也引起多方面的風險，令人擔憂。對人工智能弊端的討論主要集中在個人隱私、監控、大資料裡隱含的偏見和資料的可解釋性等方面。隨後這些爭論又延伸到對於人工智能的信任，醫生的去技術化，以及智慧醫療的脆弱性等方面，並由此引發了對智慧醫療中的權利，責任，價值的再思考。這篇文章對智慧醫療進行了全方

位的利弊分析，是一篇很全面的總結，有利於讀者了解這個領域的概況和爭論。

雖然對於人工智能利弊的爭論涉及眾多領域，各有側重，但是既有爭論中隱含了一個共同點：對於人工智能的討論大都是在“人機關係”的話語體系裡展開的。比如，“科技向善”的概念，無論對於“善”的理解是否不同，呼籲發展人與機器和諧共處，相輔相成的關係。而眾多對人工智能的批評與擔憂，則是更強調人與機器的對抗性。比如，機器可能搶了人的工作，導致大量失業；比如，人可能失去對演算法的控制，因此存在巨大安全隱患等等。在這個“人機關係”的話語體系裡，人和機器是對應主體。話語邏輯強調人和機器的相互塑造作用：人創造了技術，但是技術本身的邏輯能夠塑造人的行為。這個話語體系對於分析人工智能的利弊很實用，因此不僅被用在智慧醫療領域，也習慣性得被學者，政策制定者，實踐者，立法者用以討論人工智能在不同領域的應用。

只不過，這種“人機關係”的分析框架有可能忽略了目前對人工智能批判分析的癥結所在。對人工智能的應用進行利弊分析的確很重要。但是其重要程度，以及利弊程度，都很可能是被“人機關係”這個話語體系所塑造。誠如“人機關係”體系裡強調的技術邏輯能重塑人的行為，那麼這個話語體系的邏輯本身就預見了分析結論。因此，建立在這個話語體系上的分析就會產生一個話語悖論。一方面，批評者擔心智慧技術把病人簡化成資料，因此看病這件事就只剩技術演算法，而無人性。另一方面，這樣的批判本身就是一個技術視角：只見技術，而忽略了技術背後社會關係和社會規範的變化。因此，這個視角下得善惡好壞的分析通常還是在一個零和博弈：或善或惡，黑白分明。這樣一種簡單線性的分析框架忽略了一項新技術的發展與演變的可能性，也沒有討論導致這種演變的結構性條件。那麼，問題就產生了：如何用一個導致這個悖論的話語體系來分析解決這個悖論呢？很可能人

性的善惡有多複雜，技術的利弊就有多複雜。“人機關係”的話語體系表面上強調“人”，而簡單二元對立的邏輯卻恰恰忽略了“人”的複雜性和技術的多元性。因此，當下對於人工智能批判性分析中，缺失的恰恰是對用以分析的話語體系的批判性思考。

但是對“人機關係”話語體系的反思並不是否認技術（包括但不限於人工智能）本身的邏輯，更不否認技術對人的影響。只不過，“人機關係”的話語體系把新科技時代的問題簡化成“人與機器”的關係問題，從而忽略了新科技時代對“人與人”之間關係的重新塑造，因此帶來局限性，並也有可能誤導對於新技術的政策和法律制定。為什麼一個機器人殺人要比一個人殺死另一個人更讓人恐慌呢？為什麼機器人診斷錯誤要比一個醫生診斷錯誤要更為糟糕呢？為什麼病人要假設機器診斷的準確率會比醫生更高呢？在醫療事故中，誰要為智慧醫療系統的失誤負責任呢？在傳統專業主義的框架裡，醫生與病人之間的關係，無論信任與否，是建立在醫生具有病人不具備的專業知識的前提下。因此，醫生的權威是知識的權威。在智慧時代，機器提供的是資訊還是知識？它能加固還是弱化醫生的權威？很可能在智慧時代，醫生的專業性和權威性，以及可被信任性所需要的知識基礎會發生改變。因此個體醫生的去技術化不是個體醫生的問題，而是整個行業知識的更新換代。認識到這一點，相應的政策就不會只著眼於是否使用機器；而要更長遠地考慮如何讓促進醫生的發展，協調知識發展不均衡的不同醫生之間的關係。我們常問“智慧時代，人何以為人？”這個哲學思考其實不應只針對作為個體的人與機器有何不同，而也應關注到在不同技術環境下，人類之間關係該如何演變，以便定義“人”的群體性本質。

簡而言之，善良的機器也好，或邪惡的機器也罷，在機器的科幻浪漫色彩背後，歸根結底，是人與人的關係演變。生活在智慧時代，關鍵並不是如何處理人與機器的問題。而是當機器成為生活中必要工具的一部分的時候，人如何調整自己與他人的關

係。智慧時代，真正改變的不是人和機器的關係，而是人和人的關係。因此，我們的立法，政策以及倫理討論或許應該回歸到演變的人與人的關係的話語體系裡。

## 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.