



# Is Human Nature Computable?

## On the Relationship between Artificial Intelligence and Social Science Research

Hu Anning

**Abstract:** Along with the prevalence of big data and machine learning techniques, quantitative social science research gradually starts to use the new methods of artificial intelligence and machine learning to handle themes that were ignored or simply infeasible to study. Relative to the interpretation-oriented approach, quantitative research using algorithms in the fields of artificial intelligence and machine learning sets focus on prediction. Against this background, it is of academic urgency to examine the relationship between artificial intelligence methods and social science research. There are generally three pathways to address this question: First, based on the literature on quantitative research, we compare the traditional statistical modeling and the emerging algorithm-based modeling, so as to illustrate the paradigm contrast between interpretation and prediction. Second, from the perspective of computational science, we discuss over the potential limitations of artificial intelligence methods, pointing out its dilemma in handling non-patternized problems. Furthermore, we articulate that methods based on artificial intelligence lack the capacity of abstract thinking. Also, they are often applied separately, rendering the whole process to be short of systematic thinking. Third, we return to the classic dichotomy of sacred versus secular, and evaluate the predictive research from the perspective of social science theory. In particular, the principle of secular life is rationality, which results in the prevalence of instrumental rationality in the Weberian sense. For this type of rationality, we believe it is predictable based on computation. Compared with the secular sphere, the sacred sphere concerns supernatural, permanent, and perfect existences. Clearly, they cannot be gauged in a rational way, and in a sense, they can only be understood through personal experience. In this regard, we argue that artificial intelligence techniques cannot simulate sacred feelings. Only through the vivid human life can we understand and experience, reaching the status of inter-subjectivity, which is exactly one of the most important tasks for social science. Hence, artificial intelligence and prediction-based methods are applicable to the secular instead of sacred sphere. Simulation of human rationality and program design cannot direct us toward the supernatural sacred values. In summary, social scientist should be open to the new techniques, but at the same time keep cautious about the future application.

**Keywords:** artificial intelligence; computational science; social science; sacred versus secular

**Author:** Hu Anning earned his bachelor's and master's degree in law in 2005 and 2008 at Fudan University, and the master's degree in statistics and PhD in sociology at Purdue University. He has been the full professor of sociology at Fudan University since 2016. He is now serving on the editorial board of SSCI journals of *Social Science Research* and *Chinese Sociological Review*. Dr. Hu's research interests include culture and religion, education, social inequality, and quantitative methodology. his representative works include, *Sociology of Religion: Paradigm Shift and Chinese Experience*, and the *Theoretical Foundation of Causal Inference in Social Sciences*.

# 人心可以計算嗎？

## ——人工智能與社會科學研究之關係

胡安寧



**[摘要]** 隨着大數據和機器學習演算法的普及，社會科學的量化研究也逐漸試圖採用以人工智能和數據挖掘為基礎的新興研究手段，來處理過去無法處理或者忽視了的研究議題。與傳統以“詮釋”為導向的統計分析相比，基於人工智能與機器學習演算法的量化分析手段更加注重模型的預測能力。在此背景下，如何處理好人工智能技術與社會科學研究的關係就顯得尤為重要。解決這一帶有方法論性質問題的途徑有三條：其一，基於現有的量化研究方法論文獻，對“傳統”的基於統計模型的量化研究範式與正在興起的基於演算的量化研究範式進行對比，釐清“詮釋”與“預測”所代表的不同研究範式。其二，從計算科學的角度，對人工智能技術上的局限性進行討論。不僅要指出

人工智能的模式化分析方法在處理非模式化問題上的困境，而且要指出，基於人工智能演算法的分析技術缺乏抽象思維的能力，以及它各自為戰，缺乏整體性、系統性的思考能力。其三，回到社會學家塗爾干的“神聖—世俗”二分法論述，從社會學理論思辨的角度對人工智能在社會科學研究中的預測分析進行評價。具體而言，世俗領域的指導原則是理性化原則，因此，世俗生活需要個體具有馬克斯·韋伯所闡釋的工具理性，這可以通過演算法來預測。與之相比，神聖領域涉及超驗的、永恆的、完美的存在。這種存在無法用理性來衡量，甚至祇可意會不可言傳。因此，以演算為基礎的人工智能技術無法對超驗的神聖領域進行模擬。唯有通過人類活生生的感受，纔能夠體驗、理解，達到一種通感，而這恰恰是社會科學的重要任務之一。故而，基於人工智能的預測僅僅局限於世俗領域而非神聖領域，對人類理性計算的模擬和程序設計並不能夠對超驗的神聖價值提供指導。社會科學研究者在對新興技術手段持開放態度和胸襟去接納的同時，也要對人工智能在社會科學研究中的應用前景保持一種審慎的樂觀態度。

**[關鍵詞]** 人工智能 計算科學 社會科學 神聖—世俗

**[作者簡介]** 胡安寧，2005年、2008年在復旦大學分別獲得法學學士和法學碩士學位，2011年、2012年在普渡大學分別獲得統計學碩士和社會學博士學位，2016年起受聘為復旦大學社會發展與公共政策學院教授、博士生導師，兼任*Social Science Research*、*Chinese Sociological Review* 編委；主要從事文化與宗教社會學、教育社會學、社會不平等、社會科學量化方法研究，代表性著作有《宗教社會學：範式轉型與中國經驗》《社會科學因果推斷的理論基礎》等。

在過去的幾十年裏，建基於統計模型的量化研究方法在社會科學的不同學科得到了廣泛的應用。這些量化分析技術基本上是基於實驗數據或者調查數據，利用比較成熟的統計模型，對所關心的總體特徵（例如，某個參數）進行統計、估計和推論。學者們進一步對所估計出的係數進行詮釋，由此產生了大量具有啟發價值的經驗研究成果。<sup>①</sup>而隨着大數據和機器學習演算法的普及，社會科學的量化研究也逐漸開始試圖採用以人工智能和數據挖掘為基礎的新興研究手段，來處理過去無法處理或者忽視了的研究議題。<sup>②</sup>由於以演算和數據導向為基礎的人工智能技術可以讓社會科學研究變得具有預測性，因此被視為量化社會科學分析的一場劃時代革命。在此背景下，一個隨之而來的帶有根本性的方法論問題是，人工智能技術與社會科學研究究竟是什麼關係，如何處理？本文擬從多個角度對人工智能手段與社會科學研究之間的內在聯繫進行反思。

### 一 社會學量化統計模型的“解釋”屬性

關於傳統統計模型在社會科學中的應用，現有的相關文獻中已有論述。<sup>③</sup>比較著名的是分類樹算法的創始人布雷曼（L. Breiman, 1928—2005）的討論<sup>④</sup>。他指出，傳統的統計分析技術的核心在於理解和詮釋。這是因為，在採用統計模型進行經驗研究時，學者的關注點在於特定變量的解釋力，而並非整個模型的預測力。很明顯，這一分析範式也是社會科學量化研究的主流範式。例如，社會科學研究中最廣泛採用的方法，就是綫性回歸模型<sup>⑤</sup>。其中，“Y”代表了某一個因變量，而一系列的“X”則用來解釋“Y”。對於大多數社會科學研究者而言，一個比較普遍的區分是將“X”分為兩類：一類是人們關心的自變量，另外一類則是人們不關心但又不得不考慮的控制變量。以個體的工資水平分析為例，教育社會學家和教育經濟學家最關心的可能是教育水平的作用。按照傳統的“敏瑟模型”（Mincer model），除了考慮個人的教育成就之外，還需要引入工作經驗、性別、地區等控制因素；在中國，有時還要考慮工作部門、單位類型等等。此時，人們並不在意這些控制變量對於工資收入的影響，而是將關注點放在控制了這一系列因素之後，教育成就變量的純效應上。這體現在教育成就變量的回歸係數。研究者通常考察這一係數是否在統計上具有顯著性。如果顯著為正，人們便能夠得出結論，在控制了一系列社會學—人口學背景因素之後，教育能夠顯著地提升個體的收入水平。如此，一個具體社會科學研究議題的經驗分析便完成了。

與這一分析進路相比，基於演算的人工智能範式則不關心某一特定變量的解釋力。按照布雷曼的觀點，採用數據挖掘進路的研究者通常盡可能地納入各種與被解釋變量“Y”相關的因素，然後考察這些因素合併在一起，如何能夠在新的環境下盡可能準確地預測“Y”。正因為如此，基於演算法的研究往往將數據分為多個組進行交叉驗證，以求得最優的模型形式進行預測分析（比較常見的是將數據分為十組，利用九組數據建立模型，然後利用第十組數據看前面九組數據所訓練出來的模型是否能夠有比較強的預測能力）。這方面比較典型的應用或許是客戶分析。將大量的已有的使用者數據登錄，以此訓練出一個具有比較強預測力的模型工具。之後，當新的用戶出現時，這一使用者的某些信息便能夠輸入模型，並以此預測出使用者的某些屬性。例如，當人們在亞馬遜網站上買書時，一旦輸入社會學類書籍，網站會自動基於過往用戶的購買習慣，向

① 謝宇：“奧迪斯·鄧肯的學術成就：社會科學中用於定量推理的人口學方法”，《社會》3（2008）：81—105。

② 沈浩、黃曉蘭：“大資料助力社會科學研究：挑戰與創新”，《現代傳播（中國傳媒大學學報）》8（2013）：13—18；R. Axelrod, “Advancing the art of simulation in the social sciences”, *Simulating social phenomena* (Berlin: Springer-Verlag, 1997), 21-40.

③ K. Burnham, & D. Anderson, *Model selection and multimodel inference: A practical information-theoretic approach* (New York, NY: Springer Science & Business Media, 2003).

④ L. Breiman, “Statistical modeling: The two cultures”, *Statistical science* 3(2001): 199-231.

⑤ J. Stevens, *Applied multivariate statistics for the social sciences* (New York: Routledge, 2012).

這一新的用戶推薦一些其他書籍。誠然，如果這一新的用戶是社會學專業的大學生，而亞馬遜過往的演算是基於大量社會學專業大學生的購買習慣，這些新推薦的書籍很有可能正是新使用者所需要的，由此既給用戶以便利，又提升了網站的營業額。

通過對比上述兩個具體的例子，不難發現，至少到目前為止，社會科學中的量化研究方法主要還是布雷曼所談到的第一種範式，即以詮釋為導向的模型擬合：針對特定的“Y”，人們關心特定的“X”在控制一系列因素後如何影響其變化。通常而言，這也恰恰是社會科學理論產生的契機。比如，教育成就對於個體收入的積極效果被用來建立和完善人力資本理論、市場信號理論等等；如果一個國家的國內生產總值（GDP）提高以後，人們更加關注個人的情感和意見表達，則這一經驗模式可以用於支持英格爾哈特（Ronald Inglehart）的後物質主義理論；如果同樣的社會學—人口學背景下，女性的收入水平低於男性的收入水平，人們便有證據支持勞動力市場的性別歧視，等等。與這種“詮釋”範式相比，人們很少見到社會科學家會基於自己的研究模型做出具體的預測（例如，很少有政治社會學家預測何時會發生社會運動）。鑒於此，當越來越多的研究者將基於演算的、以預測為目的的人工智能技術引入社會科學研究的時候，人們似乎見到了社會科學新的發展方向，甚至是社會科學研究的範式革命。但是，真的會如此嗎？

## 二 “人心難測”：計算科學的視角

在計算科學中，人工智能是一個非常寬泛的領域，其中基於不同的應用方向有不同的分支學科。但究其根本，與社會科學密切相關的人工智能應用應當是機器學習。<sup>①</sup>所謂機器學習，是通過提供給機器（亦即電腦）特定的數據，讓電腦利用演算法，尋找出這些數據背後的模式。之後，當新的數據出現以後，人們便能夠用這些模式來套用數據，做出預測。雖然具體的機器學習演算法有很多（例如，決策樹，神經網絡，等等），但基本的分析邏輯大致如此。

那麼，如果用這套工具來試圖理解人類社會，做社會科學的研究，是否可能呢？答案是：介於肯定與否定之間。之所以有肯定的一面，是因為人類的某些社會性行為具有高度模式化的特點。例如，在給定特定資源的情況下，一個理性選擇的個人通常會做出利益最大化的決策（比如，在給定資金的情況下進行投資決策）。此時，機器可以對這種模式化的行為進行分析，甚至在立足於大量已有投資“經驗”的基礎上，“創造性”地通過演算給投資者指出一種常人可能忽略了的最優投資策略。從這個角度來看，人工智能的確可以幫助生物智能進行更為合理的判斷，其在社會科學中的應用應當得到肯定。

但是，除了這一類模式化的研究之外，如果簡單地認為人們可以利用人工智能技術完全預測個體行為，並建立預測性的社會科學，那就不免有些過於樂觀了。正是因為這一點，上述問題的答案中又包含否定的一面。具體而言，利用人工智能技術進行社會性行為的分析會面臨三個困境。

第一個困境，在於機器學習的模式化分析範式。正所謂成也蕭何、敗也蕭何，人工智能雖然能夠對於模式化的行為進行基於演算法的預測和分析（例如，“阿爾法圍棋”可以擊敗中國、韓國很多圍棋高手），但是，一旦人們面對的問題不是模式化問題時，那麼，人工智能就難以處理了。一個簡單的例子是：如果問機器“1+1等於幾”？機器會立刻回答“等於2”。但如果告訴機器，這裏的“1”代表一堆沙子，那麼，兩堆沙子放在一起是什麼呢？人們會說，還是一堆沙子。而機器卻無法做出類似的回答。因為，兩堆沙子混成一堆不再是一個數學化的模式問題，而是一個生活經驗問題了。

這個例子也將人們帶到了人工智能在處理社會問題時所面對的第二個困境，即缺乏抽象思維的能力。目前文本學習非常流行，其中有一些演算法能夠說明研究者分析一段文字所表達的作

<sup>①</sup> J. Friedman, T. Hastie, & R. Tibshirani, *The elements of statistical learning* (New York: Springer, 2001) .

者情感。<sup>①</sup>但是，這種分析到目前為止還比較機械，其背後的機制在於特定詞彙的計算特徵。因此，使用這種技術，人們所得到的信息也是非常初步的。換句話說，人們目前很難寄希望於人工智能能夠讀出字裏行間的潛意思。而對於社會科學而言，恰恰是一段文字的字裏行間纔能夠反映出作者的真實想法。如果僅僅根據詞頻或者其他表面上的文字使用模式，人們很難區分作者的真實意思和反諷之間的區別。試想，這樣的人工智能是否能夠理解魯迅（1881—1936）的著作，特別是他撰寫的雜文呢？

最後一個困境是，人工智能在處理具體問題時往往各自為戰，缺乏整體性、系統性的思考。2012年，有一部以棒球為題材的美國電影《點球成金》在全球公映，講的是一個棒球隊利用統計數據而不是個人經驗來選擇球員的故事。當然，這個故事的結局是基於數據的決策要優於基於經驗的決策。但即使如此，電影中通過統計數據所展示出來的信息依然要匯總到一個活生生的個體那裏，通過人腦的加工纔能夠形成最後的決策。而人腦在這裏恰恰起到了一個統籌的功能。在將人工智能應用於體育領域時，通過演算法和數據所提供的信息針對的是球場上特定位置的球員。例如，找到一個最好的後衛、一個最好的中鋒等等；但是，這個最好的後衛或中鋒能否與全隊有效配合，產生化學效應，人工智能卻難以判斷。如果出現“一山不容二虎”的局面，兩個最好的球員在一起所產生的效果可能反而不如兩個普通球員在一起合作的效果好。而這些超越個體的整體性判斷，則需要教練員與不同位置的球員進行“人與人”之間的溝通纔能做出。顯然，這些已經不是機器所能做到的了。對於社會科學而言，學者們感興趣的恰恰是人與人之間的聯繫，以及這種聯繫所產生出來的，超越個體的簡單加總所形成的具有獨立存在意義的“社會”。

### 三 “人心難測”：社會學的理论解釋

得出“人的社會性行為無法完全由演算來預測”這一結論，不僅已被計算科學研究所證實，而且也可以從社會學的經典理論中得到支持。從某種意義上說，能夠經由演算法來進行預測的人類行為往往是高度理性的行為。這種理性可以體現在多個方面。最為簡單的理性行為就是簡單重複勞動（例如，一個生產綫上的某一道工序）。對於這種重複勞動，一些簡單的演算法便能夠勝任，這在大量使用機器取代人工的工廠實踐中可見一斑。當然，理性的行為和邏輯並不僅僅是重複工作，還包括大量的基於形式邏輯的高級推導與計算行為。這裏比較具有代表性的，就是谷歌公司設計製作的“阿爾法圍棋”在圍棋上與各國頂尖棋手對弈時的“全勝”。

但是，人類的各種稟賦，尤其是那些構成人類社會本質特徵的元素並不完全是理性化的行為。甚至可以說，人類社會之所以成為社會，恰恰是因為每個人不是像機器般“算計”，而是有着非理性的情感、思想和價值。在這方面，法國社會學家塗爾干（E. Durkheim, 1858—1917，一譯“迪爾凱姆”“杜爾克姆”）的理論具有啓示價值。他在《宗教生活的基本形式》一書中，將人類社會區分為“神聖”與“世俗”兩大領域。<sup>②</sup>所謂的神聖領域，是指那些永恆的、完美的社會特徵；而所謂的世俗領域，則指那些與人們日常生活息息相關的“柴米油鹽醬醋茶”。在世俗領域，人們生活的指導原則是理性化原則。例如，如何養家糊口，如何完成某一個項目，等等。面對這些世俗活動，人們需要採用一種基本的取向，考慮如何採用特定手段來完成特定的任務。從這個角度來看，世俗生活需要人們有馬克斯·韋伯（M. Weber, 1864—1920）所闡釋的工具理性。<sup>③</sup>但是，人類社會之所以能夠存在，生活於其中的個體之所以能夠彼此團結起來而不至於一盤散沙，所依賴的並不是這種“精緻的利己主義”。相反，人們需要的是一種超驗的，在人們看

① R. Feldman, & J. Sanger, *The text mining handbook: advanced approaches in analyzing unstructured data* (Cambridge university press, 2007). 孫豔、周學廣、付偉：“基於主題情感混合模型的無監督文本情感分析”，《北京大學學報（自然科學版）》1（2013）：102—108。

② [法]愛彌爾·塗爾干：《宗教生活的基本形式》（上海：上海人民出版社，1999），渠東、汲能譯。

③ [德]馬克斯·韋伯：《社會學的基本概念》（上海：上海人民出版社，2000），胡景北譯。

來不那麼善變的、完美的存在。這種存在無法用理性來衡量，甚至祇可意會不可言傳。由於這種存在超越人們的日常生活，人們纔能夠在彼此之間找到共同的社會屬性，進而形成一個具有團結性的社會。<sup>①</sup>人們堅持這種價值，便成就了韋伯意義上的價值理性。<sup>②</sup>

這種世俗與神聖的區分，對於人們理解人類社會與社會科學研究的關係非常有價值。以演算法為基礎的人工智能技術或許對於預測世俗行為很有用，但卻無法對超驗的神聖領域進行模擬。對於後者，唯有通過人類活生生的感受，纔能夠體驗、理解，做到一種通感，而這恰恰是社會科學的重要任務之一。例如，婚姻的形成過程往往涉及很多理性化的考慮。尤其是在今天，夫妻雙方在最後走入婚姻殿堂之前可能在家庭義務、雙方家庭的物質支持等方面已經有一系列的討價還價。在這裏，雙方都在努力的做到利益的最大化，因此，理性化的原則通常會佔上風。然而，這些思慮在一般人看來都是“俗”的，因為這裏涉及的是斤斤計較和患得患失。但是，婚姻涉及的不僅僅是這些功利性的考慮，更重要的是雙方的愛情基礎。而愛情卻是一種神聖性的存在，人們通常認為它是永恆的、完美的，不容討價還價和妥協的。對於大多數的婚姻而言，世俗的元素與神聖的元素都是並存的。有些可能更加偏重對愛情的追求而不在意經濟利益，而有些則將經濟利益看的很重，不太重視愛情。但無論如何，一個基本的共識是，僅僅用理性計算來衡量婚姻是無法把握婚姻的本質的。通過這個例子，可以看到人工智能技術的應用點，但同時也看到了其局限性。很明顯，在涉及理性計算時，人們完全可以採用演算法來推理和預測個人的選擇和傾向，但這種預測僅僅局限於世俗領域而非神聖領域。換句話說，對人類理性計算的模擬和程序設計並不能夠對超驗的神聖價值提供預測。那麼，很自然的一個推論是，人們並不能夠利用人工智能來完全理解婚姻行為。

按照塗爾干的理論，人們所關注的社會恰恰是一種個人無法直接感受的超個體存在。而一切事物的神聖性，恰恰來源於其社會性。因此，祇要人們堅持認為社會具有超越個體的獨立性，即一種社會唯實論的主張，便不能夠認為人工智能演算法能夠解決“人心的預測”問題。對於社會科學家而言，更為適用的方法是直接的訪談和體驗，或者最多是傳統意義上的統計模型了。

總之，人工智能與社會科學研究的結合是近幾年學術界的熱點，但人工智能技術在社會科學研究中所出現的過度模式化、缺乏抽象能力和整體思考能力等問題卻表明，人工智能的演算邏輯更多適用於塗爾干意義上的世俗世界，而神聖世界是難以用演算和邏輯進行研究的。對於新興的技術手段，社會科學研究者應當具有開放的態度和胸襟去接納，但對於人工智能在社會科學研究中的應用前景，則需要保持一種審慎的樂觀態度。因為，如果變得技術至上，則有可能本末倒置；畢竟社會科學對於人類社會的主要貢獻在於提供理解人類社會的新思想和新視角，與之相比，工具的重要性理應是第二位的。

① [法] 艾米爾·塗爾干：《社會分工論》（北京：生活·讀書·新知三聯書店，2000），渠東譯，第46頁。

② [德] 馬克斯·韋伯：《社會學的基本概念》，第31頁。