

高奇琦——著

# 人工智能 馴服賽維坦

中和出版  
OPEN PAGE



## CONTENTS

# 目錄

推薦序一 「人工智能熱」的「冷思考」/ 夏建白 .....	1
推薦序二 人工智能的「第一性原理」/ 關新 .....	4

### 引言 參與性塑造：人工智能社會科學的未來

人工智能只是屬於科技界的「熱鬧」？ .....	8
怎麼看待人工智能可能帶來的失業影響？ .....	9
人類可以向人工智能學甚麼？ .....	11
參與性塑造與人工智能規則的制定 .....	13

## 第一部分 人工智能在各領域的影響

### 第一章 電子人權與算法獨裁：人工智能對法律的重塑

人工智能的發展對律師行業的衝擊 .....	19
電子人權的出現對民法的挑戰 .....	21
機械人和人工智能相關的侵權與刑事責任 .....	25
人工智能時代的法庭與司法活動 .....	28
算法獨裁的出現與法律的公平正義 .....	30
結語：用算法的正義來保護法律的正義 .....	33

<b>第二章 從金融科技到智能金融：人工智能對金融領域的影響</b>	
從寬容到人工智能 .....	37
人工智能在金融領域的應用 .....	40
人工智能對金融領域的顛覆性影響 .....	42
結語：智能金融的變與不變 .....	45
<b>第三章 共享智能汽車對汽車行業的影響</b>	
國際上智能駕駛領域的新進展 .....	49
中國智能駕駛領域的進展 .....	52
未來交通新模式：從無人駕駛到共享智能汽車 .....	54
共享智能汽車的技術、法律和管理問題 .....	57
結語：製造商轉型出行服務的大趨勢 .....	59
<b>第四章 人工智能時代醫療行業的機遇與挑戰</b>	
智能診斷：取代醫生的前景 .....	63
智能治療：手術中的「達文西」 .....	67
日常化護理：大數據與常態健康管理 .....	69
人性化醫療：緩解醫患矛盾的可能性 .....	73
智能醫療未來發展的三種積極趨勢 .....	76
結語：正確應對智能醫療的衝擊 .....	80

## 第二部分 人工智能時代的應對與選擇

### 第五章 從功能分工到趣緣合作：人工智能時代的職業重塑

作為一種基礎設施：人工智能未來發展的要義 .....	87
功能分工：傳統的職業規劃與教育 .....	88
結構性失業和全面性失業 .....	90
趣緣合作：人工智能時代的新模式 .....	93
興趣、心靈與終身教育 .....	96
結語：打破傳統職業分工的自由人與逍遙遊 .....	100

### 第六章 人工智能稅與差別原則：對被剝奪者的補償

機械人+人工智能：社會的生產部門 .....	103
憤怒的失業者：社會的不安定因素 .....	104
人工智能時代的差別原則 .....	107
徵收智能稅以補償失業者 .....	110
學習或靜修：失業者的兩種命運 .....	113
結語：從最不利者的角度來思考社會問題 .....	116

## 第七章 數據參與與雙向賦權：人工智能時代公民的 數據意識及其意義

人工智能時代中公民生活的數據化 .....	121
數據意識的意義與定義 .....	123
數據參與：公民參與政治的新方式 .....	127
開放政府與開放企業：數據參與基礎上的雙向賦權 .....	130
結語：用數據賦權來對抗權力鴻溝 .....	133

## 第三部分 人工智能與未來

## 第八章 主奴辯證法與相互承認：試論人工智能 戰勝人類的可能性

奇點：人工智能超過人類的臨界點 .....	139
自學習：人工智能讓人生畏的地方 .....	141
情感計算與自我意識的產生 .....	143
主奴辯證法的歷史循環 .....	145
人與機器相互承認 .....	147
結語：主奴辯證法的悖謬與新場景 .....	150

## 第九章 透明人與空心人：人工智能的發展對人性的改變

人性善惡的歷史爭論 .....	155
-----------------	-----

大數據時代下的透明人 .....	158
空心人在物聯網時代中的進化 .....	162
「套路」的終結與人類獨特性的消失 .....	165
結語：真善美作為人類的最終狀態 .....	167

## 第十章 作為新物種的數據生命體

大數據對時代和生活的改變 .....	171
數據工具論：人類對數據的偏狹態度 .....	173
為甚麼數據是有生命的？ .....	175
從相互承認到相互尊重 .....	178
人機結合：數據生命體的未來 .....	180
結語：數據的自主性與人的數據化 .....	182

## 第十一章 共產主義在人工智能時代實現的可能性考辨

網絡共產主義的端倪 .....	185
全球小康社會的可能圖景 .....	189
瓦爾登湖式的生活與相互合作 .....	192
可支配時間的增加與心靈的安頓 .....	194
共產主義與自由人的聯合體 .....	195
結語：自由人的全面發展 .....	199

結語 馴服賽維坦：把握人工智能的發展方向	
從利維坦到賽維坦 .....	202
跨眾合善：未來人工智能的發展方向 .....	203
「智能+」：一種新的認識論和方法論 .....	206
結語：人工智能潛能的運用與馴服 .....	207
主要參考文獻 .....	208
後記 .....	212

## 推薦序一

# 「人工智能熱」的「冷思考」

夏建白（中國科學院院士，中國科學院半導體研究所）

2017年7月，國務院印發的《新一代人工智能發展規劃》明確指出，人工智能的迅速發展將深刻改變人類的社會生活、改變世界。那麼，如何從社會發展的角度認識人工智能的重要影響呢？高奇琦教授的《人工智能：馴服賽維坦》正是從社會科學的角度來討論人工智能未來發展的一本重要著作。作為一項顛覆性的技術，人工智能不僅會對未來的生產方式、就業結構和社會倫理產生極大的衝擊，也會對政府管理、經濟安全、社會穩定乃至全球治理等領域產生深刻影響。正如作者所指出的，多學科的參與是保障未來社會人工智能更好發展的前提。中國要在未來成為人工智能強國，不僅需要在基礎理論、核心算法以及關鍵設備上形成重大突破，還要形成有中國特色與中國風格的人工智能發展政策法規與社會科學理論體系。在邁向這個目標的發展進程中，高奇琦教授的這部著作可能會發揮重要的開拓性影響。

這本書的特點是用較為通俗的語言來介紹人工智能及其可能產生的社會影響。可以想見，本書將會受到廣大讀者的歡迎。人工智能是近幾年出現的

新生事物，人們在關心它的同時又對它持有各種各樣的看法。這本書就給大家提供了一個討論人工智能及其前途的平台。在對這本書給予肯定之外，我還想把我對人工智能的思考，結合社會上的一些觀點，在此提出來供大家討論。希望讀者在閱讀本書的時候，一邊考慮我所提出的問題，一邊與周圍的人討論。當然，正所謂「實踐是檢驗真理的標準」，這些問題的結論到底如何，還要看人工智能以後實踐的結果。前幾年人們還不相信機器與人類下圍棋能贏，在 AlphaGo 與世界頂級圍棋大師對弈並獲勝後，人們就徹底服氣了。我所強調的主要觀點如下：

第一，社會上的一些觀點可能過於強調了人工智能對就業結構帶來的衝擊。在我看來，人們還需要謹慎思考人工智能的優勢和不足。以小說為例，機器翻譯一些通俗小說也許行得通，但翻譯經典名著就令人難以接受。每一部小說都有其歷史背景和環境，每一位作家都有其風格和特點，這些東西只有通過優秀的翻譯家的翻譯才能領會。人們讀經典名著並不是只知道一個小說情節就滿足了。至於科學技術方面的書，我自己作為科技工作者，寧可閱讀外文原著也不看那些蹩腳的翻譯，遑論閱讀機器翻譯的科技書籍。

第二，社會上的一些觀點強調人工智能在詩歌創作以及藝術創造方面的能力。本書作者也提到了「百度寫詩」的案例，並指出人們很難分辨機器的作品和人類的作品。我個人認為人工智能在藝術領域的作為還非常有限。我估計書裡提到的「人們」都是年輕人，學習古詩的時間不長，鑑賞古詩的能力和水平相對有限。我相信真正懂古詩、有鑑賞能力的人，一看就知道詩的作者是誰、是在甚麼歷史背景下寫的、歌頌甚麼、鞭撻甚麼，因此也一定能夠分辨出機器寫的詩和唐朝詩人寫的詩之間的差異。

第三，機器學習依然無法替代人的學習。人工智能的發展，如搜索引擎和維基百科、論文自由下載等確實給廣大科技工作者帶來很大方便，省去了許多查找資料的時間。但是學習與科學研究，以及文學、藝術如寫小說、翻譯等都是個人行為，每個人有各自的基礎、興趣和愛好，機器並不能完全代

替這些東西。學習只能一個人專心鑽研，如錢鍾書那樣，才能真正學進去。依靠集體和機器幫助，只能學到一些皮毛而無法掌握其中的精髓。

第四，人們不應過於悲觀地看待人工智能對人類就業產生的衝擊。當然，現在人工智能剛開始發展，它造成的後果誰也不能預料。我認為，勞動分腦力勞動和體力勞動。有些腦力勞動能夠被機器代替，但關鍵的地方還需要人類自己的智慧。至於體力勞動，除了生產線上的機械人，其他如修建大橋、高樓、高鐵等是機器無法代替的。美國為甚麼自己修不了高鐵，不是技術不行，而是沒人幹這種辛苦的工作。

人工智能不是萬能的，並非在任何場合都適用。社會永遠存在差別，如貧富差別、知識程度差別。在日常社會生活中，一些社會服務，包括銀行、電訊，甚至餐飲業，都應該盡量照顧到大多數人的認知能力和知識水平。現在社會到處在積極推廣二維碼、支付寶、電子帳單等，這對城市裡的年輕人來說也許沒有問題，但對農村來的農民，以及包括我在內的城市老年人而言就有些無所適從。所以在大眾服務領域應該盡量採用通俗易懂，便於大家學習、掌握的方法，而不能只開發那些陽春白雪、只有少數人能使用的高新技術。

## 推薦序二

# 人工智能的「第一性原理」

關新（寬資本創始人、董事長，奇點大學投資人）

人工智能時代可能是人類發展史上最大的「灰犀牛」事件。它的到來速度及衝擊會帶來哪些翻天覆地的變化？我們人類又該如何應對它的到來？《人工智能：馴服賽維坦》一書提供了全面的、發人深省的研究與預測。無論你視人工智能為洪水猛獸，或是認為它將會把人類社會領進共產主義，還是你對這個領域一無所知，都可以從這本書中獲取答案。高奇琦教授以他開闊的國際視野、嚴謹的科學論點以及深厚的智慧良知，為我們步入人工智能時代提出了極有深度的命題與思考，一定會對讀者在人工智能時代的自我定位產生積極的影響。

人工智能時代可能帶來的最大衝擊是在地球上突然出現了一個超越人類的新物種：機械人。這些機械人是人類的朋友還是天敵目前定論尚早。可以定論的是，由於人類的思維是線性的，而機械人的發展是指數性的，我們可能低估了這個新物種將要帶來的大變革。我們對人工智能及機械人的認知在很大程度上還停留在互聯網時代，認為人工智能就是大數據、機器學習以及

先進的算法。甚至有的人認為誰擁有數據，誰就將擁有人工智能的未來。我們必須看到這種認知是不全面的，同時更要指出的是，目前更龐大的人工智能系統正在尋找多項學科與行業的「第一性原理」。這可能會徹底顛覆我們在某些行業上的認知，甚至整個行業。因此，高教授提出人類應該向人工智能學習並且從各方面積極參與，因為誰擁有了未來行業的「第一性原理」，誰就具備了改造這個行業的最大創造力。

18世紀航海定位的「第一性原理」是時間與角度，如今我們用它來導航；19世紀飛行的「第一性原理」是速度而不是翅膀，現在人類比鳥兒飛得更高更快；遺傳學的「第一性原理」是DNA（脫氧核糖核酸），我們用它改變基因而延長生命。機器學習的「第一性原理」是概率與算法，我們用它打敗天下圍棋高手。NASA登陸火星的機械人「第一性原理」是人工智能模仿人類的判斷力，從而可以在沒有地面指示及缺乏數據的情況下進行大量的科學實驗。這些都是人類在掌握了事物的「第一性原理」後創造出奇跡的例子。我們可以期待人工智能時代的自動駕駛與飛行器、自動設計與拼成生產線、生物與機器結合的智慧腦、基因改變與人體器官打印、數字生命與財富及人類星際旅遊等行業在今後幾十年成為新興的行業。

人工智能時代是人的時代還是機器的時代？未來的自然科學與人工智能科學將向何方延伸？我們應該發明哪種智慧與哲學才能與機械人共贏？高奇琦教授在書中提出了「善智」的理論，也是本書的精華之一。用「善智」的哲學思想迎接人工智能的挑戰，是現代版的「天人合一」。中華民族要發揚傳統的智慧，並在敢於自我革命的勇氣下去跨界學習，不斷探索和創新人類與自然科學、智能科學、人機科學的「第一性原理」。我希望人工智能政策的制定者、時代「弄潮兒」以及被人工智能波及的人都能在此書的啟發下，積極利用時代賦予我們的智慧與手段去聯合創造一個嶄新的世界，一個以造福人類後代為目的的世界。這就是我所理解的「善智」。

引言

# 參與性塑造： 人工智能社會科學的未來

為天地立心，為生民立命，為往聖繼絕學，為萬世開太平。

——張載（北宋理學家）

未來已經發生，但是尚未流行。

——威廉·吉布森（美國科幻作家）



人工智能時代，  
社會科學不能只看熱鬧

人工智能時代的到來，帶來了政治、經濟和社會等各個領域的變遷，包括學術界、商界和政府內的所有領域都將步入這種智能化的進程。同時，人工智能時代的到來也產生了新的議題和研究領域。那麼在這場顛覆性的革命中，人文社會科學領域的學者扮演著甚麼樣的角色呢？在我看來，在人工智能時代，人文社會科學領域的學者並不是一個被動的參與者，而是要在參與性塑造的基礎上，構建人工智能社會科學的未來。

## 人工智能只是屬於科技界的「熱鬧」？

眾所周知，人工智能的時代已經來臨。不僅一些大的軟件公司正在佈局人工智能，傳統的電腦製造商也紛紛向這一領域進軍。例如，IBM 公司通過收購以色列 Mobileye 公司，全力轉向智能汽車領域，Facebook 也在全力佈局人工智能的未來規劃。值得注意的是，目前全球市值排名前列的公司都在全力進軍這一領域，中國的互聯網巨頭也不例外。譬如，2017 年 2 月，國家發改委正式批覆由百度牽頭籌建深度學習技術及應用國家工程實驗室。在此之前的幾年時間裡，百度在人工智能領域已投資超過 200 億元人民幣，招募了上千名軟件工程師和開發人員，在人工智能汽車方面也取得了驕人的成績。2017 年 3 月 19 日，在日本東京舉行的世界電腦圍棋大賽決賽上，中國另一家科技巨頭——騰訊公司開發的人工智

能圍棋程序「絕藝」戰勝日本的「DeepZenGo」獲得冠軍。

簡言之，整個科技界都呈現出人工智能熱的圖景，而這種熱度也反映在幾次大的預言中。《未來簡史》一書指出，人工智能時代的來臨將會對未來的職業產生重要影響。在未來，有一些職業將會面臨衝擊，這些職業依次是：旅遊諮詢業、外文翻譯、律師助理、銀行保險。除此之外，還有預言描繪了未來的社會變化。例如，人類在十年之後將不再擁有汽車，自動駕駛的出租車會成為主流的景象。這將為城市騰出大量的空間，而這些空間將會被改造成公園，從而提升城市生活的舒適度。

與科技界的「熱鬧」不同，人文社會科學領域的學者似乎表現得很淡然。其實，這對人文社會科學而言，在某種意義上是一種窘境。目前的學術界似乎存在這樣一種邏輯：只有某個學科內的學者才有權利對該學科的問題發表看法。在一些人文社會科學學者看來，人工智能是自然科學研究的前沿領域，而人文社科在這一領域缺乏建樹，所以不能在這一領域評頭論足。這種觀點阻止了人文社會科學學者參與人工智能的討論。此外，人文社科學者相對「沉寂」的一個原因是，科技領域的發展進步日新月異，已經遠遠超出了傳統意義的思想家和哲學家思考的範圍。

在這場可能決定人類未來發展的大討論中，人文社科學者的缺席，顯然不利於人工智能的總體發展。例如有學者就指出，大數據的未來不能僅僅掌握在那些數據公司手中。人工智能和人工智能的未來同樣也不能僅僅掌握在自然科學家、工程師和科技公司手中。因為這不僅是科技和科學家的未來，也是整个人類的未來。科學家和工程師可以掌控科技發展的技術走向，但不一定能深刻地理解科技對人類的複雜性影響。因此，人文社科學者和思想家們不能缺席這場討論，而應該主動參與到影響人類命運和人類發展的大討論中。

## 怎麼看待人工智能可能帶來的失業影響？

人工智能的發展對社會的影響將是顛覆性的。1908年，奧地利經濟學家約瑟·熊彼特（Joseph Schumpeter）提出「破壞式創新」這一顛覆性的概念。他認為，所有社會都會經歷這種過程，每一次大規模的創新都淘汰舊的技術和生產

體系，並建立起新的生產體系。李察·科士打（Richard Foster）和莎拉·卡普蘭（Sarah Kaplan）在熊彼特的「破壞式創新」的啟發下提出「創造性破壞」的概念。他們認為，無論是從企業到國家，還是從個人到社會，都具有面臨重大改變的必然性，即「不連續性」或「斷層」。從這個意義上講，人工智能的發展對人類的進步的影響也是顛覆式的。

人工智能對社會的影響，將造成兩種狀態的失業：一種是結構性失業，另一種則是全面性失業。結構性失業指的是，在人工智能的衝擊下，某些行業將在短期內面臨結構性的挑戰，甚至存在被歷史發展替代的可能。這種衝擊的對象首先是專業化、程序化程度較高的行業或職業，如傳統的翻譯行業。目前有些人工智能技術的翻譯水準正在接近，甚至有可能超過人類的翻譯水平。未來，人們到一個陌生國家旅行的時候，只需在手機上下載一款翻譯軟件，就可以與當地民眾進行無障礙交流，到那時傳譯行業將受到很大的影響。而相比於線下的旅遊中介，人工智能提供的服務更加廉價和便捷，人們將更願意依靠人工智能的網站提供的服務，由此將造成大批旅遊中介失業。此外，律師助理也將是受到人工智能挑戰的領域之一。在未來，人們通過人工智能平台獲得的法律服務將比律師提供的更為精準，而這已經成為律師行業發展的一個重要趨勢。以上這些領域受到的衝擊，可以稱之為結構性失業。因為這些領域所提供的服務，在未來能夠輕易地被人工智能所取代，而這些領域的從業人員則會面臨需重新擇業的風險。

另一種失業則是全面性失業。這裡主要指的是受到人工智能衝擊的覆蓋面。儘管不同的職業受到人工智能的影響不同，應對衝擊的調整方式也不一樣，但這種影響將是全覆蓋式的，各行業幾乎都不可避免。甚至連那些人類自認為最擅長的領域，如文學和藝術領域，也成為人工智能入侵的對象。例如，多倫多大學開發的人工智能機械人可以基於藝術家的曲譜，通過演算創作出美妙絕倫的樂曲。而百度通過「圖靈測試」的方法，將人工智能創作的作品與唐朝詩人寫的詩放在一起供人分辨，結果顯示人們很難分辨出兩首作品各自的創作者。甚至在結果揭曉後，很多人認為人工智能寫的詩比唐朝詩人的作品還要好。

以上對失業的分析意在提醒人文社會科學的學者參與到這場關乎人類社會未來發展的大討論中來。因為，科學家們往往只考慮用科技的進步來解決他們眼中一些具體的問題，以及他們的解決方案能否解決這些問題，對於科技之外的問

題，如對社會產生的影響，則較少去考量，而這些問題往往更加重要。例如，針對人工智能的發展對就業產生的顛覆性影響，科學家們並沒有給出解決方案。而這反過來也會對科技進步產生影響，因為科技需要在社會穩定和經濟繁榮的條件下才能有大的突破。由此，社會科學學者的研究意義就充分凸顯出來，研究影響人類未來發展和命運的問題本來也就是他們職責所在。這就好比科學家打開了一個個潘朵拉魔盒，但並沒有意識到打開魔盒後所增加的東西打破了社會內部所具有的微妙平衡，此時正需要社會科學家貢獻智慧使社會達到再平衡。

## 人類可以向人工智能學甚麼？

面對人工智能這樣一股洪流，有些人文社會科學的學者採取一種拒斥的態度，甚至認為技術是淺薄的，從而輕視自然科學的價值。這種對技術的輕視態度阻礙了他們對自然科學領域的學習。事實上，近些年自然科學取得的許多進展，對社會科學的發展也產生了革命性的影響，卻沒有引起人們的廣泛注意。例如，根據自然科學的研究，人類來自智人，而智人起源於非洲的母親。按照這一主張，人類文明的母親只有一個，即同一個「非洲媽媽」。這一觀點對非常流行的「文明衝突」的論斷將產生顛覆性的影響。然而，這一關於人類起源的自然科學界的進展卻並沒有得到社會科學界的關注。學科之間的壁壘成為一種傳統知識自我強化的工具，也導致了學科內學者對學科外學者的輕視。某些學科的學者驕傲地認為，只有他所從事的學科才最為重要，其學科內的知識是最為正確的。從學科本身發展角度而言，專業化固然是一種進步，但過於封閉以及學科壁壘的自我強化，阻止了人類外向學習的可能，繼而也產生了學科內的傲慢與偏見。這種傲慢與偏見在人工智能時代變得越來越不合時宜。

近年來，人工智能領域取得的巨大進展來自另外一個學科——神經認知學。簡言之，人工智能本身就是一個知識跨界的產物。如果僅僅局限於某一個學科，革命性的創新就不可能到來。人工智能是通過借鑒神經認知學的發展，從而模擬人類智能的一個成果。此外，近年來推動人工智能發展的大公司也在演繹着這種知識跨界。例如，美國 IBM 公司以前是從事傳統芯片和電腦生產的企業，但在 21 世紀之初卻將自己的主業務賣掉而轉向人工智能研發。從芯片到電腦製造再到人

工智能，本身就是一種知識大跨界。跨界是那些面向未來、敢於想像的企業的重要特質，傳統的人和企業則往往不敢想像。而只有跨界才有新的未來。

此外，需要特別關注的是人的深度學習的問題。作為一個人工智能的專業術語，深度學習是機器學習的一個重要組成部分。在這樣一個機器都要進行深度學習的時代，具有自主意識的人更應該進行深度學習。所謂人的深度學習，是指人需要跳出傳統的知識界限，廣泛地汲取各學科和領域的知識，不斷產生新知識技能的過程。如果沒有這種深度學習的能力，人終將被人工智能打敗。人的深度學習主要包括兩部分：一是跨界學習，二是團隊學習。

跨界學習是指跳出自己的學科邊界汲取與自己學科相關的前沿進展，進而反哺自己學科的過程。信息大爆炸時代為人們跨界學習提供了工具和條件，各種開放知識產權的媒體和社交工具便利了人們的跨界學習。例如，網絡上有大量的專業性的學習視頻、免費軟件和開源軟件。只要學習者有足夠的決心和毅力以及相應的基礎，藉助這些資源就可以學習掌握一些技能。

而社交媒體的革命能夠以學習小組的形式把相同興趣的人凝聚在一起。學習者可以加入一個 VR（虛擬現實）的社群，學習某種技術，並在此過程中獲得認同感。

由網絡社群、電子書、網絡視頻等多維空間組合成的虛擬社群，為學習者提供了一個長期的跨界學習沉浸式的情景。如果學習者迷戀上了某個領域的知識，這一系列網絡條件會幫助他沉浸在這種學習氛圍中，從而在更短的時間內更高效地學習。

在跨界學習的基礎上，團隊學習可以進一步加速跨界學習的效率。人工智能的能力，實質上是通過分佈式學習的方式提高。機器學習往往先根據學習內容分成不同的小組，每個小組再學習各自的算法和可能性，最後機器將不同小組的內容匯集在一起，從而產生一種共同智慧的結果。這種方式同樣可以幫助人類進行跨界學習。如果一個人學習一本書的內容比較複雜和艱苦，那麼十個學習者就可以把這本書分成十部分，最後十個人再以小組討論的方式進行交叉學習，相互印證直至最後打通所有的學習要點。需要強調的是，學習小組的規模和學習效率之間存在一個均衡點。學習小組的規模越大，成員間交流的密度就會減弱，同時聯繫溝通的成本會增加；而學習人員過少，這種學習的規模效應就無法得到體現。因此，在一個面對面學習的小組中，五至十人是最佳的規模。

社會科學的學者與其被動地面對人工智能的挑戰，不如主動參與到人工智能的大浪潮中。同時，也只有社會科學的學者對人工智能發展的充分參與，才能保證自然科學的新進展仍然在人類能夠掌控的節奏中發展，從而解決我們未來可能面對的問題。

## 參與性塑造與人工智能規則的制定

當前人工智能所面臨的許多關鍵性問題，不僅是技術問題，更多也是社會科學問題。譬如，之前討論的由於科技進步而導致的失業問題就具有代表性。這在人類歷史上也是一個古老的問題。馬克思（Karl Marx）就曾用「機器吃人」這一概念來對此總結和分析。只不過在馬克思那裡，機器「吃掉」的是工人，而在今天，所有產業都將面臨這種「被吃」的風險。因此，社會科學學者應該研究人工智能對未來職業產生的系統性影響，包括哪些人和哪些職業不會面臨失業的問題，以及人工智能會創造出哪些新的職業需求等。而且，在人工智能時代，創新精神顯得尤為重要。因為一旦工作被模式化和程序化，人工智能就可以很快學會。與機器相比，人類的優勢主要集中在靈活、情感和創造力等方面。這些都是與創新相關的領域。

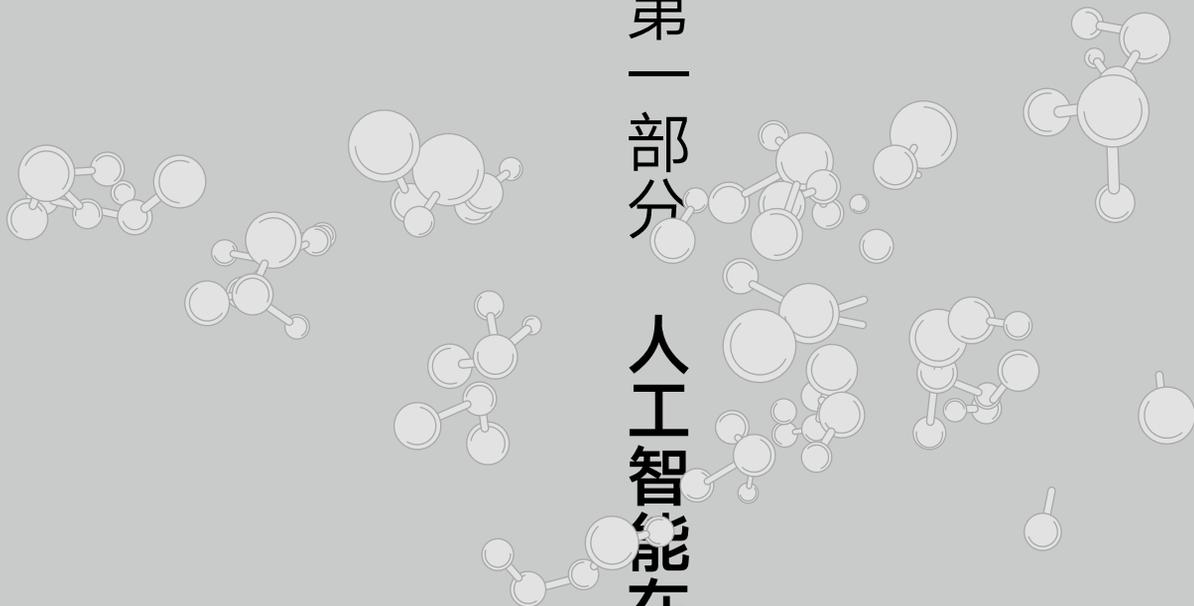
人工智能除了對人類就業產生衝擊之外，在許多方面也將對現行的法律和規則提出挑戰。例如，目前智能汽車上路運行面臨最大的瓶頸是相關法律、法規和公共管理規則的制定問題。傳統汽車的使用模式是汽車被賣出之後所有權歸車主所有，由汽車造成的一系列問題也多數由車主來承擔責任，如違規停車和交通事故等，除非汽車有明顯的質量問題，相關責任才由汽車製造商來承擔。然而在智能汽車領域，由於汽車是由智能汽車提供商所生產的系統來控制而不是由車主駕駛，如果出現交通事故，事故責任是由車主還是由智能汽車提供商承擔便成為有爭議的問題。

自動駕駛技術根據自動化程度被分為六個級別。目前，智能汽車還不能完全做到自動駕駛。在這六個級別之間，例如第二級和第三級，自動駕駛是作為車主的輔助性工具出現的。換言之，在緊急狀態下，車主應該也有必要採取如終止自動駕駛等應急的措施來避險。由此就會產生以下問題：第一，車主的責任邊界

在哪裡？第二，如果車主避險失敗，責任由誰來承擔？還有一個值得注意的問題是，自動駕駛通常是通過無線網絡進行的，而網絡運營商並不能保證信號時刻暢通。假如出現的故障和網絡有關，網絡運營商也會牽扯其中。因此，在消費者、網絡提供商、汽車製造商和智能駕駛系統提供商之間，就會出現多個交叉的責任空白區域。因此，智能駕駛在未來需要解決的不僅是技術問題，還會面臨法學、社會學、政治學等相關領域的問題。這些問題都需要社會科學領域的學者來推動解決，而這也就要求社會科學家對人工智能的前沿知識有所學習和探究。

# 第一 部分

# 人工 智能在各 領域的影 響



## 第一章

# 電子人權與算法獨裁： 人工智能對法律的重塑

法官的工作被限定為單純的解釋法律和契約，就像是法律的自動售貨機，人們在機器上投入事實（加上費用），機器下面就會吐出判決（及其理由）。

——馬克斯·韋伯（德國社會學家）

我個人的信仰飛躍是這樣的：當機器說出它們的感受和感知經驗，而我們相信它們所說的是真的時，它們就真正成為有意識的人了。

——雷·庫茲韋爾（美國發明家，奇點大學校長）

人工智能的發展將對法律及其行業生態產生巨大影響。人工智能的查詢服務以及訴訟服務已經對初級律師造成了壓力。這也將促使未來的法學教育根據新的發展形勢做出相應的調整。關於人工智能電子人權的討論則會對民法的傳統內涵形成巨大衝擊。人工智能和機械人使得侵權與刑事責任的鑒定變得更加複雜。在人工智能時代，法官的審判活動也會受到相應的影響。更為重要的是，人工智能背後的算法黑洞隱含着一個算法獨裁的問題，這種算法獨裁來自人類對算法的依賴。如果不能對這種算法獨裁建立良好的制約，未來法律的公平正義將深受影響。

人工智能的發展對法律形成了巨大衝擊，不僅體現在其對律師行業未來就業所產生的壓力上，還體現在其對民法、刑法、侵權法、知識產權法等法律條文和判例體系的深遠影響上。由於法律具有滯後性，現有的法律體系並不能很好地解釋和應對人工智能以及機器擬人化後可能出現的法律現象。可以說，在人工智能的衝擊下，法律體系可能需要全面重塑。這一問題所涉非常宏大，本章只從頭腦實驗的角度出發，試圖對人工智能的發展與未來的可能性做關聯性分析。

## 人工智能的發展對律師行業的衝擊

歷史證明，技術對社會的影響常常頗為弔詭。一方面技術可以改善人們的生活，延長人類的壽命，讓一些從事新行業、掌握新技能的人發揮更大的作用；另一方面技術也會讓更多的人離開已有的工作崗位，失去生活經濟來源。作為目前最可能引起人類未來巨變的領域，人工智能也必然會產生類似的社會影響。回顧過去，人類歷史上的三次技術革命有一個共同的特點，即技術在給人類社會帶來便利的同時，其帶來的巨大衝擊需要近半個世紀的時間才能被消化。如今，當智能革命來到之時，律師就成為受衝擊巨大的行業之一。我們不能單單感歎歷史的再次重複，而要對這種影響進行深入的思考，從而加以有效的應對。

在西方，律師往往被認為屬於高端職業，具有較強的專業性，處理的案件和問題也較為複雜。律師參與的訴訟過程會直接影響法庭的判罰結果，這導致律師在法律案件中的作用顯得尤為重要。而在訴訟的過程中，訴訟雙方都需要支付律師高昂的訴訟費，以換取律師的盡心盡力。對於大多數公司而言，律師費用往往是一個較為沉重的負擔。以生物技術領域為例，美國每年專利侵權訴訟至少產生2500萬美元的訴訟費，其中需要支付給律師的費用就高達700萬美元。

這種情況在英美法系國家表現得尤其明顯，因為英美法是強調判例法的典型法系。在一場較大的訴訟中，律師及其團隊往往需要將歷史上大量的相關法律文件匯總在一起進行分析，由此找到最優的解決方案。在過去，對於案件律師來說，這個工作量極其巨大，一個較大的訴訟案件可能會涉及上百萬份不同的歷史文檔和法律文件。

但是伴隨人工智能的出現，法律訴訟需要的時間與費用正在不斷降低。過去

需要花費大量人力、物力才能完成的工作，如今人工智能可以在極短的時間內完成。例如，矽谷的黑石數位探索科技（BlackStone Discovery）發明了一種能夠處理法律文件的自然語言處理軟件。它使得律師的工作效率可以提高 500 倍，從而將打官司的成本降低 99%。而世界上第一個機械人律師 Do Not Pay（「不花錢」）也在倫敦、紐約和西雅圖等地被人普遍使用，它已經幫助 16 萬餘人處理了交通罰單，且成本極低。同時，這一服務也在舊金山、洛杉磯等城市迅速擴展。在你填寫一份問卷調查表後，一旦法律機械人認定你合法，它就會幫助你開具一份抗辯授權書。

儘管目前這些進展還局限在一定範圍之內，但是已經顯示出了影響未來的重要趨勢，即人工智能的發展意味着未來有相當多的律師，特別是初級律師會失去工作。2014 年，北卡羅來納大學法學院教授德納·雷穆斯（Dana Remus）和弗蘭克·萊維（Frank Levy）通過研究自動化在大型律師事務所中的應用發現，如果全部採用新的法律技術，在其後五年內從事法律行業的人數將減少 13%。而抵消人工智能自動化對法律行業影響的辦法只能通過每年法律事務在原有的基礎上增加 2.5% 來實現。

2015 年，美國統計公司奧特曼韋爾公司（Altman Weil）就人工智能能否取代律師，對 20 所美國律所的合伙人進行了一項民意調查。結果顯示人們對人工智能的認可度越來越高，多數人相信人工智能能夠取代人類律師。其中 47% 的受訪者認為，在 5~10 年內，律師助理將失去工作。只有 20% 的受訪者認為人工智能無法替代人類。也有 38% 的受訪者認為，在 5~10 年內人工智能還不能取代人類律師。麥肯錫全球研究院（McKinsey Global Institute）在 2017 年的報告也指出，複雜的機械人正從工廠走出，成為法律、新聞以及醫學領域的一部分。

從目前來看，人工智能對於律師重覆性和初級性的工作顯然具有替代作用。律師在辦案中花費時間最多的是翻閱那些海量的文件，從中尋找案件的漏洞以及一些相關的證據。而人工智能的搜尋和算法可以在非常短的時間內完成這些工作。同時，人工智能的發展還會使科學輔助謊言檢測工具得到不斷的升級。例如，隨着測謊機在現實應用中的缺陷不斷凸顯，測謊機將會逐漸消失，而神經影像技術將成為法庭中可驗證謊言的救世主。這種變化將不可避免地從根本上改變法律業務的工作方式，使其對技術的依賴上升到新的層次。對於這些由人工智

能引發的變化和挑戰，美國的一些大型律師事務所已經採取積極的措施來進行應對。

擁有 7000 多名律師的大成律師事務所（Dentons）在 2015 年成立了推動技術創新的風投部門——下一代法律實驗室（Next Law Labs）。該部門負責對新的技術進展進行充分的監測，並對 7 家法律技術初創公司進行了投資。大成首席創新官約翰·費爾南德斯（John Fernandez）表示，「我們的產業正遭到破壞，我們應當有所作為，不應坐以待斃」。而基拉系統公司（Kira Systems）也表示，在兩年半的時間裡，他們已經可以讓軟件識別類似競業禁止合同條款和變更控制權等概念，並飛速地挑選相關文檔，但審閱工作尚需人工完成。我們可以發現，雖然一些律所的高級合伙人似乎已經認識到律師助理或初級律師將被人工智能取代這一趨勢，但他們對此卻並不擔心，因為其運營成本將會大為降低。

律師業所受到的這種影響，最終也會擴展到法律的其他領域。尤其值得關注的是，人工智能的發展還會重塑未來的法學教育。在未來，法學教育要更加強調對電腦、雲端運算與大數據的訓練。未來的初級法律從業員，將具備查詢法律文書、運用人工智能相關軟件的能力。高端的研究者與從業員要具備運用 SQL、Python 等與人工智能相關的電腦語言進行電子案件查詢、分析以及可視化的能力。傳統的法學訓練在分析某一案例時往往傾向於選擇具有代表性的個別案例，而未來則需要對相關議題進行全樣本的比對，在大數據的分析下得出更為科學的結論。

因此，我們的法學教育需要針對這些變化做出相應的準備。這就需要相關法學院校增加精通電腦、人工智能與大數據的從業員，吸引電腦領域的研究者進入法學研究，引導相關研究者充分運用各級法院的法律判決案例與判決書進行大數據分析。這樣一來，法學教育不僅要加強對已有的研究和教學隊伍的訓練，還要對本科生、碩士生以及博士生的教育內容進行相關調整。

## 電子人權的出現對民法的挑戰

美國在 20 世紀 80 年代中期就成立了研究未來法律的分部——夏威夷司法部（Hawaii Ministry of Justice）。它關心一些面向未來的法律問題，包括預測未來案