政治科學論叢／第 96 期／民國 112 年 6 月／頁 139-178

人工智慧協作下的公共行政研究：對公部門組織議題的多層次反思*

黃心怡**、陳敦源***

摘　要

近十年「人工智慧」（Artificial Intelligence, AI）技術與應用的成長與突破，讓政府組織認眞思考採用與普及的可能性，以期協助政府數位轉型。除了科技樂觀論者的倡議，在政治傳播、市場經濟與社會發展領域浮現許多負面影響後，政府官僚的核心價值與管制角色重新回到議論中心。本文認爲在探究 AI 的組織影響時，應重新以韋伯式理想官僚理論與組織人本主義的碰撞作爲上層理論視角，再探公共行政研究中的組織角色、控制與溝通等核心問題，而最關鍵的研究問題是：「這些新興數位科技與機器將如何影響政府正式組織的模式與運作？」作爲此系列研究的開門磚，本研究旨在分析 AI 進入公部門實務下的公共行政理論與應用，以多層級分析觀點：分別從鉅觀、組織與微觀層次，討論三個關鍵問題。一、公共價值錯置的問題。二、組織內多重代理人的問題。三、AI 裁量權與自動化偏誤的問題。本文於結

* DOI:10.6166/TJPS.202306_(96).0004
本研究紀要的研究文獻與資料的收集與整理，感謝國發會委託研究案「公部門機器演算法應用之制度調適與路徑分析」（NDC-MIS-110-001）與科技部專題研究計畫「官僚組織中的人工智慧應用與人機合作：組織理論的再探」（MOST-111-2410-H-002-069）的研究經費支持。作者也特別感謝三位審查人與編輯委員們的建議與提點。

** 荷蘭萊頓大學政府與全球事務學院公共行政系助理教授，E-mail: h.i.huang@fgga.leidenuniv.nl（通訊作者）。
*** 國立政治大學公共行政學系教授，E-mail: donc@nccu.edu.tw。

收稿日期：111 年 7 月 15 日；通過日期：112 年 3 月 21 日

臺灣大學學術期刊資料庫

論中提出政府組織面對 AI 浪潮時，應該重新思考政府組織的變革底線、途徑選擇、才能管理、以及分工「再」設計等反思面向。

關鍵詞：組織人本主義、官僚理性、人工智慧、價值錯置、多重代理人、自動化偏誤

臺灣大學學術期刊資料庫

AI 不需要是邪惡的才會去消滅人類，因爲如果 AI 推進的目標剛好碰到人類成爲阻礙，它想都不想地就會排除障礙，並且沒有一絲難過。

伊隆・馬斯克（Elon Musk）[1]
電動車大廠特斯拉（Tesla）總裁

壹、前言：當「人」不再是管理的唯一核心，組織該往哪裡走？

提升政府對內行政效率與效能並強化對外與公民的責任與服務品質，是新公共管理與組織再造理念中的重要核心之一，而政府組織「不斷改革論」的重心，不論是推動者與接受者，都是以「人」爲核心的理論發展，但是，20 世紀兩個思潮的發展，逐漸在 21 世紀的上半葉，產生了巨變。首先，1980 年代「新公共管理」（New Public Management, NPM）思潮興起，使傳統「韋伯式」的科層體制在理論與實務都受到嚴厲挑戰，包括參與式管理、扁平組織、治理取代政府、以及社會領導政府等理論的蓬勃發展，政府組織的中心性角色鬆動，數千年以來，以少數菁英擁有權力爲中心的組織運作模式受到挑戰，當然，從政治視角出發來觀察這樣的趨勢，基本上正是美國政治學者杭庭頓（Samuel Huntington）描繪「第三波民主化」（The Third Wave Democratization）對於政府治理所帶來的餘波影響；再者，20 世紀末與 21 世紀初「資訊通訊科技」（Information and Communication Technologies, ICTs）與「社群媒體」（Social Media）的快速發展，一方面強化了前述新公共管理以顧客爲需求的去中心化與由下而上的政府治理趨勢，但也讓學者們反思 NPM 理論的缺陷與不實際之處，如行政數位化所需要的再整合與政府監督，進而發展出「數位時代治理」（Digital-Era Governance）應思考的新管理典範（Dunleavy et al., 2006；Margetts & Dunleavy, 2013）。

[1] 原文如下：“AI doesn’t have to be evil to destroy humanity – if AI has a goal and humanity just happens to come in the way, it will destroy humanity as a matter of course without even thinking about it, no hard feelings.”

臺灣大學學術期刊資料庫

然而，這個發展看似以「人」爲中心，但隨著資通訊科技被更廣泛的應用的過程，且被視爲能達成此目標之利器，自新公共管理以降去中心化並且極大化個人選擇的改變，逐漸地也走向去人性化的「機器」控制之再中心化。舉例而言，資訊科技的發展讓政府與官僚體制得面臨新挑戰與問題（Barth & Arnold, 1999），如 Liou & Korosec (2009) 發現在推動資通訊科技（ICTs）較積極的美國州政府中，許多任務逐漸被 ICTs 機器所代勞，除了漸漸減少的組織官僚層級，也開始縮減政府人員的聘僱。或許人力被迫縮減改變，有些學者認爲資通訊系統的使用並不會改變公部門組織的三角形科層結構現狀，它只是複製原本結構與指令流動的樣貌，但上級管理者透過資訊科技將更能掌握基層文官的決策作爲，利用更多科技人造物與資料結果以限縮官僚執行的能力（Danziger & Andersen, 2002；Jorna & Wagenaar, 2007）。另早期研究「資訊決策系統」（Decision Support System, DSS）導入公部門組織的研究也指出，資訊與通訊化使政府人員由面對民眾的「第一線官僚」（street-level bureaucrats）轉爲僅面對電腦系統的「系統官僚」（systems-level bureaucrats），認爲基層官僚的自由行政裁量空」間已縮小甚至是移轉爲系統裁量（Bovens & Zouridis, 2002）。在 Bovens 與 Zouridis 提出論文的 20 年後，我們看到資通訊科技與新興科技的蓬勃成長，如大數據、區塊鏈、虛擬實境、物聯網 Web 3.0 與「人工智慧」（Artificial Intelligence, AI）技術等新興科技的潛力與應用企圖，正以更先進的技術在模擬人類的諸多行爲，進而開發處理與自動化更困難工作的科技工具，這些工具遠期來說，都有去人性化的展望，並且常隱含機器比起人更能有效無誤地完成工作，這些新興技術對於政府治理效能的提升潛能，或是說科技趨動的治理轉型，不論從學界或實務界，都產生偌大的正面期待（丁玉珍、林子倫，2020；黃東益，2017）。

特別是 2022 年年底，OpenAI 公司的 ChatGPT 橫空出世引起震撼之後，各類組織面對 AI 來勢洶洶的影響，本文針對公共影響的系統性反思顯得更加急迫，然而面對此新興資訊科技具有的破壞性創新意涵，從社會科學的理論建構層面，本文認爲在思考 AI 對公共行政的理論影響時，應該跳脫單純研究組織對此科技的採用與擴散，抑或全然以爲法律制度的建立將優於科技發展與社會應用。公共行政作爲高度社會應用性的學科，認爲社會與

臺灣大學學術期刊資料庫

科技的互動與整合是持續不間歇的，唯有不斷務實地辯證與反思，才能讓人類在此過程中進步與掌握控制權，特別是進入政府之後的人工智慧，學界與實務界都應投以更多的關注（黃心怡等，2021）。當代 AI 的發展與成長極爲快速，公部門得以應用之領域十分多元，此不易回頭的數位趨勢，在可遇見的未來或有以下的組織變化，如：可能是 AI 取代人類不願從事或智能自動化後可更有效率的工作，也可能是人類需學習接受人工智慧機器的資訊作爲輔助自己工作與判斷的工具、抑或者是兩者有相同的決策與行動力－如美國國防部所研擬的人類與「機器人團隊」（human-robot teaming），能共同協作完成任務。這當中存在新的機會但也存有不確定與風險。如知名導演史蒂芬．金曾說：「AI 的興起究竟是有史以來對人類來說最好、還是最壞的事情，我們並不知道」。[2] 但從組織層面來思考這個課題時，AI 的好或許是能更貼近韋伯所描繪的官僚理性。而 AI 的壞則可能來自對於人性的基本原則的妥協，因此本文的旨趣在於討論這些新挑戰將如何與行政管理學派中古典韋伯理論的「理想型官僚」（ideal-type bureaucracy）精神，以及強調以人爲本的「組織人本主義」（organizational humanism）（Denhardt & Catlaw, 2014）碰撞後的對話。[3]

---

[2] 原文如下：The rise of powerful AI will be either the best or the worst thing ever to happen to humanity. We do not yet know which (by Stephen Hawking).

[3] 審查人特別提及韋伯官僚理論與 AI 的關係在本文需要釐清，韋伯理論是眾所周知的行政管理學派代表與官僚理論的重要起源之一，新的 AI 發展與韋伯「理想型」官僚制度的原點開始對話，是近年許多討論 AI 與官僚體制論文的必有橋段（Bullock et al., 2022；Stivers & DeHart-Davis, 2022），本文亦是依循這樣「從原點談起」的論述框架。透過此文，作者們希冀更加釐清兩件事，第一，本文除了依循目前學界討論 AI 對傳統官僚體系的影響，從韋伯的理論爲起點之外，我們也加入「組織人本主義」的這個概念來討論，主要原因是在於，組織人本主義的「呼籲」，過去有部份討論是針對韋伯官僚體制論內「去個人化」的反動，因此目前 AI 的出現，技術上來說恐怕是更能協助韋伯達成「去個人化」理想型組織的情況下，此時此刻面對 AI+ 韋伯官僚體系的「去個人化」風潮，挑選組織人本主義切入討論是深具意義的；第二，審查人有提及「人群關係學派」在本文當中的定位，作者必須聲明，人群關係學派是對科學管理學派的一種經驗上的反思，經過 20 世紀初「霍桑實驗」（Hawthorne Studies）的展示，組織管理者需要將個人或小團體的動態納入導引組織績效的考量。當然，AI 的發展也將再次引起組織理論的專家們重視，討論人類進入 AI 化社會的適應問題（Vogt, 2021）。因此，本文選擇範圍較廣的組織人本主義來討論 AI 與傳統官僚制度的對話媒介，而不是組織理論中人群關係學派的經驗

臺灣大學學術期刊資訊庫

面對新興數位科技的積極湧進，充分認識且理解此變革對公部門組織的深層影響至關重要。作爲一篇「研究紀要」的論文，筆者期待以本文向學界與實務界介紹此研究之趨勢現象，以及 AI 對公共行政研究中組織研究之核心爭議的回顧。討論政府治理行爲與模式將被如何影響。如提出「演算法治理」（algorithmic governance）概念的 Rouvroy 等人所言，在不久的將來，政府治理勢必不能忽略新興資訊科技的核心－AI 演算法，以及它所嘗試建構的社會秩序與管制工具（Rouvroy et al., 2013）。透過文獻回顧方法，本文檢閱重要公行領域的國際期刊，近十年來所發表的 AI 與政府之相關論文，搭配組織理論與行爲重視的多層次分析來進行鋪陳與論述 AI 發展下，人機協作與政府治理的不同挑戰。此文定義的人機協作爲：「人類公務員與 AI 智能機器或系統，共同存在政府組織中，各自具備完成政府業務的能力，或許是各自單獨執行被指派之任務，又或是雙方存在合作或輔助關係，共同完成任務」。搭配 Burke & Litwin (1992) 之組織三層次模型，綜合分析討論公部門數位轉型趨勢下應用 AI 科技帶來之影響。下面接續針對本文的文獻資料與分析架構進行說明。

貳、文獻分析與焦點論述

從各國陸續公布之 AI 政策白皮書可發現，公部門多積極將大數據 AI 與機器學習演算法視爲是推動政府數位轉型所需佈署的重要技術之一，寄望以此提供公民智慧服務或輔助公務工作。然有關 AI 技術究竟對組織與其中之個人產生什麼影響，學者間尚無定論，且對機器的智能發展有不同的想像。樂觀派學者認爲 AI 未來能全方位協助且賦能人類、如 1950 年代設計經典「圖靈實驗」（imitation game）的「圖靈」（Turing）好奇電腦能否模仿人類的說話方式，並讓人類受試者無法分辨是機器在與自己對話，該實驗精神對後續 AI 的發展有著極大影響。懷疑派學者如 Hubert Dreyfus (1972) 指出 AI 最終都不會達到人類的智慧能力，首先是它永遠無法模擬人類的經

研究觀點而已，作者也認爲，組織人本主義由於概念較廣，應該與人群關係學派有高度的重疊。

臺灣大學學術期刊資料庫

驗，只能透過歷史資料學習、處理和計算事實。許多當代學者也附和這個看法，認爲人類大多的智能屬於隱性知識，現在的 AI 技術與演算法都不足以達到「自主代理」的目標（Beverungen & Lange, 2018），未來應該也不會。然隨著西洋棋「深藍」（Deep Blue）與圍棋 AlphaGo 的開發，乃至自動車、智慧語音、智慧聊天機器人等應用技術的成熟，開始有務實派學者認爲 AI 的能力進程將漸近地達到具有思考、學習與解決特定任務的能力，可能強化也或許會弱化人類的部分能力，但不至於完全取代人類（Makridakis, 2017）。

然而，無論是樂觀派、悲觀派或是務實派的未來學者們的想法，多單純僅從科技的智能是否與人類智能類比的觀點出發，聚焦在單一機器（人造物）與單一個人的關係。但從公共行政與組織理論的角度來看，作爲一務實應用的學科，人的社會性與組織性鑲嵌，讓公行學者們不應也不容忽視 AI 對於公共組織與內部行爲的不同層次之影響。因此在開展本文論述與討論前，下一段落，我們將依序介紹本文盼與其對話的「組織人本主義」與「韋伯式理想官僚」理論。

一、組織人本主義

組織人本主義理論重視個人（非工具性）的「內生動機」（intrinsic motivation）與組織經濟效率的關係。源於新公共行政對於 20 世紀早期的科學管理學派「去人性化」（de-human）的批判（Levine et al., 1975），也是從著名的「霍桑實驗」（Hawthorne effect, 1927-1932）中所發展出人際關係學派所謂 Y 理論的價值基礎。首先，組織人本主義的發展，根據不同管理學者的描繪，主要是在組織相關的管理理論發展中，不斷回頭檢視「以人爲本」（“put the people first”）的原則，呼籲在手段選擇的過程中，也在組織生存的競爭中，皆不能被犧牲（Alderfer, 1994；Larkin, 1986；Ramsey & Calvert, 1994）。

相關的討論基本上都圍繞在 1960 年代開始的三波理論發展（Melé, 2003），第一波是討論人們的動機，這個時期的組織管理理論，試圖從科學管理學派將人當作機器的窠臼中掙脫出來；第二波是討論管理的人本主義原則的應用，包括領導統御與組織文化等管理介入技術的發展，也如何能

臺灣大學學術期刊資料庫

同時達到組織目標又不違背以人爲本的原則，最後一波是在 20 世紀末，討論一個組織如何營造適合人們生活的環境，因此組織公民行爲、組織靈性、甚至後來的個人與組織配適度和工作/生活平衡等領域，都是這一個組織人本主義的脈絡下的發展。綜括來說，組織人本主義可以簡單定義爲：「一種不斷檢視新興的科技與管理技術所發展出來的管理介入工具，是否能夠同時兼顧以人爲本與管理效率（能）的追求，並保障組織中人們之生命存續、尊嚴維護、與期望滿足等的組織理論視角（Ramsey & Calvert, 1994）。」但無論是從「控制理論」（control theory）、「權變理論」（contingency theory）、或是「科層秩序」（hierarchical order）的角度來思考組織人本主義，都會發現要達成組織的「人本原則」（humanistic principle）是相當困難的（Alvesson, 1982），也容易與技術理性的論述產生衝突。固然不少組織人本主義論者認爲，這種「人本」原則應放諸組織內正式與非正式各個階層的職務，但已有實證研究顯示，如職等越低的人員與任務，越容易與經濟以及技術效率妥協，並犧牲人員的自我實現等人本原則。

本文聚焦在 AI 這個自網際網路發明以來，最重要的數位破壞性創新，藉此文討論在政府推動數位轉型過程中，是否可能引發更多的人本主義危機。從技術發展的角度來看，AI 科學家與工程師在發展這類科技時，背後的終極目標，是希望他們的發明與技術能讓其 AI 或智慧機器達到自主代理，或是所謂的「自動代理人」（autonomous agents）。它意味著機器可以超越初始的規則與編碼，並開始「學習如何學習」（learn how to learn），當資訊系統與機器人的工作表現，能與人類一樣甚至更好且低成本時，人類將與機器一同工作，甚至最終被取代，因此，傳統組織理論與行爲中的人類角色與功能，勢必受到嚴厲的挑戰（Barth & Arnold, 1999；Makridakis, 2017）。

因此，AI 對於組織的影響可能是根本性的，它改變了組織內「工作」的意義（Shestakofsky, 2017），兩位挪威學者指出（Wilson & Broomfield, 2022），由於 AI 的知識對於組織疆域的破壞性極強，組織內的管理者要如何藉由動態的信任與治理網絡，進行 AI 相關的知識分享與學習，是當代組織管理者一個重要的議題；而早在 1999 年間，公共行政的學者就已開始尋找 AI 對於政府組織運作「特殊」的影響，當然，那時基本上對於 AI 可以輔助人類降低決策錯誤或是公開設立決策目標、動機與價值有著較爲樂觀

臺灣大學學術期刊資料庫

的評價，如 Barth & Arnold (1999) 所言，AI 系統整合且強化組織決策者的能力，可望發展成具備目標、動機、與價值的智能工具，同時也不用被明顯的人類決策錯誤因素：「生氣、害怕與偏見」所轄制。

持審慎樂觀態度的學者認爲，AI 等科技的發展，反而逼迫學術與實務中人，從新檢視組織人本主義中將要面對的根本性挑戰。舉例來說，如果 AI 未來的發展會導致人類在 Herbert Simon 所謂「組織生活世界」（organizational life-world）（Spender, 2013）被邊緣化，那麼第四波工業革命是否將會是一波組織人本主義的浩劫？2020 年管理學門的頂尖期刊 *Academy of Management Review* 就以主編之名，發表一篇 "The Future of Organizational Theory" 的短文（Baum & Haveman, 2020），文中提出 AI 與全球化是組織理論在 21 世紀應該面對的兩大挑戰，針對 AI 的問題，作者說（Baum & Haveman, 2020: 270）：[4]

> 如果人類無法了解機器學習的演算法是如何運作，學者要如何去發展出可以駕馭演算法的組織理論？一個可能的途徑就是，將組織適應與執行機器學習或演算法的情境，等同於過去追求創新之流行物件時的作爲。

總括來說，面對資通科技的快速發展，許多有識之士都認爲人類將面臨一個「無法回頭的奇點時刻」（The Singularity）（Kurzweil, 2006），在行焉不著、習焉不察的漸進過程下，人類的生物體將與智能科技體糾結融合，人與機器的融合將無可避免，像人類目前已經無法離開手機一般形成一個混合存在的「裝配體」（assemblage）（Lanzara, 2009），如果進一步延伸這樣的部份已經成形的未來人本發展預言來看，因爲人的獨立存在隨著科技發展而產生「衰微消失」或「重生壯大」兩種相衝突的可能，公共行政長久以來應用組織人本主義來思考公共組織的研究領域，除了從個人層次討論人機融合的議題之外，當然也應從宏觀與中觀的組織層次來討論這個未

[4] 原文如下："If humans cannot understand how ML, algorithms works, how can scholars develop theories of organizational that harness such algorithms? One possible approach is to treat the adoption and implementation of ML algorithms as we have the adoption of many earlier management innovations, fads, and fashions."

臺灣大學學術期刊資料庫

來的發展，特別是從下面韋伯的理想官僚理論開始談起。

二、韋伯式理想型官僚理論

另一個在政府 AI 推動過程中，值得用以反思的公共行政理論是「官僚理論」（bureaucracy）。自馬克斯．韋伯（Weber, 1922）提出其極具貢獻的官僚體制理論後，已然成爲當代政府組織治理結構與制度的典範。韋伯指出「理想官僚體制」（Weberian bureaucratic ideal type）所代表的「法理型權威」（rational-legal authority），是相較於「傳統型權威」（traditional authority）與「魅力型權威」（charismatic authority）更爲理想的治理模式。[5] 其官僚體制以法理型權威合理化民意代表、政府行政者與市民社會的委託代理關係，具有強調「科層化」（hierarchical）、「功績制」（merit-based）、「標準化」（standardized）、「形式化」（formalized）、「特殊化」（specialized）等機制設計，如組織具嚴明的層級制度，分工高度明確且受到統一的紀律控制。「理性」（rational）與「控制」（control）是理想官僚制度中組織結構的兩大重要屬性。一個官僚從選任開始就獨立於民眾影響（非民選），讓官僚能夠客觀地面對民眾，不受到民眾個人身份的影響。但公共行政機關與人員需理性地「服從法律」（rule of law）、「行政程序」（due process）、「適切行爲準則」（codes of appropriate behavior）（Olsen, 2006），上述的要求，國內學者多稱爲「依法行政」。再者，在升遷進用與薪資上依照表現、功績、專業度等客觀標準辦理，給予合理的待遇，並爲終生職保障官僚職涯發展。

從以上的論述可以看出韋伯的理想型官僚制度，是假設官僚組織與內部個人能具備「理性客觀、依法行政，獨立於外部利益影響，且遵守組織紀律」等進行決策與執行公務的特質。韋伯（cf. Bullock et al., 2022）曾舉例說過：

> 認爲法律是毫無漏洞的想法，通常會被世人強烈質疑。此外，當代社會對於政府司法長官僅不過是一臺「審判機器」的想法也感到厭惡。但在這樣的審判機器中，把檔案與費用被丟入機器頂部

[5] 用韋伯的語言來說，所謂「理想型」（ideal-type）並非認爲「官僚體制」（bureaucracy）是組織體制的唯一解，而是相對其他歷史上的統治權威體制而言，是較佳的選擇。

臺灣大學學術期刊論文資料庫

> 的收件入口，就可以從機器底部吐出判決書與機械性推理……相較於那些強調個性化方法的說法，不變的法律規範將對其產生負面效果，因爲他限制了「公僕」(beamte) 的那些不被監管的創造性工作，這應當被視爲是正面積極的事。[6]

當然在韋伯的時代裡，不存在此情境敘述的審判機器科技，但當時的韋伯將人類法官與政府公職人員（德文原文：beamte）形容成這樣的機器典型，也顯示了在理想型官僚體制中的公僕，被期待的不僅是理性與專業，更需有去人性甚至是「機械性」(mechanical) 的作法 (Bullock et al., 2022)，因而出現 AI 似乎更能符合理想型官僚的想法。公共治理範型從韋伯的官僚體制，到新公共管理強調最大化效率、市場邏輯與顧客滿意的變革治理。組織理論的發展早期是因著工業化生產的管理目的應運而生，發展過程中主要以「人類代理者」(human agency) 爲主要核心 (van Rijmenam & Logue, 2021)。但進入 2020 年的現代，智慧數位機器已然達到更高的技術程度與能力，現在的 AI 與機器演算法科技，透過大量資料的訓練與學習，讓機器具備能做出自動判決的能力，甚至具有後續不需依靠人類給予指令，但能持續學習與修正之能力，也衍生了是否存在「AI 代理者」(AI agency) 或「AI 機器官僚」(AI machine beamte) 的想像 (Bullock et al., 2022)。我們預見此發展對於公部門組織的現在與未來將會有決定性的影響，也似乎更直接且正面迎向韋伯理想型官僚的盼望。猶如 Olsen (2005) 的建議，在不同的時代，我們都需要持續再探韋伯的官僚理論，不僅是將它視爲組織建構的「工具」(instrument)，而是將其視爲政府組織的「制度」(institution)，因爲我們眞正需關心與努力的，是那些讓公共行政達到理想型官僚目標的組織機制。從公共組織理論的核心關懷出發，本文接續透過國外文獻搜集

---

[6] 原文如下：“The idea that there can be a law without loopholes is generally strongly contested. Also, the idea that the modern magistrate is nothing more than a ‘judging machine’ is rejected with disgust. In such a judging machine, files and costs would be thrown into the top of the machine in order that it would spit out the verdict along with a mechanical reasoning at the bottom..... In contrast to such emphases on individualized approaches stand the invariable norms that play a negative role because they restrict the never-to-be-regulated ‘creative’ work of the Beamte, which is normally seen as a positive thing.”

臺灣大學學術期刊資料庫

與整理，反思並提出近年來 AI 科技對公部門組織影響文獻的歸納與討論。

三、公部門應用 AI 的文獻搜集與回顧

在上述的兩小節當中，作者藉由討論「組織人本主義」概念，以及它與 AI 在公部門組織的發展爲開始，且介紹公共組織不得忽略的「韋伯式官僚理論」。提醒我們在新興科技的浪潮衝擊下，更需認識公共組織的本源與核心，更遑論當中可能產生的兩種治理邏輯的碰撞與衝突問題。藉由文獻蒐集與回顧，且借助 Burke-Litwin 的歸納組織理論之「宏觀－中觀－微觀」（macro-meso-micro）的多層次結構，進行智慧數位科技對組織人本主義與官僚體制的反思。接下來，本文嘗試從文獻回顧中連接組織理論的框架，對於 AI 對公共行政未來的影響，尋找組織人本主義對 AI 發展的一個論述框架。

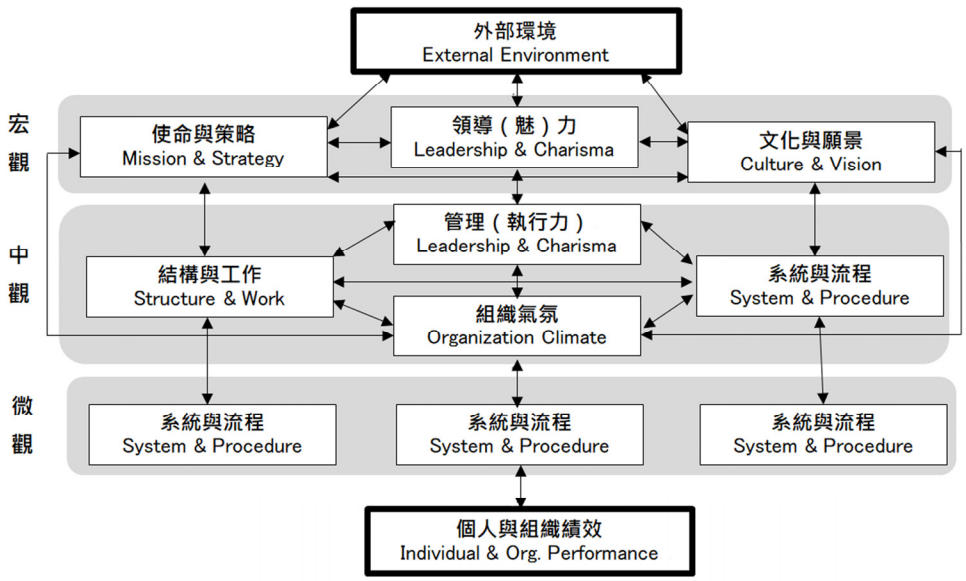
無論是國內或國外社群，以 AI 爲主的研究範疇在公共行政領域尚在發展階段，近年來此議題更被國際學者所重視，根據 Zuiderwijk、Chen 與 Salem 的研究，AI 在公共治理領域的研究也正在蓬勃發展中，甚至有學者認爲 AI 可被視爲公部門的「破壞式創新」（disruptive innovation），將改變公部門內的組織、流程與職能結構（Zuiderwijk et al., 2021）。本研究紀要希冀能以公行組織觀點研析與整理 AI 對政府組織的影響，因此並不考慮已在 AI 領域投入已久的其他工程領域或資管領域論文。我們從公共行政與公共管理領域 Q1 等級的期刊中搜尋品質較高文獻，先排除掉特定政策領域期刊（如：*Educational Administration Quarterly*, *Criminology and Public Policy*）與較關注私部門組織理論課題的管理領域期刊（如：*Administrative Science Quarterly*）。筆者最後選取國外十本公共行政與公共管理領域期刊來搜尋相關文獻。[7]

[7] 這十份期刊的名稱如下：(1) *Public Administration Review* (PAR), (2) *Journal of Public Administration Research and Theory* (JPART), (3) *Journal of Policy Analysis and Management* (JPAM), (4) *Review of Public Personnel Administration* (ROPPA), (5) *Policy Studies Journal* (PSJ), (6) *American Review of Public Administration* (AROPA), (7) *Public Management Review* (PMR), (8) *International Public Management Journal* (IPMJ), (9) *Policy and Internet* (P&I), (10) *Government Information Quarterly* (GIQ)。

臺灣大學學術期刊資料庫

選定期刊後，選定搜尋區間爲 2000 年 1 月 1 日到 2022 年 5 月 20 日。在 Web of Science 的期刊資料庫系統上，執行下列搜尋語法。[8] 搜尋結果找出了 324 篇文章，但刪除重複、書評與應用領域與政府無關的文章後，最後共 113 篇學術論文作爲本篇研究紀要的參考。本文檢閱這些文獻後，依照 Burke & Litwin (1992) 的三層次歸納模型，進行文章分類、問題意識歸納並作分類比較。如附圖 1 所示，兩位學者將組織理論與行爲分爲：(1) 組織宏觀層面（藍）、(2) 組織中觀層面（綠）、與 (3) 個人績效層面（紅）等三

圖 1　組織績效與改變模型



資料來源：修改 Burke & Litwin (1992: 528)

[8] 這語法如下：(“Artificial intelligence” OR “AI” OR “machine learning” OR “deep learning” OR “reinforcement learning” OR “supervised learning” OR “unsupervised learning” OR “neural networks” OR “natural language processing” OR “computer vision” OR “image recognition” OR “facial recognition” OR “face recognition” OR “machine algorithm” OR “speech recognition” OR “intelligence systems” OR “virtual assistant” OR “autonomous vehicle” OR “predictive analytics” OR “robotics” OR “self-driving”) AND (“governance” OR “government” OR “public management” OR “public governance” OR “public sector” OR “public administration” OR “public policy” OR “public transformation”) AND (“discretion” OR “behavior” OR “decision making” OR “behavior” OR “administrative”)。

臺灣大學學術期刊資料庫

個層次，此架構類似制度論分析層次的「宏觀－中觀－微觀」的結構，適合用來討論以「人類代理人」在 Simon 所說組織當中的「工作世界」中各種管理作爲的交叉作用場域，可以當作討論AI技術與組織發展綜合影響的完整架構。

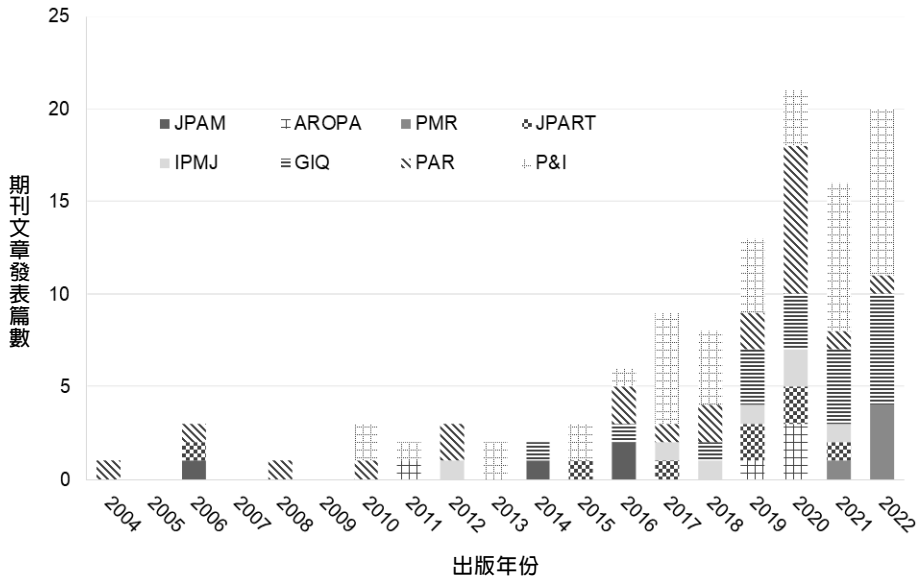
接著，圖 2a 呈現 AI／機器演算法議題相關之學術論文於選定之國外公共行政與公共管理領域期刊的刊出篇數。我們可明顯觀察到自 2016 年後，有關 AI 或智慧數位科技應用於公部門的研究有顯著成長。在過去近二十年間，在此議題上關注最多的期刊爲 P&I（42 篇）、GIQ（23 篇）與 PAR（19 篇）三本期刊。據此，筆者在閱讀搜尋出來的文獻時，根據該論文之標題、摘要與內文，將個別文章之分析層次做了分類，圖 2b 即呈現編碼後的分類統計，清楚看見自 2004 年開始，研究論文多是以組織宏觀層次爲主（共 45 篇），探討許多 AI 在政府組織的使命與策略、對公共治理的宏觀意義、智慧機器背後代表的公共價值（如效率、課責、公義、人權保障等）（Vogl et al., 2020；Zuiderwijk et al., 2021）。組織中觀層次則有 20 篇，論文探討的議題包括組織結構、組織資源與能力、組織流程與行政分工（Fountain, 2022）、組織內外學習與知識分享等（Malomo & Sena, 2017）。在個人績效層面則有官僚裁量認知與行爲決策等議題，也有探討個人的專業能力、不同職位對 AI 或智慧機器之態度、或對個人業務判斷與專業能力影響等題目。筆者也在此 113 篇文章中，觀察到有 3 篇屬於文獻探索與系統性回顧的文章，也有接近四分之一的文章屬於運用 AI 演算法如機器學習等方法來精進公行研究或強化機關對政策問題發現與政策評估的效用與表現。

另外，不少文獻嘗試提供 AI 或是「智慧機器」（intelligent machines）的分類，以及可應用於公部門組織的哪些領域。首先，筆者觀察到 AI 與演算法的關係密不可分，根據 Hurbans (2020) 的整理，AI 包含傳統的「搜尋演算法」（search algorithm），也包括新形態的「機器學習」（machine learning algorithm）、「深度學習」（deep learning algorithm）與「生物啓發演算法」（bio-inspired algorithm）。不少學者選擇用「演算法治理」（algorithmic governance）來說明一個新的公共行政治理挑戰（Danaher et al., 2017；Gritsenko & Wood, 2022；Katzenbach & Ulbricht, 2019；Rouvroy et al., 2013）。儘管此名詞的定義尚未有定論，Katzenbach 等學者的註解蠻能表達此名詞的內涵：「演

臺灣大學學術期刊資料庫

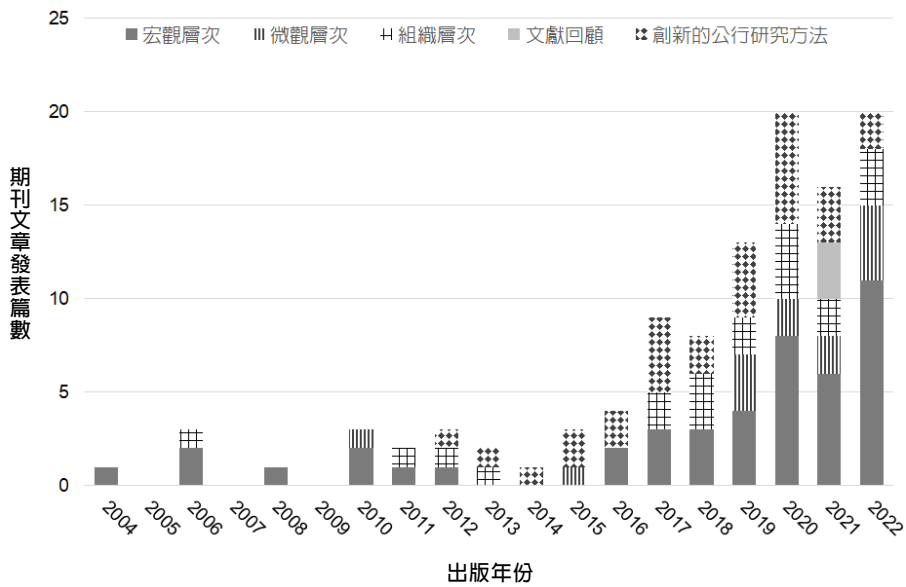
■ **與 AI 相關論文發表於國外公共行政（PA）與政府公共管理（PM）領域期刊的趨勢**

圖 2a 比較不同期刊的歷年發表篇數



資料來源：作者自繪

圖 2b 比較歷年期刊發表的研究分析層次



資料來源：作者自繪

臺灣大學學術期刊資料

算法治理是某種社會秩序的型態，基於既有社會規則並結合複雜數學與電腦計算的認知流程，使系統中不同行動者得以互相協調。」[9] 學者 Vogl 等更進一步提出「演算型官僚」（algorithmic bureaucracy）將成爲新興行政模式，意指借助電腦演算的力量以協助地方政府與機構執行公共服務（Vogl et al., 2020）。以下各節將進一步在各分析層次上，闡述數位轉型與 AI 將對公部門組織內涵與理論帶來之挑戰。

不止步於國外文獻論文的描述性分類與新概念介紹，本文將更進一步以「宏觀－中觀－微觀」結構三個組織理論層次爲分析單位，依序討論圖 1 當中，公共行政研究面對 AI 在公部門組織的進擊發展，引領與討論「組織人本主義」與「理想型官僚」碰撞後產生的理論衍生之實務挑戰與問題。

參、當 AI 進入公部門組織：組織宏觀、組織中觀與個人微觀層面的影響

一、組織宏觀層面：AI 帶來的公共價值錯置的問題

AI 對於組織的宏觀層面帶來的影響主要在於價值追求終極目標、價值選擇以及價值衝突等「價值錯置」（value misalignment）的問題。2015 年有三位英美的理工科學者 Russell, Dewey 與 Tegmark 在《AI 雜誌》上發表了一篇名爲〈追求強健與有益的人工智慧之優先研究項目〉（“Research Priorities for Robust and Beneficial Artificial Intelligence”）的文章中，提到人類社會面對人工智慧的狂潮，最重要的研究工作除了聚焦推動技術層面的進步之外，我們還要科學且審慎地去檢視人工智慧的發展（Russell et al., 2015）。一方面可以因此獲得其後的龐大利益，另一方面則是希望可以避免它給人類帶來的可能災禍。因此，他們說：「現在正是時候，我們應該聚焦在人工智慧相關研究，不只是提升其能力，還要能夠極大化它帶來的社會利益」[10]，

[9] 原文如下：“Algorithmic governance is a form of social ordering that relies on coordination between actors, is based on rules and incorporates particularly complex computer-based epistemic procedures”.

[10] 原文如下：“Progress in AI research makes it timely to focus research not only on making AI more capable, but also on maximizing the societal benefit of AI.”

臺灣大學學術期刊資料庫

特別是在高度價值導向的決策場域中，人工智慧的發展無法不積極思考，要如何更有效率地將人本的價值選擇置入人工智慧的發展願景當中；當然，這樣將一個原本技術導向爲主的科技發展項目，擴大到「機器倫理」（machine ethics）領域的呼籲，回到前述組織人本主義與韋伯官僚理論對應下的人工智慧發展，對於政府組織來說可以算是一種組織人本主義的覺醒，這三位作者當然清楚這拉近機器運算的一般性原則與人類價值體系的發展有著根本性的困難，他們說（Russell et al., 2015: 110）：[11]

> 長期來說，人工智慧的載體能自動且強大地在不同領域遊走將極可能達成，但是，將人類的偏好藉由機器倫理的概念置入這個跨域的人工智慧載體實務上會有困難，因爲，要將強大人工智慧載體的價值體系，與我們人類的價值與偏好體系，進行「同步配置」（同置，alignment）的工作是困難的。

不過，對於理工與社會科學領域的研究者來說，這個困難是必須要設法克服的。因爲，如果人工智慧所產製出來的成果雖然滿足快速、自動、與效率的原則，但是卻違反了效能價值中的結果偏好以及公平的價值，那麼前述人工智慧的發展要避免給人類帶來災難的期待就會落空，比方說，人工智慧可以輔助民主政體中的選民作出更爲「知情」（informed）的選擇，但是同樣的科技也可以被用在非民主體制中對於人民進行更完美的監控。科技的應用的跨域性與效率或許是一致的，但對於人類追求人權與民主自由的價值來說，其後果是有天壤之別的；因此，正如 Herbert Simon 早已談論過組織決策往往同時得在「眞實」（facts）與「價值」（values）中來回選擇，人工智慧的發展與應用過程所引發的價值錯置問題，或許是人類過去的決策行爲即存在價值偏誤，當然也存在著 AI 機器的演算過程與閥值將具有價值偏誤資訊又回饋給人類的情況（Bullock et al., 2022）。因此，從「複雜系統」（complex system）的角度來看，本文認爲人工智慧進入組織產生

[11] 原文如下：“In the long term, it is plausible that we will want to make agents that act autonomously and powerfully across many domains. Explicitly specifying our preferences in broad domains in the style of near future machine ethics may not be practical, making aligning the values of powerful AI systems with our own values and preferences difficult.”

臺灣大學學術期刊資料庫

的價值錯置問題，可被定義爲：「人工智慧載體所依循的價值體系，明顯與人類與組織社會所追求的價值產生不協同，導致人類與機器同工中所產生的結果，明顯違反人類的價值認知，進而產生負面的社會影響。」

接下來，本文依序從機器主權、價值衝突、以及人類錯誤三方面，更細緻地討論人工智慧衝擊到價值錯置與組織人本主義的可能發展。

首先，機器主權與人類主權：本文所強調的組織人本主義，其興起最主要的目的，就是宣示「組織理性化」（organizational rationality）的許多工具性的非人性發展中，人的存在這個「價值」應該「永遠高過」其他所有組織應用之工具的價值，不過，人工智慧的最高境界剛好是與人類一樣擁有主權，這可能根本性地衝擊到組織人本主義所要追求的目標。2016 年電影《人造意識》（*Ex Machina*）中的機器人 Ava，之所以可以突破成爲人的樣式，就是她通過與男主角 Caleb 的所謂「圖靈測試」（Turing test），展現她想逃出框架的動機下，懂得用性別吸引力騙過人類，殺了她的創造者並且逃脫。前述電影的英文名是翻自希臘文的現代拉丁文，意指「機器中的神明」（god in machinery），這在政治學是一個「主權」（sovereignty）的概念，代表機器擁有自主權，這個可能性不但凌駕了人的存在，也會成爲組織內去人性化的當代推手，這個部分的「危機」對推動政府數位轉型的相關學術與實務菁英來說，應該要正視。

再者，人類本身的價值衝突：以政府爲首的公共組織，因爲需要「全員」（catch-all）的管制與服務，通常會承載多元價值並存的現實，而組織人本主義其中一個重要的意義，就是要保障組織中人們的選擇權，比方說，效率與公平對於政府組織設計來說，起碼在流程上，一方面因爲資源的稀缺性，需要關注資源應用的效率性問題，另一方面，面對公眾利益的分配，公共組織又必須持續聚焦公平性的價值，這就是決策理論中的「多屬性效用模型」（multi-attribute utility model）希望可以在保障個人選擇權的前提下，解決組織決策的多元價值衝突的問題；因此，這個效用模型的機制中，最重要的人本價值就是保障個人的「選擇權」（the right to choose），然而，它卻可能因爲 AI 的發展而受到傷害；舉例來說，現在政府組織中篩選補助對象的流程，可能漸漸會委由機器演算法進行，其過程中個人表達意見的機會被機器表面上的客觀性給壓縮，而其結果如果有機器學習的回饋效

臺灣大學學術期刊資料庫

果，還會直接影響後續政府進行資源重分配的選擇框架，這個類似將資源配置的「議程設定權」（agenda-setting power）交給機器處理，從價值自主選擇的人本主義角度來看，這種數位轉型的改革當然會給公共事務的決策帶來新的組織人本主義危機。

最後，人類無可避免的錯誤：價值衝突雖然是組織運作最常面對並且需要處理的問題，但是，納入 AI 協作的組織還會產生一個新的價值衝突面向：人類會因自己的偏見下的故意或是無知所犯的錯誤，而將價值偏差複製並輸入演算法或是 AI 的介面，導致判斷結果的偏差。換言之，AI 對於行政流程輔助的結果可能與組織對公務代理人所期待的價值偏好存在明顯落差（Bullock et al., 2022），所謂價值錯置是有雙向影響路徑的可能；因此，即便是演算法或機器學習可以作到價值中立而無意識地照設定的流程運作（李翠萍等，2022），但是協作的人類卻可以藉著機器對價值衝突的無感，從源頭輸入各種自我認知的「正確」價值，進而對社會輸出並激化價值衝突，造成組織內外或人際之間的衝突；根據學者的說法，在 AI 支援之組織運作工具日益蓬勃的今天，人到機器或是機器到人這個「雙向互動」（two-way interaction）的偏誤發生可能，比起前人工智慧時代會更加嚴重（Sun et al., 2020），總的來看，這個情況等於是在組織舊有價值衝突之上，加添了人機互動綜效後的系統偏誤可能，讓後人工智慧時代的組織，處理價值衝突產生能力不但未提升卻還會惡化的危機。

二、組織中觀層面：AI 帶來的多重代理人的問題

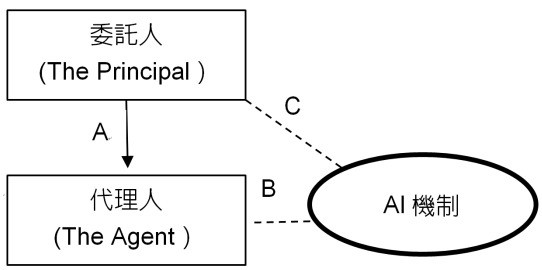
如前所述，組織人本主義非常強調宏觀層次上的人性的選擇權至上原則，但是「新韋伯式官僚國家」（Neo-Weberian State）當中談組織控制的部分卻是組織存在最重要的功能性價值之一（Byrkjeflot et al., 2018），畢竟政府組織的行政權與角色依舊得回歸到正式系統中三大課題，「結構」（structure）、「溝通」（communication）與「控制」（control）的討論。誠如 Hage (1980) 指出韋伯的科層組織理論，其實是以「權力」爲核心在討論組織的內控問題，具體來說聚焦在官僚的管理監督與紀律訓練。因此從控制學派的角度，人本與內在驅動的觀點，無法解決官僚控制中最重要的權威合理化問題。這也會導引出組織層次傳統上必須面對的「代理問題」（agency problem）（Arrow,

臺灣大學學術期刊資料庫

1985）。如有學者提醒，加入 AI 的技術之後的組織決策運作，會產生所謂「AI 黑盒子」的問題（Bannister & Connolly, 2020），由於 AI 技術十分仰賴資料，但所使用的資料屬於主觀資料或客觀資料，恐讓這個技術黑盒子導引出不同型態的風險，更加凸顯組織管理與控制必要的功能性意義。近來文獻也提到，政府組織內可能產生各種新的互動形式，除了人與人、人與機器、還有機器與機器的關係，也可能出現人類團隊、機器團隊與人機合作團隊的概念。這些新的組織成員組成，都需要我們重新設計組織互動與溝通方式以利業務的進行，但也需要新的控制與課責機制來進行風險控制（Bullock et al., 2022）。因此，越來越成熟的 AI 代理人納入社會組織中，讓組織不得不思考「多重代理系統」（multi-agent systems）的出現（Carrera & Iglesias, 2015；Daylamani-Zad et al., 2020），更重要的，學者也認爲，AI 在人類組織中持續擴大的應用，將會是未來代理人理論應用的一個全新場域（Moloi & Marwala, 2020）。

回到傳統的代理人理論，如果我們將兩個人分別賦予「委託人」（the principal）與「代理人」（the agent）的角色，在他們之間組織必須處理包括目標衝突與資訊不對稱等課責控制的管理問題（Waterman & Meier, 1998），當然，當 AI 時代來臨，如果前述的代理人角色，需要與 AI 的決策功能協力工作時，附圖 3 中委託人與代理人的關係，會從傳統 A 關係，延伸出 B 與 C 的關係，進而出現三個管理的問題：(1) AI 作爲獨立個體的身份問題（圖中的 C 關係）、(2) AI 與代理人之間的責任分工（圖 3 中的 B 關係）、以及 (3) 委託人如何獲取足夠資訊用來分辨不良決策是來自 AI 還是人類的問題（圖 3 中 A+B+C 的關係）。

圖 3　AI 進入組織的多重代理問題



資料來源：作者自繪

臺灣大學學術期刊資料庫

首先，AI 作爲代理人的一個前提，應該是它可以像人一樣負責的可能，這個可能性在 AI 的發展上一直被大力強調，但是卻很少討論這樣成爲像人一樣的潛能，在甚麼樣的發展階段必須成爲委託人與之的 C 關係中，一個可以負責任的獨立個體，事實上，就圖中多重代理人的架構來看，AI 機制已經扮演某種「準代理人」的角色，這個角色對與整個代理關係的影響，還受到關係 B 的制約，也就是說，傳統的委託－代理關係 A 的問題，因爲關係 B 與 C 的加入，AI 的導入會在未來成爲雙重代理下與傳統代理人之間，出現利益衝突嗎？關鍵在於 AI 機制能否獨立像人一樣。嚴格來說，目前科技下，AI 機制在是無意識地存在（Buttazzo, 2001），也就是說，它們並沒有像人因認知能量而產生的獨立身分，但是卻會在職場上影響人類的自我「認同」（identity）（Mirbabaie et al., 2022），但是，因爲人類演算能力的限制，AI 機制越來越成爲工作生活中不可或缺的一部分，甚麼時候從管理層面 AI 能夠與人一樣而同等地被管制與課責，仍然是無法預期的，目前是 AI 能夠協助人類工作，但是出錯了仍然是人類要出現負責的狀況。

再者，AI 機制與代理人之間的分工或是協力問題（關係B），也是多重代理人理論在過去最需要處理的問題，這個牽涉到組織理論內因爲無法分辨個人出力或是犯錯的成分，而產生整體績效的「偷懶」（shirking）或是「課責規避」（blame avoidance）的問題（Brehm & Gates, 1997）。當然，目前 AI 還無法自稱權益的前提下，這個關係 B 當中，代理人在論功的時候不需要分享給 AI 機制，但是在組織任務出現問題時可能將責任推給 AI 機制。因而未來這個多重代理人的環境中，委託人面對代理人聲稱自己的錯誤是因爲 AI 機制的「不可抗拒因素」的分辨問題，就顯得格外重要。另一種可能的情境是，由於社會課責與政治課責的箝制，社會上普遍認爲儘管是 AI 犯錯，仍是人類挨罵的情況下，代理人會更嚴格檢查與掌控演算法，使得關係 B 中展現的「演算法做裁決」（adjudicating by algorithm）與「機器人訂規則」（rulemaking by robot）的權力架構仍有（人）高（機器）低之分（Coglianese & Lehr, 2017）。因此我們可以說，過去研究認爲領導是組織控制與激勵的重要因子（Rainey et al., 2021），但在 AI 與數位轉型的發展下，科技本身更可能成爲控制與激勵的驅力，或可能成爲阻力。如在共同工作的情境下，會成爲「動機強化科技」（human-enhancement technology）（Korinek

臺灣大學學術期刊資料庫

& Stiglitz, 2017）或「動機攔阻科技」（human-curtailment technology）（Buffat, 2015）都是有可能的。

最後，傳統多重代理人問題的處理，牽涉到資訊不對稱的機制設計，比方說，亨利福特（Henry Ford）的生產線設計，將生產流程設計成先後可分離的作業程序，讓員工單獨完成個別的工作後，事後不論獎勵或處罰的資訊較爲完整，因此，前述 AI 機制導入組織管理所產生多重代理人問題的解方，可能在於個別績效資訊的蒐集與分析處理，AI 機制不但可以用來協助代理人工作，也可以應用個人化的各種資訊蒐集的技術，將個人與機器互動中的信息保存並且傳遞給管理的中心，這樣一方面可以改善組織內部績效資訊的有效傳遞，讓個人的錯誤與不滿的意見可以上達，但是這樣的應用當然也會帶來不當監控的組織人本主義危機（von Schenk, 2021）。無論如何，AI 機制給組織代理人問題所帶來的新管理挑戰，當然可以藉由 AI 機制本身擅長放大與展示特殊資訊的能力來進行解決，不過，因爲這種高度監控的環境，也可能對組織產生信任與溝通效能的影響，需要更細緻的實驗或研究來發掘（Lazer, 2015）。

無論如何，前述組織人本主義與人工智慧發展所產生的宏觀層次的問題，一方面從價值錯置的概念來提醒從事技術研發的參與者，絕對不能忽視公共行政理論中的「公共價值」的管理與領導的問題（Andrews, 2019），另一方面，人工智慧的發展對於官僚組織當中的多重代理問題也會有持續惡化的可能，特別是互動關係中的信任（Lui & Lamb, 2018）的破壞與重建，都應該被關注；接下來，我們從組織人本主義的個人層面，來討論傳統韋伯官僚系統要面對的問題。

三、個人微觀層面：AI 帶來的 AI 裁量權與自動化偏誤的問題

在個人層次上，隨著政府組織更積極追求「官僚程序自動化技術」的趨勢，對個人層次的最直接影響將是對公務人員決策與行政裁量行爲的改變。對政府機關的公職人員而言，他們關心的或許不是組織的終極價值，也不見得常常反思自己在組織結構上的委託代理關係，但 AI 科技的應用，將對他們日常的工作常規與流程產生直接影響，也讓數位資訊系統帶出新的理論意涵。

臺灣大學學術期刊資料庫

行政資訊系統當然不是新玩意，過去早已出現爲人熟知的資訊決策支援系統（DSS）。DSS 的設計通常是由專家提供系統中所需的所有規則，藉由資訊工具使其能自動化專家的判斷，屬於被動工具，其實是複製傳統組織規則，工作業務上最後的決策依舊回到傳統組織所設定的結構與流程規範。但利用大數據與「機器學習」（machine learning）、「深度學習」（deep learning）或「類神經網絡系統」（artificial neural network）等演算法爲基礎的 AI 系統，則先透過訓練資料，產生輸入資料與輸出結果的模式，建構一個預測模型。實例如 Kleinberg 等利用紐約市過去的審前羈押資料，將資料訓練成一預測函數，再據此函數之分析結果給予法官，作爲是否需要羈押的風險評估建議（kleinberg et al., 2018）。因此機器學習演算並非只是複製專家決策所設定的判斷規則，而是根據不斷更新的歷史資料與學習演算的閥值調整來產生決策演算法。表 1 爲 Janssen 等學者對傳統組織規則以及機器學習規則的比較。此比較清楚呈現機器學習規則與傳統組織規則，在知識創造、資料輸入與執行、相關風險、以及解釋力上，皆有不同的邏輯（Janssen et al., 2022）。

表 1　傳統組織規則與機器學習規則的比較表

	機器學習規則	傳統組織規則
知識創造	從資料中找到規則並推論	根據邏輯與因果假定，由人類明確定義規則
輸入資料與執行	使用歷史資料來學習	形式化規定＋經驗法則＋個案
相關風險	歷史資料本身存在的偏誤 模型不擅長處理例外與異常值 正確率侷限	個人詮釋規定的差異 對特殊衝突狀況與複雜規定所產生的有限理性
解釋力	模糊的	宣告式邏輯

圖表來源：整理自 Janssen et al. (2022)

也因此，已有研究提出未來 AI 可被視爲獨立且不重複的個體，每個 AI 系統在學習後得出的結果可能不盡相同。除了「監督式學習」（supervised learning）的作法外，近來機器學習模型中「非監督式學習」的發展（unsupervised learning），不需人工標注或被定義，即可達到分群或尋找特定模式的目的，已被認爲能執行偵錯文字錯誤、詐騙、行爲交易異常等任務。非

臺灣大學學術期刊資料庫

監督式學習也呈現未來機器可能不需要向人類學習，具可獨立學習的機會（但不可諱言仍需由人類確認結果）。此發展方向正呼應了人工智慧工程師對發展「強 AI」的企圖，希望能開發出能更適應環境、辨識他人情緒的認知型機器（Jiménez-Picón et al., 2021）。如 Bostrom (2016) 認爲當 AI 技術發展突破與成熟後，將可達到與人類相當甚至超越人類的智慧能力。因而人工智慧可成爲官僚組織中一份子（代理載體）的想法，已非僅是未來學的想像，而是現在進行式。但由於技術仍存有不確定性與可能風險，在組織中與人類行爲決策的關係爲何，仍有不同見解。

長久以來，公共行政的核心課題之一，即是研究公務人員行政裁量的合理性與範疇，如 Lipsky (1980) 認爲政策執行評估應「由下而上」進行，以實際與民眾接觸的基層執行人員爲研究對象，如：警察、老師、社福人員、清潔隊員等，因爲他們才需要在充滿不確定性的「現場」處理公共事務，在第一線於極短的時間運用其裁量空間來解決民眾問題。2000 年以後資訊與通訊技術（ICTs）的發展則讓政府組織學者觀察到，ICTs 物理性地改變政府人員的決策裁量空間，從過去面對民眾的第一線基層官僚具備較高行政裁量空間，轉變爲面對螢幕與電腦系統，裁量空間縮小且更爲被動的系統官僚（Bovens & Zouridis, 2002）。

對比過去資訊系統的靜態設計與專家邏輯定義爲主，Young等學者認爲 AI 技術的運用使得新公共應用將根本性突破過去人類裁量的獨立性，因而提出「AI 裁量權」（artificial discretion）的新概念，定義爲機器學習技術面對「非決定性任務」（nondeterministic task）的理性決策行爲，[12] 這個概念也隱含公務人員核心業務內容與行政裁量權行使的移轉（Brehm & Gates, 1997；Bullock, 2019；Liou & Korosec, 2009；Lowi, 1993；Rohr, 1986；Wilson, 1991；Young et al., 2019）。這些裁量包含了最狹義的行政裁量，亦即「行政機關基於法律之明示授權或消極默許，本於行政目的，於數種可能之法律中，自行斟酌選擇一適當之行爲爲之」。與此直接相關的例子不少，如各

[12] 根據 Young 等學者指出，非決定性任務的非決定性意指，該任務的結果無法被既定「規則」（rules）完全的決定，可能參雜多重規則與價值，並可產生多個解方（Young et al., 2019）。

臺灣大學學術期刊資料庫

縣市警察局正積極部署的科技交通執法與環保署正在研擬的智慧判煙系統，由 AI 進行自動化的偵測與違規判斷，即是數位科技改變過去裁罰流程判定方法的實例。但本文希冀探討的不僅止於公法所規範的行政裁量範圍，希望公行學者能關心更廣義的官僚裁量，以及「在行政權限範圍內的有效決策空間」（Hill, 2013: 238）。因爲在 AI 發展多元與快速的推促下，未來應有更多政府業務是透過 AI 協助資料提供、資料分析與資料判斷，進而執行既定政策，AI 可謂是間接地對公務人員之行政裁量與業務判斷產生影響。如國內研究發現，機關用自然語言處理之 AI 演算法來判斷政治新聞之輿情立場與文字情緒，事實上容易分析錯誤，產生約 20% 的電腦判斷錯誤率（Hsiao et al., 2018: 219）。若政府單位倉促採用這樣的輿情分析於政策制定或民眾回應的判斷上，恐導致嚴重問題。

在上述廣泛定義的基礎上，歸納近來相關文獻發現，AI 裁量權與政府人員裁量權之取捨，應從以下三個面向比較，包括：任務規模化、任務的成本有效化以及任務的品質優化。如：AI 裁量權在處理大量且複雜因子的任務時雖具有優勢，但如此前端的系統可能極度昂貴，不見得是具有成本效益的。另一個與行政行爲科學有關之課題是，AI 裁量與政府人員行政裁量權將如何相互影響？當委託 AI 演算法進行先期裁量成爲「組織習慣」（organizational routine），行政人員的專業自主性是否將從主動轉爲被動？行政部門人員未來是否有實際進行選擇的眞實（Oswald et al., 2018）？這些問題的答案，將會決定未來組織這個以人爲本的傳統概念，是否會產生根本性變化的關鍵所在。

此外，隨著行爲取向的政策研究已逐漸被公共行政與政策學者關注（賴怡樺等，2018），Simon (1991；1997) 的「有限理性」（bounded rationality）與 Tversky & Kahneman (1974；1991) 的「展望理論」（prospect theory），常被用來理解爲何人的決策常是次佳或是錯誤的。如 Tversky & Kahneman 提出的「捷思」（heuristics）概念，即發現在有限的認知資源與條件下，人們往往不會進行充分分析，而傾向使用捷徑式的思考方式來做選擇。而研究個體層次的學者們更是開始借用「人機工程學」（Ergonomics）的理論，研究機器與使用者互動關係。這當中以 Parasuraman & Manzey (2010) 所提出的「自動化偏誤」（automation bias）理論架構，似乎最能回應近來文獻中

臺灣大學學術期刊資料庫

對於 AI 裁量與政府人員行政裁量權將如何相互影響的提問。藉由一系列訓練航空管制員的實驗發現，即便有人在旁做最後判斷與控制，長時間工作下來，自動化機器會使人的注意力下降，進而對決策產生非預期的後果。而「自動化系統的技術效度」（automation reliability）以及受測者對新科技的樂觀態度與自我效能感都會增加「自動化自滿」（automation complacency）的情況，意即受試者不會仔細檢查自動化系統的建議，會假定一切都沒問題（Bainbridge, 1983；Parasuraman et al., 2000；Parasuraman & Manzey, 2010）。此外，對於需要高度專注的任務來說，操作機器時接受資訊的每個環節都十分重要，因此當注意力不足時，也將讓操作者容易失去情境警覺力。倘若自動化系統表現正常，則不影響最終結果，但若是此時系統出錯，則操作者恐怕無法偵錯，進而導致錯誤。

方法上，從近年國外文獻的趨勢來看，此領域課題被認爲以行爲實驗方式來檢驗更爲適合。如 Janssen 等人透過實驗調查研究發現，比起個人單獨做判斷，有機器輔助能讓受試者作出更多的正確決策（Janssen et al., 2022）。對於有經驗的專家來，雖說較能夠偵測機器學習演算法所做的錯誤判斷，但也並非百分之百。相反的，Alon-Barkat & Busuioc (2023) 的實驗研究並未發現荷蘭民眾在處理高中老師聘任案的任務上，存在 AI 演算條件下的自動化偏誤問題。他們的實驗更近一步檢驗心理學的「選擇性依附」（selective adherence）理論，意即在資訊介入時若給予符合社會刻板印象的資訊，將會強化決策者的既定想法的假設。實驗結果發現，選擇性依附的現象確實存在，但資訊介入的效果源自於人類專家，還是 AI 演算法，則並無顯著差異。以臺灣樣本進行經驗研究，筆者分別於 2019 年與 2021 年執行了兩個實驗調查研究。實驗一針對廣義公務人員與私部門就職民眾進行實驗調查（曾冠球等，2020），受試者要完成 12 次的影像辨識任務，透過將受訪民眾分爲有 AI 介入輔助與控制組的方式，觀察到明顯的自動化偏誤，意即接收到 AI 提供輔助答案的成員，更容易改變自己原先的判斷，即使機器提供的資訊是錯誤的亦然。實驗二則以全國現役觀護人進行實驗調查（陳敦源等，2022），每位受試者需完成多次的假釋再犯風險評估，利用公部門實際的組織與案例來確認理論的效果，研究結果同樣發現有 AI 輔助決策介入的組別存在更顯著的自動化偏誤現象。

臺灣大學學術期刊資料庫

上述自動化偏誤現象的實證研究結果與信效度或許會被質疑或挑戰。以時至 2022 年 AI 在公部門的應用來看，國內公務機關尚未眞正出現如航空管制員或是飛行員使用的全自動化系統或 AI 系統，甚少有「完全部署」（fully deployed）的例子，但已有些機關正在進行試驗，如財務部運用人工智慧提升稅務服務（張文熙等，2021）。此外，雖非所有政府人員執行勤務都是長工時或多工進行，且操作員對系統失靈的警覺性也非政府文官的必要條件，然以台北市交通警察大隊於 2019 年，始推動交通違規之 AI 科技執法爲例，在 AI 系統判定違規行爲後，再由交通警察隊員警進行第二道人工辨識，確認違規吻合則寄出罰單。所呈現的執法情境，越來越類似 Bovens & Zouridis (2002) 提出的「系統官僚」情境。隨著數位科技的導入，第一線文官已從與民眾面對面的基層官僚，轉變爲僅與系統互動的系統官僚，公務人員在決策流程中，越來越依賴機器。可以想見，承辦人員承擔的風險強度與資訊處理的即時性或許仍在初期階段，但政府組織需擔心的應屬認知偏誤後果，是以行政人員未來能否判斷所接受的資訊爲重要訊息、錯誤內容、亦或是雜訊。政府業務眾多且繁雜，哪些因素可能導致政府行政人員的自動化偏誤，影響其專業判斷。是相當值得研究的課題，也將對未來組織在設計導入流程與執行細節上有重要參考意義。

上述討論顯示，已有愈多的理論文章與實證研究注意到個人層次的行政裁量影響問題；這也呼應了「組織人本主義」中的困境之一（Alvesson, 1982），亦即個人業務角色與技術理性甚或是經濟效率的衝突可能更爲加劇。 但另一方面，技術更成熟的 AI 系統能否成爲組織內的技術代理，協助人類公務員達成「韋伯式理想型官僚」的專業理性、依法行政、獨立於外部利益等特質，同樣令人感到期待與盼望（Bullock et al., 2022）。

肆、結論：組織根本性的 AI 智慧變革－舊瓶無法裝新酒？

公共行政的實務應用與學術研究面對 AI 的來勢洶洶，其趨勢已經無法抗拒。故本文最主要之目的在於提供國內公共行政社群的一次系統性的反思機會，特別是學術界，我們相信未來當參與者面對研究的探究、理論與

臺灣大學學術期刊資料庫

議題的選擇、以及研究方法與設計的問題時，本文是一個不可或缺的開始。本文倡議國內公行學界應深刻思考數位轉型趨勢下，大量開發與應用 AI 科技對政府機關所帶來的影響，或許已達孔恩（Thomas Kuhn）所言之科學典範轉移的關鍵時刻，許多嶄新之課題與現象非過去組織改造或績效管理能解釋與回答，也類似西方基督教文明喜用「舊皮袋裝新酒」的說法，來比喻新興的 AI 是否能夠如「科技接受理論」（Technology Acceptance Model, TAM）（Marangunić & Granić, 2015）所言，政府組織可以透過特定的管理與教育訓練活動，讓 AI 這個新科技順利納入當下的組織體系，類似新酒順利裝入舊有的酒器當中，而不會影響公共組織在效率、效能、與公平等價值要求下的運作。因此，在可見的未來當 AI 與機器學習演算機器進入公部門後，組織的改革該如何思考、規劃、執行、檢驗，應該會成爲公共行政領域未來十年發展上的重要研究與實務的議題。本文起始於應用「組織人本主義」與「韋伯官僚理論」來應對 AI 快速發展對於公共行政核心理論的影響，接著，本文從組織「宏觀－中觀－微觀」分析結構所進行的對話論述中，找出價值錯置、代理人問題、以及自動化偏誤等三個焦點，最終，作者將在前面的討論基礎之上，於此提出公部門相關理論與實務工作者，面對 AI 必然的衝擊時，可以深度思考的四點焦點，作爲結論。

首先，AI 採用與介入的「底線」劃定是重要的：從組織人本主義出發的討論，我們看見人類在 AI 發展的初期，對於即將發生的「失控」焦慮，因此，面對宏觀層次的集體焦慮感，最直接的處理方式就是以政府的公權力介入 AI 的發展與可應用之範圍，進行抑制其負面影響的超前部屬工作，類似基因改造或早期人類幹細胞技術的發展，一旦碰觸到人本主義所重視的人倫底線，也就是產生價值錯置的風險時，政府就該在事前或事後進行介入管制。比方說，在基因複製人的體內圈養器官給另一個人商業使用，會不會有貶低人類尊嚴的倫理疑慮就是一例，對比來看，AI 不但能夠在例行事務上取代大眾決策，還可能取代人類對於倫理價值「對與錯」選擇的權利。這樣看來，政府介入 AI 的發展，首要任務就是要能夠確保人類面對 AI 漸次進入組織的當口，仍然能夠保持最後的話語與行動能力；然而，面對組織生活因 AI 而來的鉅變，公部門介入去平衡失控焦慮的行動，還是必須調和變革的幅度與控制的底線，例如，政府可以建立一種「數位創新沙

臺灣大學學術期刊資料庫

盒實驗機制」（Digital Innovation Regulatory Sandbox）（陳敦源等，2021；Huang et al., 2021），在傳統的韋伯式官僚組織中，預設一個可控制的實驗環境，一種一方面能確保人類的倫理底線發言權，但同時也能獲取組織變革所能帶來效益的創新治理工具。

第二，人機（AI）協作的「途徑」選擇是重要的：目前以 AI 爲組織變革關鍵技術的政府數位轉型的議題，組織領導者應首重考量到這個過程應該是「漸近式」（incremental）或「革命性」（revolutionary）的問題；當然，本文一開始應用特斯拉總裁的反諷話語，認爲 AI 如果輾壓人類，根本不受「人類既有能力」（human compacity）的限制。這樣看來，由人類來掌控最大的問題還是要回到組織人本主義「組織變革以人之福利爲本」的核心，其中最重要的「辨識點」（demarcation point），就是問：「AI 導入組織，是否能夠協助人類在有限能力這個根本性前題上產生根本性的改變？」；換言之，前述漸進與革命性改變的選擇焦點，在於 AI 能否讓人類自我超越 Herbert H. Simon 所提出「有限理性」，甚至物理上的天然限制（一天只有 24 小時）。如果可以，漸近式最終將產生根本性的改變而讓人類獲益，不過，這個變革途徑的選擇權，由於無法想像人類能像上帝一樣「全知全能」（omnipotence），將無法得以實踐，因 AI 恐將成爲有限理性框架下的另一種資訊來源。另一個問題是，在韋伯型官僚組織中，AI 能否協助組織降低因人類能力限制而產生的「交易成本」（transaction costs），如搜尋、協商、監督等成本。如果可以，革命性的變革途徑才有落實的可能。以現今 AI 技術的能力，風險較低的漸進改革路徑，似乎更合適回應組織人本主義精神，以及韋伯官僚組織既有穩定功能不受破壞的基本社會需求。

第三，人機（AI）協作的「才能」管理是重要的：AI 能強化組織決策與服務能力的基礎，在於組織內眾多個人身上所擁有的「才能」（talent）是否足以應對且理解持續成長的 AI 能力。因此，不論在 Burke-Litwin 的三個層次架構中的哪一個部分，如果涉入 AI 的變革，總是會有組織內部人員 AI 應用管理的才能需求，這個才能的問題也是組織人本主義討論的核心之一。觀察目前政府於 AI 變革才能管理的作法，政府對於新才能的需求滿足，多從組織以外的外包模式取得，在決定組織疆界的「購買」或「自製」決策（buy or make decision）（Bartel et al., 2014）上以前者爲主，但長期來看

臺灣大學學術期刊資料庫

恐弱化組織內人員專業能力進而影響決策。因此，韋伯型官僚組織對於人才管理的回應，除了需建構與領域專家長期合作的模式外，政府也需盤點內部相關領域人才知識的範圍與稀缺的狀態；更甚者，傳統資訊人才單純扮演政府組織中幕僚角色的傳統，也因我國政府於 2022 年成立數位發展部之後，而產生角色的質變，自此，資通科技的知識領域，不再是 Henry Mintzberg (1980) 組織設計「五大」結構中「技術幕僚」的角色而已，也會更多扮演業務部門的角色（蕭乃沂等，2022-2023），這個轉變剛好是 Burke-Litwin 組織中觀層次關於設計與流程等問題的核心議題，當然，不論從組織人本主義或韋伯官僚體制的理論來看，這問題必然成爲組織內部才能管理創新的重要環節，在未來，不論是組織中觀或是個人微觀的層面，資通科技的應用必然成爲業務單位所需才能的一部份，雖然公務人員的工作不會被 AI 所立即取代，但是公務組織或公務人員都必須設法導入與融合 AI 的相關應用知能，以期能強化政府解決民眾公共問題的能量與有效程度。

第四，人機（AI）協作的「分工」設計是重要的：從組織人本主義的角度來看，本文 AI 科技的發展，似乎隱含悲觀的想像，然而，從社會科技轉型理論（Geels, 2011）過去對科技與社會的交互發展脈絡來推測，如果先將人類被 AI 取代放在「可能但不會是當下」（possible but not now）的未來管理的意識上，AI 時代政府組織以人爲本的中心性不會立即改變的前提下，政府韋伯官僚組織設計的目標，仍然是以組織經濟學所著重「協調」與「溝通」（coordination and communication）（Milgrom & Robert, 1992）爲核心，因此，本文前面的討論中，韋伯官僚體制面對 AI 的快速發展將面臨到三個重新設計的議題：其一，組織控制機制的設計，因爲 AI 的介入，對於個人的行爲的影響，也就是前文提到組織成員與 AI 機制協調下所可能產生「自動化偏誤」的問題，需要積極重新設計處理之；其二，對於個人的工作場域的未來變化，AI 可以協助員工以更有效率的工作模式處理「多管道的服務提供」（multi-channel services delivery），並且不一定需要在特定的辦公場所工作，這些改變需要政府組織部門，在法規與管理層次，進行基層行政規則與溝通管理模式的變革；其三，目前政府改革承襲「新公共管理」（NPM）的四化策略：地方化、去任務化、資訊化、委外化，除了資訊化之外，都已經碰到瓶頸，因此，導入 AI 重新設計任務與組織分工，是推

臺灣大學學術期刊資料庫

動政府組改的唯一有未來性的手段。

最後，本文希望藉由站在「組織人本主義」與「韋伯官僚體制」的雙元理論基礎上，與快速發展的 AI 科技進行深度對話，以期能夠提供政府管理社群中學術界與實務界的成員，一個面對 AI 科技衝擊所引發政府數位轉型的狂潮中，能夠停下腳步進行整體反思的機會，並且以 Burke-Litwin 的三層次理論爲討論框架，分別從不同層次中的關鍵概念來指引這樣的反思。當然，作者清楚這樣的反思面對這個驚天動地的變革壓力，其理論與實務的密度都不足以創造出足夠且合理的變革知識內涵與動能，因此，後續針對本文所提出的多元理論建構與實務驗證的發展都是更加迫切需要的，也期待學界與實務界能夠有更多的研究與實驗的資源，投入攸關未來的知識生產，以期在這個沒有選擇的變革壓力下，以人爲本的選擇能力以及傳統韋伯官僚體制穩定現況的功能，還有持續存在的可能。

參考書目

一、中文部分

丁玉珍、林子倫，2020，〈人工智慧提升政府公共治理的應用潛力探討〉，《檔案半年刊》，19(2): 24-41。Ting, Yu-jen & Tze-luen Lin. 2020. “Rengong zhihui tisheng zhengfu gonggong zhili de yingyong quanli tantao” [The Exploration of the Potential on Artificial Intelligence Applications to Improve Public Governance]. *Archives Semiannual* 19(2): 24-41.

李翠萍、張竹宣、李晨綾，2022，〈人工智慧在公共政策領域應用的非意圖歧視：系統性文獻綜述〉，《公共行政學報》，63: 1-49。Lee, Tsuey-ping, Chu-yi Chang, & Chen-ling Lee. 2022. “Rengong zhihui zai gonggong zhengce lingyu yingyong de feiyitu qishi: xitongxing wenxian zongshu” [Unintentional Discrimination in Application of Artificial Intelligence to Public Policies: A Systematic Article Review]. *Journal of Public Administration* 63:1-49.

張文熙、王培得、陳秀宇、林哲豪、鄧宜芳、曾伯勳、平曉鷗，2021，〈運用人工智慧提升稅務服務之探討〉，《財稅研究》，50(5): 93-115。Zhang, Wen-Xi, Pei-de Wang, Xiu-yu Chen, Zhe-hao Lin, Yi-fang Deng, Bo-xun Zeng, & Xiao-ou Ping. 2021. “Yunyong rengong zhihui tisheng shuiwu fuwu zhi tantao” [The Exploration of Using Artificial Intelligence to Improve Tax Services]. *Public Financial Review* 50(5): 93-115.

臺灣大學學術期刊資料庫

陳敦源、廖洲棚、黃心怡、張濱璿，2021，《公共服務數位沙盒實驗機制預評估》，臺北：數位治理研究中心年度研究報告。Chen, Don-yun, Zhou-peng Liao, Hsini Huang, & Pin-hsuan Chang. 2021. "*Gonggong fuwu shuwei shahe shiyan jizhi yu pinggu*" [T*he Pre-evaluation of Public Service Sandbox Experimental Scheme*]. Taipei: Taiwan E-Governance Research Center Annual Research Report.

陳敦源、廖洲棚、黃心怡、張濱璿、王千文，2022，《公部門機器演算法應用之制度調適與路徑分析》，臺北：數位治理研究中心年度研究報告。Chen, Don-yun, Zhou-peng Liao, Hsini Huang, Pin-hsuan Chang, & Chian-wen Wang. 2022. "*Gongbumen jiqi yansuanfay ingyong zhi zhidu tiaoshi yu lujing fenxi*" [*The Institutional Adaptation and Path Analysis of Algorithmic Governance in Public Service*]. Taipei: Taiwan E-Governance Research Center Annual Research Report.

曾冠球、廖洲棚、黃心怡，2020，《新興科技與公部門數位轉型:個案採擷與模式建構》，臺北：數位治理研究中心年度研究報告。Tseng, Kuan-chiu, Zhou-peng Liao, & Hsini Huang. 2020. "*Xinxing keji yu gong bumen shuwei zhuanxing gean caijie yu moshi jian hou*" [*An Analysis of Evidence-based E-Governance and E-Communication Strategies*]. Taipei: Taiwan E-Governance Research Center Annual Research Report.

黃心怡、曾冠球、廖洲棚、陳敦源，2021，〈當人工智慧進入政府：公共行政理論對 AI 運用的反思〉，《文官制度》，13(2): 91-114。Huang, Hsini, Kuan-chiu Tseng, Zhou-peng Liao, & Don-yun Chen. 2021. "Dang rengong zhihui jinru zhengfu: gonggong xingheng lilun dui AI yunyong de fansi" [When AI Joins the Government: A Reflection on AI Application and Public Administration Theory]. *Journal of Civil Service* 13(2): 91-114.

黃東益，2017，〈資訊通訊科技驅動治理轉型？趨勢與研究議題〉，《文官制度》，9(3): 1-25。Huang, Tong-yi. 2017. "Zixun tongxun keji qudong zhili zhua xing? qushi yu yanjiu yiti" [Do ICT Drive Governance Transformation? Trends and Research Issues]. *Journal of Civil Service* 9(3): 1-25.

蕭乃沂、林俞君、廖興中，2022-2023，《政府機關數位發展人才職能建構與訓用連結之研析》，臺北：數位治理研究中心年度研究報告（未公開期中報告）。Hsiao, Nai-yi, Yu-chun Lin, & Hsin-chung Liao. 2022-2023. "*Zhengfu jiguan shuwei fazhan rencai zhineng jia gou yu xunyong lianjie zhi yanxi*" [*A Study of the Linkage between the Competency Construction and Training of*

臺灣大學學術期刊資料庫

Digital Development Talents in Government Agencies] Taipei: Taiwan E-Governance Research Center Annual Research Report (unpublished interim report).

賴怡樺、林水波、陳敦源，2018，〈行爲主義導向的公共政策研究：以政策工具「推力」爲焦點之初探〉，《行政暨政策學報》，67: 1-37。Lai, Yi-hua, Shoei-po Lin, & Don-yun Chen. 2018. "Xingwei zhuyi daoxiang de gonggong zhengce yanjiu: yi zhengce gongju "tuili" wei jiaodian zhi chutan" [The Direction of Behavioralism in Public Policy Research: A Review on "Nudge" as a Policy Tool]. *Public Administration & Policy* 67: 1-37.

二、西文部分

Alderfer, C. 1994. "A Response to 'a Feminist Critique of Organizational Humanism'." *The Journal of Applied Behavioral Science* 30(1): 98-102.

Alon-Barkat, S. & M. Busuioc. 2023. "Human-AI Interactions in Public Sector Decision-Making: 'Automation Bias' and 'Selective Adherence' to Algorithmic Advice." *Journal of Public Administration Research and Theory* 33(1): 153-169.

Alvesson, M. 1982. "The Limits and Shortcomings of Humanistic Organization Theory." *Acta Sociologica* 25(2): 117-131.

Andrews, L. 2019. "Public Administration, Public Leadership and the Construction of Public Value in the Age of the Algorithm and 'Big Data'." *Public Administration* 97(2): 296-310.

Arrow, K. 1985. "The Economics of Agency." In *Principals and Agents: The Structure of Business*, eds. John W. Pratt & Richard J. Zeckhauser. Boston: Harvard Business School Press, 37.

Bainbridge, L. 1983. "Ironies of Automation." *Automatica* 19(6): 775-779.

Bannister, F. & R. Connolly. 2020. "Administration by Algorithm: A Risk Management Framework." *Information Polity* 25(4): 471-490.

Bartel, A. P., S. Lach, & N. Sicherman. 2014. "Technological Change and the Make-or-Buy Decision." *The Journal of Law, Economics, and Organization* 30(1): 165-192.

Barth, T. J. & E. Arnold. 1999. "Artificial Intelligence and Administrative Discretion: Implications for Public Administration." *The American Review of Public Administration* 29(4): 332-351.

Baum, J. A. & H. A. Haveman. 2020. "Editors' Comments: The Future of Organizational Theory." *Academy of Management Review* 45(2): 268-272.

臺灣大學學術期刊資料庫

Beverungen, A. & A. C. Lange. 2018. “Cognition in High-Frequency Trading: The Costs of Consciousness and the Limits of Automation.” *Theory, Culture & Society* 35(6)*:* 75-95.

Bostrom, N. 2016. *Superintelligence: Paths, Dangers, Strategies.* Reprint ed. New York: Oxford University Press.

Bovens, M. & S. Zouridis. 2002. “From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control.” *Public Administration Review* 62(2): 174-184.

Brehm, J. & S. Gates. 1997. *Working, Shirking, and Sabotage: Bureaucratic Response to a Democratic Public*. Ann Arbor: The University of Michigan Press.

Buffat, A. 2015. “Street-Level Bureaucracy and E-Government.” *Public Management Review* 17(1): 149-161.

Bullock, J. B. 2019. “Artificial Intelligence, Discretion, and Bureaucracy.” *The American Review of Public Administration* 49(7): 751-761.

Bullock, J. B., H. Huang, & K. C. Kim. 2022. “Machine Intelligence, Bureaucracy, and Human Control.” *Perspectives on Public Management and Governance* 5(2): 187-196.

Burke, W. W. & G. H. Litwin. 1992. “A Causal Model of Organizational Performance and Change.” *Journal of Management* 18(3): 523-545.

Buttazzo, G. 2001. “Artificial Consciousness: Utopia or Real Possibility?” *Computer* 34(7): 24-30.

Byrkjeflot, H., P. du Gay, & C. Greve. 2018. “What is the ‘Neo-Weberian State’ as a Regime of Public Administration?” In *The Palgrave Handbook of Public Administration and Management in Europe*, eds. Edoardo Ongaro & Sandra Van Thiel. London: Palgrave Macmillan, 991-1009.

Carrera, Á. & C. A. Iglesias. 2015. “A Systematic Review of Argumentation Techniques for Multi-agent Systems Research.” *Artificial Intelligence Review* 44(4): 509-535.

Coglianese, C. & D. Lehr. 2017. “Regulating by Robot: Administrative Decision Making in the Machine-Learning Era.” in https://scholarship.law.upenn.edu/faculty_scholarship/1734. Latest update 15 April 2022.

Danaher, J., M. J. Hogan, C. Noone, R. Kennedy, A. Behan, A. de Paor, H. Felzmann, M. Haklay, S. M. Khoo, J. Morison, M. H. Murphy, N. O’Brolchain, B. Schafer, & K. Shankar. 2017. “Algorithmic Governance: Developing a Research

臺灣大學學術期刊資料庫

Agenda through the Power of Collective Intelligence." *Big Data & Society* 4(2): 1-21.

Danziger, J. N. & K. V. Andersen. 2002. "The Impacts of Information Technology on Public Administration: An Analysis of Empirical Research from the "Golden Age" of Transformation." *International Journal of Public Administration* 25(5): 591-627.

Daylamani-Zad, D., H. Agius, & M. C. Angelides. 2020. "Reflective Agents for Personalisation in Collaborative Games." *Artificial Intelligence Review* 53(1): 429-474.

Denhardt, R. B. & T. J. Catlaw. 2014. *Theories of Public Organization*. 7th ed. Boston: Cengage Learning.

Dreyfus, H. L. 1972. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge: The MIT Reress.

Dunleavy, P., H. Margetts, S. Bastow, & J. Tinkler. 2006. "New Public Management Is Dead—Long Live Digital-Era Governance." *Journal of Public Administration Research and Theory* 16(3): 467-494.

Fountain, J. E. 2022. "The Moon, the Ghetto and Artificial Intelligence: Reducing Systemic Racism in Computational Algorithms." *Government Information Quarterly* 39(2): 101645.

Geels, F. W. 2011. "The Multi-level Perspective on Sustainability Transitions: Responses to Seven Criticisms." *Environmental Innovation and Societal Transitions* 1(1): 24-40.

Gritsenko, D. & M. Wood. 2022. "Algorithmic Governance: A Modes of Governance Approach." *Regulation & Governance* 16(1): 45-62.

Hage, J. 1980. *Theories of Organizations: Form, Process, and Transformation*. New York: John Wiley & Sons.

Hill, M. 2013. *The Public Policy Process*. 6th ed. Harlow: Pearson.

Hsiao, N., Z. P. Liao, & P. Y. Chen. 2018. "From Naive Expectation to Realistic Progress: Government Applications of Big Data on Public Opinion Mining." In *Big Data in Computational Socieal Sciences and Humanities: An Introduction*, eds. Shu-heng Chen & Tina Yu. Switzerland AG: Springer International Publishing, 207-220.

Huang, H., P. H. Chang, Z. P. Liao, & D. Y. Chen. 2021. "A Matter of Risk Management: The Effects of the innovation Sandboxes on Citizens' Risk Perceptions." In *DG.O2021: The 22nd Annual International Conference on Digital Government Research,* ed. Association for Computing Machinery.

臺灣大學學術期刊資料庫

New York: Association for Computing Machinery, 281-285.

Hurbans, R. 2020. *Grokking Artificial Intelligence Algorithms: Understand and Apply the Core Algorithms of Deep Learning and Artificial Intelligence in this Friendly Illustrated Guide including Exercises and Examples*. 1st ed. Shelter Island: Manning Publication.

Janssen, M., M. Hartog, R. Matheus, A. Y. Ding, & G. Kuk. 2022. "Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-supported Decision-Making in Government." *Social Science Computer Review* 40(2): 478-493.

Jiménez-Picón, N., M. Romero-Martín, J. A. Ponce-Blandón, & L. Ramirez-Baena. 2021. "The Relationship between Mindfulness and Emotional Intelligence as a Protective Factor for Healthcare Professionals: Systematic Review." *International Journal of Environmental Research and Public Health* 18(10): 5491.

Jorna, F. & P. Wagenaar. 2007. "The 'Iron Cage' Strengthened? Discretion and Digital Discipline." *Public Administration* 85(1): 189-214.

Katzenbach, C. & L. Ulbricht. 2019. "Algorithmic Governance." in https://policyreview.info/concepts/algorithmic-governance. Latest update 10 May 2020.

Kleinberg, J., H. Lakkaraju, J. Leskovec, J. Ludwig, & S. Mullainathan. 2018. "Human Decisions and Machine Predictions." *The Quarterly Journal of Economics* 133(1): 237-293.

Korinek, A. & J. E. Stiglitz. 2017. "Artificial Intelligence and Its Implications for Income Distribution and Unemployment." In *The Economics of Artificial Intelligence: An Agenda,* eds. Ajay Agrawal, Joshua Gons, & Avi Goldfarb. Chicago & London: The University of Chicago Press, 349-390.

Kurzweil, R. 2006. "Reinventing Humanity: the Future of Machine-human Intelligence." *The Futurist* 40(2): 39.

Lanzara, G. F. 2009. "Building Digital Institutions: ICT and the Rise of Assemblages in Government." In *ICT and Innovation in the Public Sector*, eds. Francesco Contini & Giovan Francesco Lanzara. London: Plgrave MaCmillan, 9-48.

Larkin, T. J. 1986. "Humanistic Principles for Organizational Management." *Communication Studies* 37(1): 36-44.

Lazer, D. 2015. "The Rise of the Social Algorithm." *Science* 348(6239): 1090-1091.

Levine, C. H., R. W. Backoff, A. R. Cahoon, & W. J. Siffin. 1975. "Organizational Design: A Post Minnowbrook Perspective for the "New" Public Administration." *Public Administration Review* 35(4): 425-435.

Liou, K. T. & R. Korosec. 2009. "Implementing Organizational Reform Strategies

臺灣大學學術期刊資料庫

in State Government." *Public Administration Quarterly 33(*3): 429-452.

Lipsky, M. 1980. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Service*. 30th Ann. Expanded ed. New York: Russell Sage Foundation

Lowi, T. J. 1993. "Legitimizing Public Administration: A Disturbed Dissent." *Public Administration Review 53(*3): 261-264.

Lui, A. & G. W. Lamb. 2018. "Artificial Intelligence and Augmented Intelligence Collaboration: Regaining Trust and Confidence in the Financial Sector." *Information & Communications Technology Law* 27(3): 267-283.

Makridakis, S. 2017. "The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms." *Futures* 90: 46-60.

Malomo, F. & V. Sena. 2017. "Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector." *Policy & Internet* 9(1): 7-27.

Marangunić, N. & A. Granić. 2015. "Technology Acceptance Model: A Literature Review from 1986 to 2013." *Universal Access in the Information Society* 14(1): 81-95.

Margetts, H. & P. Dunleavy. 2013. "The Second Wave of Digital-Era Governance: A Quasi-Paradigm for Government on the Web." *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences* 371(1987): 20120382.

Melé, D. 2003. "The Challenge of Humanistic Management." *Journal of Business Ethics* 44(1): 77-88.

Milgrom, P. & J. Robert. 1992. *Economics, Organization and Management*. Englewood Cliffs: Prentice Hall.

Mintzberg, H. 1980. "Structure in 5's: A Synthesis of the Research on Organization Design." *Management Science* 26(3): 322-341.

Mirbabaie, M., F. Brünker, N. R. Möllmann, & S. Stieglitz. 2022. "The Rise of Artificial Intelligence-understanding the AI Identity Threat at the Workplace." *Electronic Markets* 32: 73-99.

Moloi, T. & T. Marwala. 2020. *The Agency Theory. Artificial Intelligence in Economics and Finance Theories*. Unam: Springer.

Olsen, J. P. 2006. "Maybe It Is Time to Rediscover Bureaucracy." *Journal of Public Administration Research and Theory* 16(1): 1-24.

Oswald, M., J. Grace, S. Urwin, & G. C. Barnes. 2018. "Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and 'Experimental'

臺灣大學學術期刊資料庫

proportionality.” *Information & Communications Technology Law* 27(2): 223-250.

Parasuraman, R. & D. H. Manzey. 2010. “Complacency and Bias in Human Use of Automation: An Attentional Integration.” *Human Factors* 52(3): 381-410.

Parasuraman, R., T. B. Sheridan, & C. D. Wickens. 2000. “A Model for Types and Levels of Human Interaction with Automation.” *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 30(3): 286-297.

Rainey, H. G., S. Fernandez, & D. Malatesta. 2021. *Understanding and Managing Public Organizations*. 6th ed. Hoboken: Wiley.

Ramsey, V. J. & L. M. Calvert. 1994. “A Feminist Critique of Organizational Humanism.” *The Journal of Applied Behavioral Science* 30(1): 83-97.

Rohr, J. A. 1986. *To Run a Constitution: The Legitimacy of the Administrative State*. Lawrence: University Press of Kansas.

Rouvroy, A., T. Berns, & E. Libbrecht. 2013. “Algorithmic Governmentality and Prospects of Emancipation.” *Reseaux* 177(1): 163-196.

Russell, S., D. Dewey, & M. Tegmark. 2015. “Research Priorities for Robust and Beneficial Artificial Intelligence.” *Ai Magazine* 36(4): 105-114.

Shestakofsky, B. 2017. “Working Algorithms: Software Automation and the Future of Work.” *Work and Occupations* 44(4): 376-423.

Simon, H. A. 1991. “Bounded Rationality and Organizational Learning.” *Organization Science* 2(1): 125-134.

Simon, H. A. 1997. *Administrative Behavior.* 4th ed. Florence: Free Press.

Spender, J. C. 2013. “Herbert Alexander Simon: Philosopher of the Organizational Life-world.” In *The Oxford Handbook of Management Theorists,* eds. Morgen Witzel & Malcolm Warner. Oxford University Press, 297-358.

Stivers, C. & L. DeHart-Davis, 2022. “Introduction to the Symposium Issue on Reappraising Bureaucracy in the 21st Century.” *Perspectives on Public Management and Governance* 5(2): 77-83.

Sun, W., O. Nasraoui, & P. Shafto. 2020. “Evolution and Impact of Bias in Human and Machine Learning Algorithm Interaction.” *PLOS ONE* 15(8): e0235502.

Tversky, A. & D. Kahneman. 1974. “Judgment under Uncertainty: Heuristics and Biases.” *Science* 185(4157): 1124-1131.

Tversky, A. & D. Kahneman. 1991. “Loss Aversion in Riskless Choice: A Reference-Dependent Model.” *The Quarterly Journal of Economics* 106(4): 1039-1061.

van Rijmenam, M. & D. Logue. 2021. “Revising the ‘Science of the Organisation’:

臺灣大學學術期刊資料庫

Theorising AI Agency and Actorhood." *Innovation* 23(1): 127-144.

Vogl, T. M., C. Seidelin, B. Ganesh, & J. Bright. 2020. "Smart Technology and the Emergence of Algorithmic Bureaucracy: Artificial Intelligence in UK Local Authorities." *Public Administration Review* 80(6): 946-961.

Vogt, J. 2021. "Where is the Human Got to Go? Artificial Intelligence, Machine Learning, Big Data, Digitalisation, and Human-Robot Interaction in Industry 4.0 and 5.0: Review Comment on: Bauer M. (2020). Preise kalkulieren mit KI-gestützter Onlineplattform BAM GmbH, Weiden, Bavaria, Germany." *AI & Society* 36(3): 1083-1087.

von Schenk, A. 2021. "Predictably Satisfied: Contributions of Artificial Intelligence to Intra-Organizational Communication." in https://ssrn.com/abstract=3856479. Latest update 20 March 2022.

Waterman, R. W. & K. J. Meier. 1998. "Principal-Agent Models: An Expansion?" *Journal of Public Administration Research and Theory* 8(2): 173-202.

Weber, M. 1922. "Bureaucracy." In *Classics of Public Administration*. 7th ed. eds, Jay M. Shafritz & Albert C. Hyde. Wadsworth: Cengage Learning, 63-67.

Wilson, C. & H. Broomfield. 2022. "Learning How to Do AI: Managing Organizational Boundaries in an Intergovernmental Learning Forum." *Public Management Review* 2002(4): 1-20.

Wilson, J. Q. 1991. *Bureaucracy: What Government Agencies Do And Why They Do It.* 1st ed. Mew York: Basic Books.

Young, M. M., J. B. Bullock, & J. D. Lecy. 2019. "Artificial Discretion as a Tool of Governance: A Framework for Understanding the Impact of Artificial Intelligence on Public Administration." *Perspectives on Public Management and Governance* 2(4): 301-313.

Zuiderwijk, A., Y. C. Chen, & F. Salem. 2021. "Implications of the Use of Artificial Intelligence in Public Governance: A Systematic Literature Review and a Research Agenda." *Government Information Quarterly* 38(3): 101577.

臺灣大學學術期刊資料庫

Public Administration Research on Human and AI Collaboration: A Multilevel Reflection on Public Sector Organizational Issues*

*Hsini Huang** and Don-yun Chen****

Abstract

The growth and breakthroughs in innovative artificial intelligence (AI) technologies have provided many governments and public organizations with the possibility of achieving digital transformation. Yet, apart from the proposals of techno-optimists, as more problems emerge in implementing AI in society, politics, and business, there are calls for scholarly discussions to revisit how AI might reshape and influence government values, bureaucracy, and bureaucratic operations. This paper draws from both the Weberian ideal bureaucracy theory and the organizational humanism perspective as an upper-level theoretical framework to explore the impact of AI on the roles, control, and communication issues in public organizations. As a stepping stone towards AI research in government, this study aims to analyze public administration theory and the applications of AI in public sector practices. Three key issues are discussed from a multi-level analytical framework, including: 1) the problem of public value misalignment, 2) the problem of multiple agents within an organization, and 3) the issue of AI discretion and automation bias. This article concludes that when facing the inevitable AI revolution, public organizations should rethink the foundations of innovative change, path selection, talent management, and the "re-design" of divisions of labor.

Keywords: Public Administration, Organizational Humanism, Artificial Intelligence (AI), Value Misalignment, Multiple Agents, Automation Bias

---

* DOI:10.6166/TJPS.202306_(96).0004

** Assistant Professor, The Institute of Public Administration, FGGA, Leiden University (corresponding authon).
E-mail: h.i.huang@fgga.leidenuniv.nl

*** Professor, Department of Public Administration, National Chengchi University.
E-mail: donc@nccu.edu.tw

臺灣大學學術期刊資料庫