

法律人工智慧的法哲學反省： 判斷權限、執法機制與法治理念*

陳弘儒

中央研究院歐美研究所

E-mail: marxchen@gate.sinica.edu.tw

摘要

法律人工智慧系統 (ALI) 對於法治價值的影響與關係尚無深入探討。本文主張 ALI 與法治具有內在的緊張關係。第一，法治理念中的「主體」是一種主動且負責任的主體概念觀，而日趨完善的 ALI 將凸顯出權限分配移轉的難題。第二，高效率執法機制的建置是 ALI 的核心特色，旨在以執法效率提升法律之應然與實然面向的擬合，然其卻可能損害個人自主的道德判斷空間與干擾行動者行動理據的排序。據此，本文從法哲學角度說明

© 中央研究院歐美研究所

投稿日期：110.3.17；接受刊登日期：110.9.23；最後修訂日期：110.7.1

責任校對：張文綺、王崇懿、趙麗婷

* 本文基本構想發表於中央研究院法律學研究所舉辦的「人工智慧與法律規範工作坊」(2019年6月14日)，以及2019 International Conference on Technologies and Applications of Artificial Intelligence (2019年11月22-23日)之中。爾後短文出版在《人工智慧與法律衝擊》(張麗卿主編)一書。最後，文章實質改寫並發表於中央研究院歐美研究所主辦的「AI與民主」研討會(2020年6月3-4日)。文內的核心觀點自2019年開始筆者已進行一系列演講與討論，受益許多。謝謝諸多學界先進的寶貴意見。此外，特別誠摯感謝兩位審查人的細緻與寶貴意見，讓筆者有機會再次調整本文結構以及基礎論證。最後，謝謝助理趙麗婷女士給予寶貴的行政協助與翻譯意見。

ALI 與不同法治觀的可能緊張關係、透過因果圖呈現 ALI 對於守法行動的可能介入方式，進而提出兩個基本規範性原則：第一，替代型智慧系統之禁止原則。第二，高效率執法機制禁止原則。

關鍵詞：法律人工智慧、法治、實踐差異命題、判斷權限移轉、電腦功能主義

壹、判斷權限移轉的難題

一、問題在哪裡？

自動駕駛車輛 (self-driving cars; 自駕車) 是當前人工智慧技術應用的實現場域。李開復就指出，「我們也許還無法準確預測，全功能、最高等級的自動駕駛汽車，會在什麼樣的時間點，真正走入普通人的生活。但毫無疑問的是，在這次人工智慧的熱潮中，自動駕駛一定是最大的應用場景」(李開復、王詠剛，2017: 9)。從 SAE International J3016 (2014) 標準的區分可以見到人工智慧系統在設計上以功能履行為主要設計理念的框架，SAE International J3016 的技術分級意味著自動駕駛技術的等級區分是以機器是否可以妥適履行人類在動態駕駛任務中所使用的各式功能，例如油門、方向之控制、感測外界環境以及駕駛環境出現緊急狀況時的反應機制等等。

自動駕駛系統的例子具體表現了人工智慧在社會生活的運作特色：取代 (或協助) 原本由人為執行主體的功能運作。而在法律中，人工智慧的運用也屢見不鮮。例如，2018 年初，五位商務律師與法律人工智慧系統 (Artificially legal intelligent; ALI)，¹ 針對五份保密條款中 30 個爭點，進行辨識爭點的競賽。結果是這五位律師平均花了 92 分鐘閱讀完契約文件，爭點辨識比率高達 85%，而

¹ Artificial legal intelligence 一詞是來自於 Mireille Hildebrandt 的文章。她在〈演算法管制與法治〉(“Algorithmic Regulation and the Rule of Law”) 一文中首度提出這個概念 (Hildebrandt, 2018)。Hildebrandt 將 ALI 區分為兩類：第一類是「若此則那」(if this then that; IFTTT)。第二類是資料驅動為主 (data-driven) 的人工智慧系統。筆者雖然對於 Hildebrandt 的區分沒有完全贊同，但是本文的 ALI 可以包含這兩類法律人工智慧系統。作為一篇以法哲學為基礎的文章，我希望指出 ALI 的一般性問題。關於 Hildebrandt 對於 ALI 的區分，請參考本文肆、一。

法律人工智慧系統僅花了 26 秒就看完所有文件，爭點辨識率是 94%。雖然五位律師內最好的爭點辨識率也是 94%，但我們不知道該位律師是不是花了最長的閱讀文件的時間（156 分鐘）完成此項任務（LawGeex, 2018）。此外，臺灣交通部鐵路局預計在各個車站設計智慧人流監控系統——運用人臉辨識及人流移動提早發現可能意外狀況（交通部鐵道局，2019）；北臺灣的萬里隧道所創設的區間測速系統，計算車輛在兩點間之平均行駛距離，而推斷是否有超過該路段之速度上限；² 法律扶助基金會、東吳大學與 SAS 公司曾一起合作開發法扶案件分流系統；³ 清華大學的王道維、林昀嫻所合作開發的親權判決預測系統（林昀嫻、王道維，2020）。⁴ 甚至行政執行署所推動的交通正義專案，希望結合車牌辨識扣押欠稅車輛等等。⁵

上述的發展展示出一個關鍵特色：**判斷權限移轉是可能的**。更精確的說，人工智慧系統的研發就是要將原本屬於人的判斷權限移轉到人造物身上。即使人工智慧系統是否具有判斷能力仍是一個重要問題，但是無疑地它們在許多事務上已經在幫人類做判斷了。本文的基本問題便是要探尋，在法治的脈絡下我們可以如何思考法律人工智慧系統基本議題？

法學可以有以下方式看待 ALI 的相關議題：（一）探討人工智慧的應用所產生的實證法問題。例如自駕車如果造成死傷會是由誰需要負責？是系統設計者？車商？駕駛人或是由保險基金處理。由於

² 請參考新北市政府警察局交通大隊（2021）。

³ 請參考東吳大學（2018）。根據筆者後續與法扶基金會聯繫，此一合作並未持續下去。

⁴ 關於其人工智慧系統之應用網頁，請參考國立清華大學（n.d.）。

⁵ 請參考法務部行政執行署桃園分署（2020）。筆者曾與桃園與彰化承辦人員聯繫過，該兩區域之執行方式不一樣，唯基本上都是使用車牌影像辨識系統完成執法任務。

法律體系是一個分配權利與義務的規範體系，因此當任何人工智慧系統之應用對於社群成員產生傷害時，法學必須針對具體事件進行處理以及回應。這種處理是一種事後觀點 (ex post)，主要目的在於確定具體個案中的權益關係等課題。這會屬於法律裁判與法律釋義學的核心課題 (吳從周，2019; 李榮耕，2019; Farnsworth, 2007)。(二) 則偏向採取事前觀點 (ex ante)，思考要如何在人工智慧系統之應用產生問題之前管制 (或是設立規範) 以妥適分配權利與義務的關係，這種以事前觀點為切入取徑的方式通常是以管制理論的角度出發。⁶ 其主要思索 AI 對於法律既有的價值保障是否會產生扞格之處。例如以資料驅動 (data-driven) 為主的人工智慧系統所產生的偏見、歧視、隱私權等爭議。(三) 從更後設立場出發，不僅思考法律既有價值是否會妥協於人工智慧系統之應用，更思考這背後的可能性是如何產生的。本文是採取第三種研究取徑來看待 ALI 的課題。此外，本文在研究取徑上的基本預設是：ALI 雖然是一種技術，但是技術之設計與研發往往會具有價值蘊含在其中，⁷ 因此必須思考既有的運作體系的價值內涵產生變遷之可能。

奠基在第三種的研究取徑上，筆者探討，ALI 與法治的緊張關係是如何產生的。我將指出，ALI 與法治的內在緊張關係來自於 ALI

⁶ 例如，美國的 Executive Office of the President (2016) 提及自動駕駛車輛或航空器的技術發展與其進行政策管制的必要性。林勤富與李怡俐 (2020) 則奠基於對人工智慧對於社會、經濟與政治層面之影響，從國際人權機制思考管制人工智慧的可能方針。

⁷ 技術設計者對於技術與設計所摻雜的價值或偏見，常見的例子大概是美國紐約要通往瓊斯海灘 (the Jones Beach) 的橋樑設計。城市規畫者 Robert Moses 責成設計師將橋樑高度降低以不利雙層巴士的通行，檢少或避免有色人種利用大眾交通工具使用海灘 (Caro, 1975)。數學家 Hannah Fry 亦在其書中直接引用這個例子 (漢娜·弗萊，2018/2019)，作為技術與價值關係的探討。而 Thomas Campanella (2017) 則指出 Moses 雖然是個種族主義者，但是對於橋樑設計上排除有色人種之規畫並不是 Robert Caro 書中那樣單純。

的兩個基本特色：第一，權限分配的無差別性。第二，高效率執法機制的運作。現今許多法律的運作已經採納 AI 來協助法律行動者（法官、律師、研究者、一般人民等）完成特定的法律任務或是判斷。雖然在概念上我們可以區分輔助型的 ALI 以及替代型的 ALI，但是 ALI 的設計者並未體認到替代型 ALI 對於法治的破壞可能性。這不是意味著唯有工程人員才有可能擔綱法治的守護者。恰恰相反，本文企圖指出我們如何可以從工程學界對於 AI 的一般界定與特徵找出法學的深度切入點，並進而捍衛以下的價值立場：當 ALI 在規畫、設計與應用時，應該對於法治損害的可能性與類型納入考量，從而堅持輔助型人工智慧系統的方向。⁸

然而更深層的問題是：如果 ALI 做的比人類還要好，為何要堅持將特定權限交給人類來做？在某些情況下，將決策權限交給 ALI 來做似乎沒有不可。因此，這就有了一個需要分析的難題，我稱之為權限分配的難題。如何從法律觀點思考 ALI 的功能履行考驗著我們如何分析與論證法治的價值。因此，以下筆者將先簡短說明判斷權限的問題如何產生以及本文如何處理相關問題的架構與主張。

⁸ 關於替代型與輔助型人工智慧系統的相關說明，請參考 Pasquale (2019: 6)。Frank Pasquale 主張在管制方向上應該要以智能輔助型 (Intelligence Augmentation) 為主要理念，不應該在設計與管制政策上朝向替代人力或是人類決策的方向發展。筆者在這一點上實質贊成 Pasquale 的觀點。而 Pasquale (2020b) 也在新書《機器人新法則：人工智慧時代下捍衛人類專長》(*New Laws of Robotics: Defending Human Expertise in the Age of AI*) 中，多元與細緻地展現了人機互動模式中的這種智能輔助的運作更為恰當。

Pasquale 的呼籲其實並不孤單，因為 Stuart Russell 作為人工智慧研究的重要先驅者，也提出類似的擔憂，他認為必須要小心 (Russell, 2019: 9)。筆者認為，Russell 的呼籲值得重視的原因，不在於他是人工智慧的研發者，而是他指出那種以普遍研發起人類智慧機器的想法是如此根深蒂固地存在於工程學界，以致於他的想法是「新」的想法。

二、誰來駕駛？一個假想的例子

先從一個常見但假想的例子開始。約翰和瑪麗是好朋友，某天他們決定要共同參觀一處從未拜訪過的城市。約翰是賽車手，對於駕駛技術非常在行，而且對於即時路況的判斷也很專業。但是約翰的方向感不好，常常迷路。有趣的是，瑪麗不會開車，但擁有非常好的方向感，而且對於如何抵達目的地擁有極佳的抽象與路徑規畫能力。他們兩人決定要共同合作齊心抵達目的地，因此在任務分配上，他們做各自擅長的事情：約翰駕駛車輛，瑪麗指示方向。他們也順利地抵達目的地。

場景一：約翰的方向感還可以，但是沒有瑪麗那麼好。當他們第一次開車時，約翰遵循著瑪麗的指示前進。即使約翰有還可以的方向感，但是因為瑪麗的方向感更好，因此在路徑的選擇上，約翰有好的理由去遵循瑪麗的指示。

場景二：約翰的方向感很差，但他跟瑪麗已經造訪了這個城市很多次了。雖然他方向感不好，但由於造訪多次的經驗，約翰有理由按照自己過去聽從瑪麗指示與自身的駕駛經驗，來選擇路口的左右轉。雖然，約翰有理由倚賴於自己過去造訪成功的經驗來判斷，但由於瑪麗的方向感比約翰好，因此如果瑪麗說：「這裡右轉比較好」，約翰也有理由遵循瑪麗的指示。

在場景一即使約翰的方向感還算可以，他仍舊有理由要聽瑪麗的指示，因為瑪麗的方向感明顯地比他好。若假設約翰的方向感比瑪麗差很多，那麼他就更有理由要聽瑪麗的指示。在場景二，雖然約翰的方向感不好，但是由於他已經成功地聽從瑪麗的經驗，所以他逐漸學習到哪裡該左轉與哪裡可能會塞車的經驗，因此約翰有理

由遵循自己的判斷，但是這個理由是初顯的 (*prima facie reason*)，因為它可能被「瑪麗有更好的方向感」這個理由給超越。不管如何，瑪麗判斷方向的能力比約翰好很多，也因此約翰必須思考他是否要遵循瑪麗所給的方向指示。

假設約翰對於方向的判斷跟瑪麗一樣好時，站在瑪麗的角度，她有理由不需要向約翰指示怎麼走。即便瑪麗給了約翰指示，約翰也有理由遵從自己的判斷，但約翰也可以選擇遵從瑪麗的判斷。這意味著，「誰來判斷方向」作為一個被分配的決策權限，不論是交給約翰或是交給瑪麗都可以。這可以稱之為**決策分配的無差別性**。這是一種事前觀點，當然也可以採取事後觀點檢查到底當初在某個路口如果聽瑪麗的會比較好或是比較差，但是這涉及到反事實判斷 (*counterfactual judgment*)，並非本文處理的重心。⁹

讓我們改變一下剛剛的例子。約翰是人類的駕駛者 (*human driver*)，而瑪麗是 Google Map。可能的改變條件是：第一，目的地是約翰單獨決定的，而不是約翰與 Google Map 共同決定 (*decide together*)。第二，由於目的地是約翰單獨決定，約翰僅需要考量 Google Map 的「能力」即可，而不需要考量 Google Map 是否「是」一個人。第三，即使目的地是約翰自己設定，但是這仍舊沒有改變約翰需要 Google Map 協助的可能性，約翰依舊方向感很差等等。因此，在第一次造訪目的地時，約翰有理由在開車的時候聽從 Google Map 的導航到目的地去。第四，當約翰開了一次兩次三次之後，他就沒有理由一定要聽從 Google 的導航系統，除非他需要更準確地知道自己何時要抵達目的地。但是，在某些情形下，約翰

⁹ Judea Pearl 等人在《統計中的因果推論：入門介紹》(*Causal Inference in Statistics: A Primer*) 一書中涉及到反事實判斷的因果討論時，就以高速公路或一般道路的道路選擇作為反事實論證的具體例子。請參考 Pearl et al. (2016: 89-91)。

更有理由依賴於 Google Map，而規畫與安排自身的一系列行動。例如，當 Google Map 在約翰手機跳出一個訊息：「本路段 30 分鐘後將有車潮湧現的可能性」。約翰有理由即刻啟程。甚至，在一般的情形下，約翰也有理由開啟 Google Map，因為 Google Map 會即時顯示車流資訊（綠色、橘色或紅色），約翰可以藉由著車流顏色來判斷接下來可能會發生的變化。因此，當人類運用人工智慧要達成某個自身所設定的目標時，在判斷決策權限的分配上，並沒有太多的理由認為一定要將判斷決策交由人類來做。相反地，可能更有理由讓人工智慧來替我們做某些事情。¹⁰

三、文章核心主張與架構

從上述假想例子進一步拓展到法律之中，問題就是：法律是否允許法律人工智慧系統將法律事務的判斷權限全然由人類移轉到機器（或是人工智慧系統）身上呢？

就如同羅爾斯 (John Rawls) 在《正義論》(*A Theory of Justice*) 中，指出「正義是社會制度的首要德性，如同真理之於思想體系一般」(1971: 3)。法治也是法律的首要德行。法治的主要目的是限制國家公權力之運作，將統治者的個人意志 (will) 對於權力運作之影響降到最低，因此，對比於法治的概念是人治 (rule of men)。然而，所謂的法治不是統治者可僅「用」法律的形式來統治，換言之不是「藉法而治」(rule by law)，而是統治者必須「依法而治」，或者更精確地說是依理性 (reason) 而統治。

¹⁰ 例如，在前面所提及的商務律師跟 ALI 的爭點辨識競賽上，事後商務律師被訪問提及，是否會擔心工作被 AI 所取代？律師一般都認為不擔心，因為他們反而可以有更多的時間去做更多的事情。此類規律性的爭點辨識就可以交給 AI 來做。

思考 ALI 與法治的關係上，概略有兩種觀點：相容論與不相容論。相容論認為，對比於人治，法治要擺脫統治的意志恣意性要素，ALI 在設計與應用上就是讓法律的執行與判斷與人的意志恣意性有所區隔，因此 ALI 與法治並沒有扞格之處。¹¹ 相容論有以下幾個理由：(一) 系統在運作時並不會涉入個人情感（根本沒有個人這一回事，僅有機器）。(二) 資料的分析也是大量（例如資料驅動的人工智慧系統）與某種程度是客觀的，雖然資料可能有偏見，但是由於個人判斷也會出現偏見，因此偏見與否不是機器特有的問題。(三) 既然偏見是同時存在於人與機器之中，我們便不應該以較為嚴苛的標準評價 ALI 的優劣。需要注意的是，相容論有兩個特色：第一，檢視是否採用 ALI 的標準應該與檢視一般法律行動者的標準相一致。第二，ALI 本身所具有的問題可以透過技術改良來解決，或許問題之解決遙遙無期，但只要該問題也會發生在人身上，那就不應該成為反對應用 ALI 的理由。

相較於相容論，本文要捍衛的觀點是不相容論。不相容論主張，ALI 與法治有內在的緊張關係，這個緊張關係的來源有兩個：(一) 依法而治的主體想像要求權限判斷必須交由人來做。(二) 高效率執法機制對於法律主體的自主性的損害。雖然，這種緊張關係不必然會造成法治的嚴重損傷，但仍必須小心 ALI 的運用界線。運用界線的議題無法一概而論，這仍有賴後續的研究，然而本文若可將 ALI 與法治的緊張關係分析清楚，那麼至少有了一個清明的出發點。

本文主張，法治的核心理念想像了法律主體可以行使自主的道德判斷，法治並非期待個人成為法律指令被動接收者而已，而是期望個人藉以衡量法律之內容與其自身持有理據的輕重關係。換言

¹¹ 這也出現在許多宣稱使用 ALI 的優點之上，例如客觀以及不帶太多個人主觀情感等。

之，行動者必須思索諸多涉及法律理據的重量，從而判斷自身是否要臣屬於法律的要求之中。然而，ALI 的發展越來越會威脅此種奠基在人文主義精神開展出來的價值，且試圖壓縮個人自主道德判斷的空間。或許在初期這種緊張關係不太明顯，但隨著 ALI 的功能履行越來越多樣或效能越來越好時，這種緊張關係會越來越凸顯。雖然法治同時要求國家公權力與公民遵守法律，但法治（即便是形式意義上）對於什麼是法律主體 (legal subjects) 具有實質的價值蘊含與目的導向的界定，ALI 如果忽視了法治的這個目的導向，將會導致法治本身的規範性基礎的喪失，這是當前急迫需要被重視的。

本文的基本架構與內容如下：

第一部份主要闡述問題意識與架構，並且簡單界定 ALI 的概念以及筆者的核心主張跟分析方法。

第二部分進一步指出在思考 AI 與人類之間的一個根本問題，我稱之為「權限分配的難題」，並且以此難題來定位 ALI 與法治的關係。我將指出在功能履行的觀點下，當 ALI 對於特定法律任務的履行跟人類不相上下時，若沒有特別需要考量的理由下，那麼權限如何分配原則上是沒有差別的，這就是權限分配的無差別性。權限分配的無差別性是相容論的核心主張，其關鍵論證是能力宣稱 (Capacity Claim; CC)。權限分配的無差別性幾乎是所有 AI 的特徵，但是 ALI 還有另外一個重要的特色：高效率執法機制。使用 ALI 可以更有效率地執行法律任務以及更加落實法律規定。因此，效率 (efficiency) 的提升與實效 (efficacy) 的落實也是支持使用 ALI 的理由。我在第二部分將說明此類 ALI 的特色。

第三部分主要處理法治的概念觀。Richard Fallon (1997) 梳理出的四種法治概念觀（歷史理念型、形式理念型、法律程序理念型以及實質理念型）可以讓我們看到不同法治理念的價值取向跟實踐

目的。我將進一步抽象化上述四種法治概念觀，指出雖然其四種法治概念觀對於法治的規範性基礎有所爭論，但是「人的圖像」卻是法治諸多理念的共同特色。法治雖然是將統治脫離人治，但並不是將人變成法治的客體，法治有一個實質的價值內涵，那就是維持受法律所統治的個人之道德自主判斷空間，這表現在法治雖然要求個人遵守法律，但是個人對於如何遵守法律以及判斷守法與否仍保有最終決定權限，法治將遵守法律的理據之優先次序交由個人（或法律主體）來決定。

第四部分則是針對ALI的規範性批判。這個部分著重兩個主題：第一個主題是對於權限分配的無差別原則的批判。我將以第三部分的法治概念觀為基礎，逐一分析相容論的可能論點，然後說明其不足之處以及最可能成立的論點何在。第二個主題是針對ALI所具有的行動協調功能進行反省，我將指出原則上ALI的應用是輔助行動者更加理解遵守法律的一階理據，而不是著重以制裁為主的管制方式。第五部分則是結論。

貳、權限分配的分析與法律人工智慧系統之特色

一、同等履行下的無差別原則與能力宣稱

在上述約翰與瑪麗的駕駛例子表達出一個初步的觀點：針對特定事務的履行，兩個人的能力不相上下時，不論將該事務履行所涉及的判斷權限分配給哪一個人都是許可的 (admissible)。這個觀點以下述原則表達出來：當A跟B有同樣的能力履行某項任務時，讓A或B單獨行使該決策權限都是許可的。筆者將此原則稱為「同等履行的無差異原則」(Indifference Principle under Equal Performance; IPEP)。權限的分配者可以有兩種反對IPEP的理由：第一，

分配者可以主張，A 跟 B 的能力並不一樣，例如 A 比 B 更有能力。第二，分配者可以主張，即便 A 跟 B 能力相同，但是分配者擁有更多一些理由讓 A 來做。第一種反對 IPEP 的理由稱之為非同等履行的論證，第二種反對理由為附加理由論證。需要注意的是，這兩種反對理由的著重點不太一樣。非同等履行論證針對的是探討 IPEP 的前提是否成立，檢視是否有同等履行。附加理由論證是著重，有額外的理據支持為何要讓其中某一個人擁有判斷權限。前者的例子是，當瑪麗下指示時，約翰可能會說，你這幾天都在感冒，腦袋不清楚，所以我來判斷怎麼走。後者的例子是，約翰對瑪麗說，你等一下有重要面試，我開車時你就好好沈澱心情，準備等一下的工作面試。就前一種回應，約翰提出的主張是指，瑪麗沒有好的能力來履行方向判斷的任務，第二種回應之理由是，瑪麗即使有能力給予指示，但是有另外一個更好的理由不讓瑪麗給予指示。

IPEP 的運用上有一個關鍵的宣稱：能力宣稱。能力宣稱主張，分配者在思考如何分配判斷權限時，他其實無法真正的預知未來會是如何，因此，較好的比較基準點是思考被分配對象在功能履行上的能力是否相同。若相同，則 IPEP 就成立，若不相同，IPEP 不成立，因此分配者必須探詢其他理由分配權限。

CC 幾乎是在思考機器可否代替人類做判斷的特色。多數人工智慧學者都認為，當機器在履行特定功能時與人類有一樣的能力時，那麼將特定事務的判斷權限交給機器來做是允許的。而判斷機器是否擁有跟人類一樣的某種能力時，這端視機器在任務之履行上是否達成跟人類一樣的效果。要注意的是，AI 不需要擁有比人類更強大的能力，它僅需擁有跟人類差不多一樣的能力即可取代原本人類要做的判斷或任務執行。此外，功能的運作也不需要依賴於相同的物質基礎，換言之，不同的物質構造物可以擁有相同或者類似的

能力。例如，人類的大腦具有記憶的功能，而電腦的硬體也可以具有記憶的功能。但是人類的大腦跟電腦的硬體在物質構造上是不一樣的。這種物質構造的差異性但卻可以擁有相同能力或者功能的履行，就是計算機功能主義 (computer functionalism) 的基本想法。¹²

關於 CC 有幾點必須加以說明：首先，這裡的能力不是指意識能力 (capacity of consciousness)，所謂的意識能力，Susan Schneider 指出人工智慧意識的問題 (Problem of AI Consciousness)，是指「AI 的運作是否會以某種內在方式而體驗到」(2019: 18)？但 AI 是否有以及是否可能有意識能力不是筆者關注的重點。其次，capacity 與 capability 雖在中文都譯為「能力」，但其實 capability 比 capacity 更高一個層次。Martha Nussbaum 的能力取向 (Capability Approach) 是一種實質的規範理論，主要關注在人可以做什麼以及成為什麼 (what people are able doing and being) (Robeyns & Byskov, 2020)。然而，筆者此處所關注的 capacity 不必然是要自然人或是生命體，其 capacity 擁有者也不一定要針對自身存在之目的、意義與價值加以反省。¹³ 這裡所提出的能力是指，促使其完成某任務或履行某功能的條件。能力雖然實際上看不到，但卻可以透過任務履行的方式被展現出來。例如，Siri 的語音輸入表現越來越好就意味著，其在履行語音轉換成文字的能力越來越好。因此，筆者暫時擱

¹² 關於計算機功能主義，最為人熟知的大概是 Ray Kurzweil (2012) 在 *How to Create A Mind* 第八章。筆者認為幾乎所有的工程師都有某種計算機功能主義的傾向。例如，Max Tegmark (2017: 70-107) 也說明關於記憶與運算等能力，如何在人類與機器相異的物質組成下，而可實現相同功能。而根據 Tomas Hauer (2018) 的說明，計算機功能主義的想法就是認為心靈是一種抽象功能的運作，可以在任何物理結構中被實現，但心靈無法被化約到物理結構的法則或是概念。請參考 Hauer (2018: 100-106)。

¹³ 對於自身目標之設定以及反省是一種很高階的能力，通常已經到了 wisdom 的階段。蔡政宏稱之為 artificial wisdom，請參考 Tsai (2020: 939)。

下能力 (capacity) 背後的形上學爭論，而著重在這個能力宣稱是與功能履行相互搭配的這一個面向上。採用這個方式有個好處是，容易與人工智慧工程學者進行對話，例如 Stuart Russell 與 Peter Norvig (2010: 40-41) 在《人工智慧：現代取向》(*Artificial Intelligence: A Modern Approach*) 中指出 AI 在設計上首要之事就是特定化其任務環境 (task environment)，這個任務環境由四個要素所構成：履行測量、環境、執行器與感知器 (performance measure, environment, actuators, and sensors)，稱之為 PEAS。其中的 performance (履行) 就是用來確認人工智慧系統對於任務完成所需要的目標。在筆者的界定下，當我們說某個人工智慧系統可以適當實現某個目標，就意味著它有實現該目標之能力。

CC 幾乎是所有人工智慧設計者的核心想法。因為工程師在設計 AI 時，他們所設想的系統並不是憑空出現，而是對於這套系統要用來做什麼 (what an AI is made for) 有非常特定的想法。AI 必須擁有工程師所設定的目標、系統能夠生產出實現被給定之目標的策略、執行該項策略，並且該執行可以影響與系統所互動的環境等等。因此在 AI 的規畫設計與製造過程中，工程師必須將一個特定任務切割為不同次任務，確認次任務的執行是可以被電腦所計算等等，當諸多履行不同的次任務的系統可被建構出來時，這些不同的次系統可以整合起來執行一個特定的任務。如果被整合起來的系統可與其他系統再度整合，並且完成另一個特定的任務時，那麼這個系統就可以移作他用。換言之，能力可以被整合以及被提升。也因此，從此角度來看，對於能力宣稱的反省是一個非常重要的面向。¹⁴

¹⁴ Russell (2019: 94) 在書中描繪了一種想像超級智慧機器的可能，雖然他並未區分 capacity 與 capability (他反而是使用 capabilities 一詞)，但是卻也開啟了一種能力整合的想像可能。

二、效率、實效與實踐差異做為法律人工智慧系統的特色

根據上一節的討論可以見到，能力宣稱幾乎是所有支持 AI 的特色。而效率是 AI 在應用上的一個重要考量，當系統建置完成後，由於機器運作的特色，因此它在特定任務履行上相較於人類的確可以更有效率的完成。¹⁵ 然而，法律人工智慧系統除了在應用上相對有效率外，還有一項特色：實效。實效一般而言，就是指法律內容落實的程度，在法哲學中實效通常是與效力 (validity) 的概念作為對比，因此 H. L. A. Hart 在《法律的概念》(*The Concept of Law*) 一書中指出，「說某個既存的規則是有效，就是肯定它已通過所有承認規則所提供的判準，並成為法體系規則中的一員」(哈特，1961/2010: 133-134)。¹⁶ Hart 進一步指出，「如果『實效』意味著，一項規範某種行為的法律規則大部分時候都會被遵守，那麼很清楚地，任何規則的『效力』與其『實效』之間實在沒有必然的關係。除非，該體系的承認規則在其諸判準中包含以下這種有時候被稱為『廢棄規則』的條款：任何規則如果早已不再具有實效，則不能被

¹⁵ 此處講的效率不是指法律制度設計中的效率概念，例如訴訟制度之設計如可以達成特定法律目的，也不是專指法律經濟學上的效率概念。筆者在此是以一種較為寬鬆意義使用效率一詞，泛指一種對於使用人工智慧系統的理據內容，認為當系統在研發完成後，使用者使用人工智慧系統履行特定任務跟自己履行特定任務所需要付出的資源是較小，或是更小。的確，「效率」不論在經濟分析或法學中具有相當重要性，相關討論請參考 Farnsworth (2007: 13-23) 與 Bix (2004: 59-60)。關於法律經濟分析在法學應用的討論，請參考王鵬翔、張永健 (2019)。該文中，王鵬翔與張永健詳細區分一階效率與二階效率的概念，作為論證經濟分析如何以效率原則與傳統法學方法相互搭配。張永健在其專書中，更進一步指出效率是經濟分析評價結果的核心標準，但對於什麼是效率容有不同的概念觀。筆者由於能力所限，無法針對「效率的概念觀」進行深入探討，有興趣的人請參考張永健 (2020: 58-64)。

¹⁶ 原文請參考 Hart (1994: 103)。

視為該體系內的規則」。¹⁷ 因此，實效是指法律規範在特定社群中落實的程度。當我們說一個法律規範具有實效時，意味著在這個社群中，至少某些成員普遍地遵守系爭法律規範，或者，社群成員的行為與規範內容相一致。¹⁸

從實效的角度來看，ALI 的使用就是要影響社群成員的行動，讓社群成員的行為與規範內容相一致。社群成員對於法律的態度可以是分歧的，但成員所表現出來的行為至少是與法律規範之內容相一致的。用更理論的術語來表達就是，ALI 要縮小規範性內容與實效間的落差，將「應然」與「實然」擬合起來。這種擬合的方式就是要讓 ALI 之應用對於行動者的行動產生差異。

AI 應用在法律的層面可能在以下幾個運作層面對於行動者產生影響，例如針對事實判定、法律解釋、法律推理與結果預測等等。

舉例來說，在事實判定上，無論是測速照相機 (traffic enforcement camera) 或是英國的 CCTV Enforcement Vehicle (Gateshead Council, 2013: 1)，都是利用機器確認違規的法律事實，透過技術的改善讓違規事實的認定更為迅速。此外，在法律解釋上，法律推理的計算機模型 (computational model of legal reasoning; CMLR) 係關注於將法律條文以及法律推論給程式化，藉以讓使用系統的人可以更快地知悉某個案件之可能推理過程。而結果預測的最典型運用，則是透過機器「閱讀」判決文本，建構出判決結果預測的演算法模型，以供其他類似個案運用。結果預測的 ALI 之運用脈絡可以

¹⁷ 需要特別說明的是，筆者在此譯文完全採用中譯本翻譯，但是中譯本在這段引文最後一句是逗號，將原文內的下一個段落直接連接起來。為了忠實原文起見，筆者是以句號表示。原文請參考 Hart (1994: 103)。

¹⁸ 與 Efficacy 相同的詞彙亦有 effectiveness，但都是指規範確實被遵守的情形。請參考 Kelsen (1991: 138)。

很多樣，例如可以用來知悉法院在判決時的可能考量、協助當事人進行訴訟策略規畫，或是讓訴訟代理人知悉個案的相關訴訟成本等等。

不論是上述的哪一種 ALI，其關鍵特色就是希望影響社群成員的法律判斷或是行動。而支持採用 ALI 的人往往會以效率與實效作為使用的規範性理據。這個理據可以稱之為「執法效率」的提升：系統要達成之效果是讓法律判斷（事實認定、法律解釋或是結果預測）與法律的落實更為迅速，藉以彌補人類的有限判斷能力與成本（例如記憶力有限、無法長時間關注或是判斷失誤等等）。從這個角度來看，支持運用 ALI 的人有兩種相關聯但不相同的論據。第一個論據是效率論據 (efficiency argument)，效率論據是指當某個法律任務交給人工智慧系統來做時，系統執行上比人更有效率，系統可以比人更快地執行完任務。第二種論據是實效論據 (efficacy argument)，應用人工智慧系統可以提升遵守法律的程度。

本文使用高效率執法機制的概念來指涉，結合效率與實效論據的法律人工智慧系統。具有高效率執法機制特色的 ALI 的主要目的是用來降低法律規範與法律規範之落實之間的落差。在當代社會之中執法機制的存在擔保了法律最低程度獲得遵守的條件。擁有執法機制雖然無法確保法律百分之百會獲得遵守，但是沒有執法機制法律的實效性一定會出現問題。因此，ALI 的主要效用是將法律規範的內容與現實世界社群成員行動之間的落差給降低，甚至是減縮至零。ALI 如何降低法律規範內容與現實世界社群成員之行動間的落差呢？**關鍵在於，實效的實現取決在系統有效率的運作，高效率執法機制透過即時告知社群成員可能違法事實（或與法律內容偏離之事實）的發生，讓成員調整自身行動，或是告知成員該違法事實會被人工智慧系統所偵測到，而改變成員的行動。**

讓筆者以幾個實際例子，說明這種高效率執法機制的特色。¹⁹

區間測速執法是執法單位在舉發超速的方式之一，主要是將某車輛「特定兩點的距離」除以「通過兩點距離的時間」所得出的平均速度。若該平均速度高於法定限速，那麼在計算上可推論該部車必定有「某個時間點的速度高於法定限速」。為確保區間測速的確實執行，系統必須要能夠完整且全面的捕捉特定車輛通過兩點的時間與車輛辨識，以及自動計算該車輛通過特定兩點的平均速度。²⁰

筆者先擱下區間測速的執法設備的適法性議題，而集中在區間測速的運用如何展現出 ALI 高效率執法的一面：它要求特定車輛在兩點間速度**必須由駕駛人控制**在法定限速之內。不可否認的，現行的交通法規下駕駛人的確負有遵守法定限速的義務。然而區間測速的執行不僅僅是告訴駕駛人必須遵守法定限速，也告知駕駛人在特定兩點間**有時時遵守**法定限速的義務。²¹ 這種對於駕駛人遵守法定

¹⁹ 審查人針對本文所使用的兩個例子（區間測速與欠稅車輛之偵測）提出了建設性的意見，認為這兩種機制著重在事實判定，因此可能不算是法律人工「智慧」系統，例如它們的運作方式都不具備關聯性預測，所以也沒有智慧可言。審查人在此精確地指出此類系統的「智慧疑慮」，筆者完全同意。不過，筆者使用這兩個例子主要是說明高效率執法機制的特色。此外，筆者認為此類執法機制的應用本身雖然是事實判定，但是其系統設計本身是可以應用人工智慧技術的，因為影像辨識現在非常普遍是透過機器學習技術進行。也因此，即使僅是事實判定，但是從工程的角度來看，系統也可以被稱為人工智慧。審查人另外一個關鍵批判是，因為「區間測速與欠稅車輛之辨識」僅做出事實認定，因此其並未取代執法機關的法律判斷與決策權限。針對這一點筆者有不同意見，將於下文中說明。

²⁰ 在具體運用上區間測速一開始運用在北部的萬里隧道，之後擴展到北宜公路，最後擴展到全國的運用。

²¹ 可能有人會說，執法單位有舉發的裁量權，因此法定限速之違反並不一定會引發違法的法律效果。但是這只是將問題延後而已，因為既然裁量權是執法單位所擁有的，執法單位也可以限縮自身的裁量範圍，或者捨棄裁量。如此一來問題仍舊存在。

速限義務的提升是一個質的改變，並非量的改變而已。因為，定點測速照相同樣也具有告知駕駛遵守法定速限義務的功能，但是它並未要求駕駛人時時刻刻需要將速度侷限於法定限速之內。道路駕駛作為一個具體判斷力的實現場域，要求駕駛人判斷具體情境如何與法律要求相容。可以這麼說，駕駛人是將法律之要求作為一個自身行動的理由與其自身所面對的具體情境所產生的其他理由相互衡量。區間測速所取消的是這樣的一種衡量可能性，因為透過區間測速所產生的高效率執法的效果，它告訴駕駛人以下這個訊息：遵守法律，否則就處罰。

第二個例子則是交通部所推動的「交通正義專案」。這個交通正義專案的主要目的是提高對於欠稅車輛之查緝效率。該專案將「辨識欠稅車輛」的任務外包給人工智慧系統，讓行政執行署官員或是停車費收費員可以攜帶車牌辨識系統，於執行勤務時同時尋找「可能欠稅車輛」的蹤跡。該任務分配的方式主要有三種：(一) 由停車場收費員於收費時輸入車牌，同時辨識該車輛之車主是否有欠稅紀錄，若有則通知行政執行署的官員，視情況是否進行扣押。(二) 由行政執行署官員直接開車，配載車牌辨識系統，由官員開車上街直接尋找是否有欠稅車輛。(三) 與車牌辨識的停車場合作，將進入停車場之車牌辨識與比對之後的結果傳給行政執行署的官員，告知其停車場內是否有欠稅的車輛。雖然這三種任務分配的方式各有其法律爭點，但仍可見到高效率執法機制的實現是其主要目標。以車牌辨識為主的 ALI 所具有的能力，遠超過一般人類的執法者，因為它儲存的車牌資訊可以無限量、辨識車牌的時間與地點也不受侷限。諸如此類的 ALI，即便其所執行的法律任務是非常單一（辨識出欠稅車輛之車牌），其執行任務所需的能力也非常單一（影像辨

識)，但是它對於執法效率的提升卻有顯著的效果。²²

透過上述兩個具體實例，筆者希望指出的是，高效率執法機制是一個亟需討論的課題，高效率執法機制的目標是將法律規範與其實現程度的落差 (gap) 給降低甚或是擬合 (fitting)，這樣的特色雖可能被認為是法律的實現，但是 ALI 是如何達成這個落差的呢？是一個需要透過理論思辨而討論的課題。

既然高效率執法機制是要促成行動者的行動改變，那麼一個根本問題是：**法律如何促成行動改變呢？**一般來說，社群成員在進行決策判斷時涉及到的是實踐理性 (practical rationality)，實踐推理涉及到的是我應該做什麼的問題，與此相對的是理論理性 (theoretical rationality)。²³ 法律與實踐理性的課題在法哲學中有豐富的討論，其中一個重要主題就是法律的實踐差異 (practical difference of

²² 審查人指出這兩種執法系統都只是確認做出法律判斷的基礎事實是什麼，但是使用這兩種執法機制並未取代執法機關的決策權限。這是一個深刻的批判。的確在某個意義上，執法機關仍然握有最終決策權限。但關鍵問題可能是，相較於真人，一個系統若僅能做到事實認定，其認定事實就沒有判斷權限的行使嗎？筆者在這一點有不同意見。例子如下：警察甲使用手持測速照相機偵測超速車輛。假設甲看見高速公路上，中線行駛著一輛時速 90 公里的砂石車，當時因風大砂石不斷從車旁落下，其砂石車後的小客車以時速 120 公里從內側車道快速超車。甲測得小客車的時速為 120 公里，也知悉小客車是因為超車而加速。此時，甲可能選擇要不要讓「小客車 120 公里的時速」這個事實成為法律事實，進而開單告發。如果是自動測速照相，「小客車 120 公里的時速」就會成為法律事實。在甲的例子中，的確不會說手持測速照相機「取代」了對於車速是怎麼多的判斷，因為沒有測速機器的輔助，甲根本無法判斷。然而一旦甲對比於自動測速照相，我們會說要讓何種事實成為法律事實，這裡面有個判斷權限，而這個判斷權限在自動照相的例子中是移轉到機器上面。當然，可以進一步討論「判斷」到底是什麼以及機器是否真的在進行判斷，但這不是此處的主題。

²³ 簡單來說，實踐理性涉及到的是行動，理論理性涉及到知識，前者涉及到工具理性，後者涉及到信念的證成。關於這兩者之間的關係筆者無法在此加以說明，但基本上哲學家認為維持這兩者之間的區分是有必要的。例如艾莎·威克福斯 (Asa Wikforss) (2017/2021: 90-94) 就抱持這樣的想法。

law)。簡單來說，就是法律如何對於行動者的實踐推理產生差異。過去英美法哲學在這個部分有豐富的討論，筆者在此將簡單說明 Scott Shapiro (1998) 的研究，闡述法律如何創造實踐差異的兩種模型。²⁴ 此外需要特別指出的地方是，本文所指的高效率執法機制所促成行動者的行動改變是指對於一般公民而言，並未觸及到適用法律的官員。²⁵

(一) 法律造成實踐差異的兩種模型：決策模型與侷限模型

法律可以影響行動者的實踐推理，法哲學一般稱之為「實踐差異命題」(Practical Difference Thesis)。²⁶ 實踐差異命題認為，法律若具有規範性，則法律必須要能夠對於人的實踐推理產生影響，法律規則可以影響實踐推理意味著法律規則可以指引行動 (或行為)，

²⁴ 一般在 ALI 的分析之上多著重 AI 如何對於法律體系既有的價值與權利產生影響或傷害，我們可以說這是權利中心的討論。權利固然是當代法律的一個重要保障對象，筆者承認此類討論的重要性，但是筆者認為我們同時也要關注到高效率執法機制的可能影響，更精確地說是關注到高效率執法機制對於規則遵循行動 (rule-followed actions) 的影響，因為有一種可能性：ALI 並未侵害法律主體的基本權利，也沒有持續地進行監控，但是其高效率執法的特色卻會造成行動者在衡量行動理據時的重要改變，而這個改變將會損及法律的核心價值，筆者將在下一個部分說明這種可能性。

²⁵ 感謝審查人指出這個本文在分析上沒有明白闡述的前提。本文在思考關於遵守法律與否的理據結構時確是以一般公民為探討對象，經由審查人提醒，發現到適用於執行法律的官員的實踐判斷很可能跟一般人有非常不一樣的結構。由於文章篇幅所限，這個部分筆者僅能暫且忽略。以下本文所闡述的因果結構圖都是以一般公民為設想對象。

²⁶ 關於實踐差異命題，從理據論角度的中文細緻討論，請參考王鵬翔 (2012: 221-230)。據筆者知悉，中文學界較早提及的應該是顏厥安 (2006: 61-89) 的文章，其運用這個概念來說明規範縫隙理論。筆者在此不加以區分王鵬翔所提到的，強與弱的實踐差異命題的主張，根據王鵬翔從理據論的脈絡界定弱的實踐差異命題是指，法律可以產生影響行動的規範性理據，而強的實踐差異命題是指法律可以改變既有的理由結構。這部分的區分很重要，但是不是本文的關切之處，因此僅能存而不論。

因此 Shapiro 說：「唯有當規則——以某種方式——影響了行動者實踐推理的產出時，或許才能說行為是被規則所指引的」(1998: 44)。

Shapiro (1998) 在〈規則所造成的差異〉(“The Difference That Rules Make”) 一文中對於規則如何創造行動差異提出了兩個概念性模型：決策模型 (decision model) 與侷限模型 (constraint model)。²⁷ 以下我將一般性說明這兩種模型對法律如何造成實踐差異的不同之處，由於 Shapiro 在該文章中涉及的議題非常深刻，因此筆者必須有所取捨，將其部分主張排除在說明之外。

Shapiro 表示，決策模型是指「當行動者基於工具性目的採納規則時，行動者依然保有不依循規則的能力。然而，當行動者認為規則適用時，他們會選擇遵循規則」(1998: 38)。位居於決策模型的核

²⁷ 在此有必要先說明 Shapiro (1998) 這篇論文的脈絡。本篇文章主要收錄在 Brian Bix 所主編的《對法的分析：法理論中的新研究》(*Analyzing Law: New Essays in Legal Theory*) 一書之中，本書所收錄的論文主要奠基於 1996 年在 Quinnipiac College School of Law 的一場研討會，研討會主題是針對 Jules Coleman 的法政哲學思想進行討論，因此嚴格來說，Shapiro 的對話對象是針對法實證主義內的柔性法實證主義者 Coleman。在一篇人工智慧與法律的文章中「貿然引用」法哲學的討論似乎有點唐突，但筆者認為分析法理學的功能之一就是以清晰的視野說明相關的概念與主張，Shapiro 在〈規則所造成的差異〉一文中雖然以批判柔性法實證主義為主要目標，但是他對於規則如何影響行動的說明是非常清楚的，因此有助於我們理解法律人工智慧系統要履行法律任務時，可以如何影響行動的討論。然而，需要注意的地方是，筆者雖然使用 Shapiro 的討論，但是筆者不一定贊成 Shapiro 在相關法哲學上的主張，例如法與道德分離命題等等。這個關鍵主張之支持與否不影響此處的討論。此外，Shapiro 在 2002 年的《牛津法理學與法哲學手冊》(*The Oxford Handbook of Jurisprudence and Philosophy of Law*) 一書中的「權威」(“Authority”) 這個辭條中，亦使用了決策模型與侷限模型來說明，但具體例子不一樣。筆者在此先對於這部分的說明擱下不論。有興趣者請參考 Shapiro (2002: 415-430)。最後是，Shapiro 針對法律如何創造實踐差異的兩種模型的說明在其 2011 年出版的《合法性》(*Legality*) 一書中並未見到相關說明，這倒是令筆者非常好奇。一種可能的理由是，Legality 在某個意義上是 Shapiro 運用行動哲學要拓展的新領域，或許他不認為實踐差異命題是一個需要在其計畫理論中被討論的主題。關於最後這一點仍舊有待進一步探索。

心命題是被 Shapiro 稱之為**偏好命題** (Preference Thesis)，決策模型認為，行動者認為某個規則適用在個案中的這個事實會讓行動者偏好的判斷產生改變。因此，偏好命題是指：行動者是工具地被規則 R 所指引行動，若且唯若以下條件被滿足：

1. 行動者的行動服從於規則 R，
2. 且承認規則 R 適用於既存之個案中會使得行動者相信他偏好服從而非不服從，且
3. 行動者相信，若不認為規則 R 適用於自身的話，會使自身做出相反的判斷 (Shapiro, 1998: 35)。

在偏好命題的觀點下，法律是**透過改變行動者的偏好**而產生實踐差異的，而這個偏好的改變是以法律是否適用於行動者當前狀況為成立條件的。

除決策模型之外，Shapiro (1998: 39) 指出有另外一種看待法律創造實踐差異的方式，他稱之為**侷限模型**。侷限模型是指，**採納規則**的行動者是策略性地與另一個行動者互動——日後的自己 (his later self)，現在的自己 (present self) 採納了規則，決定讓自己被規則所指引，因此現在的自己透過規則侷限了日後的自己可以行動的選項。現在的自己跨出第一步，決定自己被規則所指引，而如果這第一步成功的話，日後的自己將被禁止去選擇規則所排除的選項。換言之，讓原本的可實現行動變成不可實現的。因此，規則之所以能夠創造實踐差異，不是因為行動者有能力選擇依循規則，而是規則讓行動者欠缺選擇 (*lack of choice*; 斜體為原文所有)。由此可知，侷限模型關注的是未來行動可實現性 (feasibility of future action)。佔據侷限模型的核心命題是可實現性命題 (Feasibility Thesis)。而

可實現性命題是指：行動者是工具地被規則 R 所指引行動，若且唯若以下條件被滿足：

1. 行動者的行動服從於規則 R，
2. 且承認規則 R 的適用侷限行動者要去服從 R，該承認也使得不服從 R 是不可實現的，
3. 且行動者相信，欠缺規則 R 的誘發的話，他可能不會遵守 R (Shapiro, 1998: 47)。

決策模型與侷限模型有以下幾個重要相似與相異之處。這兩者的相同之處在於：第一，兩者都是針對法律如何具有規範性的說明。第二，兩者都認為法律本身足以針對實踐推理產生影響，而不依賴於法律的內容。第三，它們都接受行動者可以工具理性地運用法律。

然而，決策模型與侷限模型的差異在於以下幾點：第一，決策模型並不考慮行動者在時間過程中產生的偏好改變的問題，例如當行動者決定適用規則時，如果他發現適用規則反而沒有辦法達成自己預期的效益時，決策模型無法說明這個狀況。第二，侷限模型區分了對於規則的思辨 (deliberation) 與對於規則的實施 (implementation)，這兩者有概念性差異。Shapiro 說：「當行動者在思辨時，他的任務，在極大化行動者效益的傾向向下，評價其他替代方案的可行性。理想上，他應該要衡量所有行動的理據，一旦找出最佳的選項後，他就可以自由選擇想要實施的選項」(1998: 49)。為何區分思辨與實施這兩者是重要的呢？Shapiro (1998: 49-50) 指出，我們往往會預先規畫，而在預先規畫時規則對於思辨的考量就會產生影響，但是預先規畫到實施規則所許可的行動中間往往會有時間落差，此時要適用某一個規則便是另一回事。此外，思辨是效益的極

大化，但是實施不是。在實施的階段，行動者需要做的是找出哪些選項是依循規則的行動，而不是哪些選項是最有價值的。

(二) 高效率執法機制如何影響行動者的行動？

現在我們知悉法律可以如何影響行動者行動，在這個基礎上可以進一步說明 ALI 如何影響行動者的行動。首先，我以遵守速限法規為例子，簡短建構出行動者遵守法律的因果圖，說明行動者要思索遵守法律與否的可能因素（變項）。其次我將說明在選擇變項進行干預上，採用決策模型和侷限模型的 ALI 會有何差異。

圖 1 是筆者勾勒出行動者在遵循法律與否的因果圖：²⁸

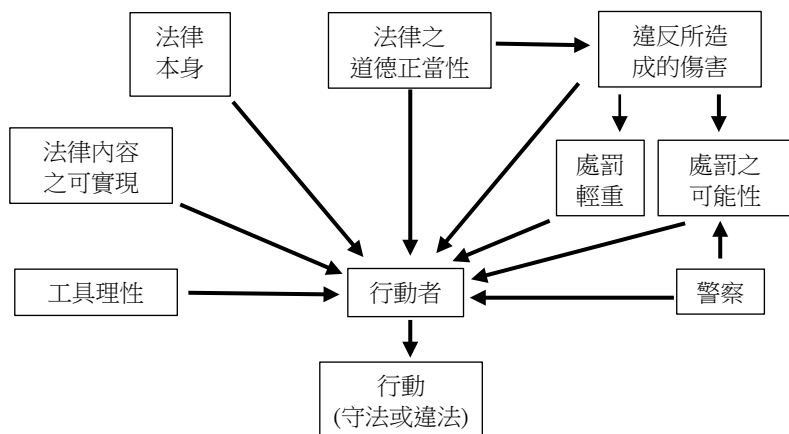


圖 1 遵守法律的初步因果圖

²⁸ 筆者曾在中央研究院民族所一場演講中提出此因果圖式的簡單版本。而郭佩宜指出，這個圖像是非常主體理性運作的特色，具有新自由主義 (neo-liberalism) 的特色。筆者也認為，這種以主體為中心的觀念就是西方法律體系的特色 (顏厥安, 2001: 3-7)，但不一定與新自由主義有概念上連接。筆者感謝那場演講中，容邵武、郭佩宜、鄭瑋寧與張仁和的相關意見交換。

行動者要做出的決定是遵循法律 (conformity to law) 或是不遵循法律 (nonconformity to law)，影響行動者的因素有以下幾個可能變項：行動者遵守法律要達成的目的是什麼、法律內容的可實現程度、法律本身、法律內容的道德正當性、違反法律造成的傷害、處罰輕重、受處罰的可能性、決策時警察是否出現等。²⁹

上圖有幾個需要注意的地方：第一，此因果結構所適用的對象為一般公民，而排除適用與執行法律的官員。第二，有八個變項會影響行動者，其中當變項間沒有邊 (edge) 聯繫時，假定這些變項是彼此獨立的。例如，法律內容之可實現程度跟法律內容之道德正當性沒有關係。³⁰ 第三，八個變項到行動者的箭頭意味著行動者會考量這些變項內的資訊，例如「工具理性→行動者」，意味著行動者會考量自身目的之達成與法律的關係為何。在這個因果圖中，邊的長短與粗細並未反映出其因果影響力。第四，這個因果圖並非是完整的因果結構圖，因為變項間需要更精確的因果影響力的表述方式，筆者因為篇幅原因並未處理。但是筆者相信，這個因果圖足以說明我們在守法與否的一般性想像。例如，在處罰可能性與警察兩個變項中可以見到，警察出現與否本身會直接影響行動者是否遵守法律，而警察出現這個事實也會間接影響到行動者對於受到處罰的判斷，因此警察這個變項是透過兩條路徑影響行動者的遵守法律與

²⁹ 上述八個變項的界定如下：(1) 工具理性是指，行動者自身所要達成之目的與手段的考量。(2) 法律內容的可實現性是指，法律所要求或禁止之事的難易程度。(3) 法律本身是指，該規範作為法律本身對於行動者的影響。(4) 法律之道德正當性是指，法律內容的道德重量。(5) 違反所造成的傷害是指，若違反法律的要求，會造成的道德傷害本身，例如違背殺人罪時，所造成的傷害是生命利益的喪失。(6) 處罰輕重是指，法律所設定的處罰程度或類型。(7) 處罰可能性是指，行為者對於自身違反法律遭受到制裁的可能性。(8) 警察是指，行動者在決定守法與否時，警察因素的考量。

³⁰ 當然這些預設 (assumptions) 是可以被挑戰與檢視的，但這就是因果圖的好處，我們可以清晰地看見推論的預設是什麼。

否的判斷。行動者對於處罰可能性的認知除了受到警察的影響之外，還會受到違法所造成的傷害的影響，這表示，如果傷害越大，行動者會認為受到處罰的可能性越高。³¹ 上述因果圖有個優點是，當 ALI 在運用時，可以較為清晰地看見 ALI 要介入哪一個變項以及該變項造成的影響是什麼。而運用決策模型與侷限模型的 ALI 在干預的變項上也會有所差異。

決策模型主要是影響行動者的偏好，而侷限模型是影響行動者的可實現選項。而不論是偏好或是行動的可實現選項都是透過「某一項法律規則的適用」為前提。換言之，行動者必須承認，ALI 正在適用某項規則，且承認行動者的當前處境正是 ALI 所適用的某項規則的範圍之中。這兩個承認，缺一不可。

運用決策模型的 ALI 可能採取以下方式改變行動者偏好。例如，它可以告知行動者該路段的速限與其判斷依據（法律的道德正當性）、公布其他行動者在該路段的平均速度（法律內容的可實現程度）、或是告知行動者即使提高速度對於目的地之抵達亦沒有太大助益（介入工具理性的考量）等等，雖然此類的介入不能擔保行動者必然遵守法律規則，但是可能增加那些變項對於行動者遵守法律的因果影響力。決策模型的 ALI 也可能選擇介入「處罰可能性」跟「警察」這兩者，例如它可以告知行動者前方有警察，或是該路段的告發比例等等，例如 100 位違規者有 80 位會被處罰。

採用侷限模型的 ALI 主要目的就是讓行動者欠缺選擇，或是不遵循法律成為不可實現的選項，因此要介入的變項，相對而言比較少。最簡單的方式就是透過高效率執法系統讓行動者未來的自己所具有的選項逐漸喪失，這意味著讓規則遵循的行為變成是強制性

³¹ 這也符合我們的直覺，因為我們有時會說：「你造成的傷害這麼嚴重你覺得你逃得掉處罰嗎？」

的行為。³² 而**要讓行動的可選擇性喪失最直接的做法就是提升執法效率**。因此，它可以選擇介入「處罰輕重」、「處罰可能性」跟「警察」三個變項，藉以加強處罰可能性對於行動者的因果影響力。在處罰輕重上，採用侷限模型的 ALI 可能可以告知行動者超速的罰單金額，在處罰可能性上，ALI 可以告知系統會自動偵測超速行為等等。

因此，可以見到如果重視高效率執法機制對於一般公民的行動管制，那麼在 ALI 的應用上其實很容易偏向侷限模型，因為 ALI 就是用來提高「惡害實現」的或然率，藉以改變行動者對於行動的選擇。³³ 另外需要注意的是，ALI 並未直接介入行動者的行動，如果是直接介入行動，那麼 ALI 就是物理性強制 (physical coercion)，而不是心理上的強制。³⁴ 然而，即便是惡害的通知，例如通知行動者「該當於某個違法且制裁實現將隨著而來」此一事實，它其實可能具有與心理上強制的相同效果，進而改變行動者的實現選項。

(三) 小結：如何對於上述分析進行規範性反省？

到目前為止，筆者針對 ALI 進行概念性分析與探討。然而，仍需要面對規範性課題，亦即 ALI **是否應該被使用**在法律體系的運作之中？現實上，ALI 在法律的許多領域中已經被大量運用了，但是對於其運用的規範性反省仍有其必要，因為它可以協助我們畫出應用的界線。

³² Shapiro 也是這樣認為的，請參考 Shapiro (1998: 51)。

³³ 關於強制的概念，請參考 Anderson (2011)。

³⁴ 感謝審查人指出物理上強制跟心理上強制的差異。筆者在初期並未仔細思考這兩者的差別，但誠如審查人指出的，惡害的通知具有和物理上強制相同的「心理」強制作用，此為更加精確的用法。

第三部分將承接著第二部分的分析進一步討論 ALI 與法治的緊張關係何在。以下先簡短勾勒筆者的觀點：法治與 ALI 的緊張關係表現在兩個部分：第一，多數的法治理論希望降低意志的恣意性，但是它們並未否定法治的核心價值在於讓受法律所規制的行動者得以成為自主與負責的行動主體，在這個意義上法治理論會要求將關鍵事項的判斷權限交給人來做，而非機器。第二，高效率執法的 ALI 會影響行動者的道德判斷自主性。

參、依法而治的主體意涵

法治 (rule of law) 是個豐富的概念，通常對比於「人治」。人治就是國家權力的行使是依照統治者的意志，而法治是指治理 (或統治) 是依據法的理性。根據 Gianlugi Palombella (2010: 10) 的界定，法治的理念在下述情況中可以獲得滿足：由人所握有的統治權力會被其他法律規則與制度所制衡，藉以避免人對於法律的創造與其內容握有壟斷性權力。將法治與人治區分開來之後，可以更細部區分法治跟「藉法而治」的差異，後者僅將法律視為統治者的工具，卻不在意法律的理念是什麼。³⁵

³⁵ 審查人特別指出藉法而治與法治的關係為何？Brian Tamanaha (2004) 在《論法治》(On the Rule of Law) 一書中將藉法而治放在法治的形式理論最薄的版本位置。他指出，藉法而治的核心是「法律是政府行動的工具」(law as instrument of government action)，但政府不受法律的限制。Tamanaha 即使將藉法而治放在形式法治的最薄版本，但是他也不認為西方法學者會完全將藉法而治與法治等同起來，因為藉法而治無法說明為何要對於政府的行動具有法律的限制，而對於政府行動具有法律限制是法治的核心。感謝審查人此處的協助釐清。的確，藉法而治的特徵是「不管政府做什麼，它應該要以法律為之」，但是這並未課與政府任何限制，它僅是說要以法律的形式為之。因此，Tamanaha 雖將藉法而治放在形式版本的一環，但是他並不同意以藉法而治來界定法治。此外，藉法而治甚至也未說到「法律應該要具備的性質」是什麼。

一般來說法治訴求的對象有兩個：第一是國家公權力。第二是一般公民。因此，Brian Tamanaha 說：「法治，在其核心之中要求政府官員以及公民必須要受法律的拘束，以及其行動必須與法律具有一致性」(2007: 3)。他認為，法治具有兩個重要功能。第一個功能是針對政府官員課與法律限制。第二個功能是維持秩序以及協調公民間的行動。就前者而言，仍然是法治的主要任務，因此 Tamanaha 主張，對於國家公權力之限制，包含要求政府官員服從既存的法律，以及針對立法權力課與法律限制。

要在本文中深入處理法治的概念幾乎是不可能的任務，因此筆者先說明預計探討方式。本文將透過兩個層次來架構起對於 ALI 的反省。³⁶ 首先，本文區分法治的概念 (concept of the rule of law) 與法治的概念觀 (conceptions of the rule of law)。³⁷ 概念觀主要是針對概念的內涵以及核心問題進行闡述以及說明，也因此，我們可以透過這組區分理解不同法治理論的側重之點。在這個意義下，Fallon (1997) 對於四種法治概念觀的整理可以作為思考的出發點。其次，筆者將指出四種法治概念觀的核心內涵是依法而治的主體價值。這個核心內涵不是指四種法治概念觀對於具體爭議或是理論建構有一樣的發展，而是指四種法治概念觀有共同的人文價值蘊含，它們認為受法律所拘束的行動主體具有一種道德自主判斷空間。筆

³⁶ 中文法學界中，已有專門對於理論家的法治概念與概念觀之研究文章，請參考莊世同 (2017)。莊世同的文章是對於海耶克的深度探討。若是對於法治概念的區分，顏厥安曾在〈論民主抵抗與法治〉(“On Democratic Resistance and the Rule of Law”)一文中探討 Tamanaha 的對於法治的類型區分。請參考 Yen (2015)。

³⁷ 除了採取概念與概念觀的區分，另一種方式可以是 Tamanaha (2004) 在《論法治》一書中所採取的區分：形式意義與實質意義的法律類型。筆者對於哪一種區分方式並沒有特別的偏好，採用概念與概念觀的區分目的在於概念觀的闡述，可以讓我們看到法治概念中更深層的價值取向。關於概念與概念觀的差異，請參考 Rawls (1971: 5)。

者在此所省略的是以單一理論家的法治與其概念觀作為探討對象，此種方式固然有重要助益，但是由於筆者的問題意識是反省 ALI 與法治的緊張關係，因此便必須依賴於 Fallon 的概念觀說明作為發展基礎，而不著重於單一理論家的深入探究。

一、四種法治概念觀

Fallon 首先指出，法治的概念是由以下幾個基本要素所組成的：

1. 法律規則、標準或原則有能力去指引人們的行動，人們必須能夠理解法律的要求是什麼。
2. 法治具有實效性，這意味著法律應該要能夠真正地指引人民行動。
3. 法治具有穩定性要素 (stability)，法律必須具有合理的穩定性 (reasonably stable)，如此一來才能歷時性地促進合作 (facilitate coordination) 與實現計畫。
4. 法治就是法的至高性 (Rule of Law is the supremacy of legal authority)，法律應該要統治官員與一般公民。
5. 法治意味著公正司法 (impartial justice) 的實現，法院應該要採用公平程序等等 (1997: 8-9)。

Fallon (1997: 9-10) 指出，下述四個原因造就了對於法治概念的爭執：第一，五個要素本身具有模糊性 (vague)，例如合理穩定的合理本身就是模糊的概念。第二，對於哪個要素具有首要重要性或是衡量其重要性的標準具有爭論。第三，某個法律體系是否有符合法治的理念本身是個程度問題。第四，討論法治不可能不討論法律。在這裡面最關鍵的問題之一是，適用法律的人 (例如法官) 是否一定受到法律的拘束 (must be ruled by law)? 若認為法官需受法律拘

束，法官考量制定法律之人的意圖時，是否就滿足這個要求？³⁸ 因此較佳方式是，梳理出不同的法治概念觀，藉以釐清不同概念觀的側重之處。³⁹

因此，Fallon (1997: 10) 進一步闡述四種法治概念觀 (ideal-typical conceptions of the Rule of Law)：歷史理念型 (Historicist Ideal Type)、形式理念型 (Formalist Ideal Type)、法律程序理念型 (Legal Process Ideal Type) 以及實質理念型 (Substantive Ideal Type)。這四種概念觀雖然有其具體的理論加以支持，但是並不是其理論的精確再述，而是探詢相類似理論中的共同要素而得出的。

歷史理念型將法治與由正當立法權威所制定的規範結合起來，主張正當立法權威所制定的規範（以規則方式呈現）優先於法律在

³⁸ 審查人提出，Fallon 是否支持「藉法而治」是法治的最低條件。非常感謝審查人此一細緻的問題。在此，Fallon 並未明確談到藉法而治是法治的最低條件，但是根據 Fallon 的論證方式，我認為他不會主張「藉法而治」是法治的最低條件。原因如下：Fallon 所提出的四個法治概念觀，雖然他本身偏向第三與第四種，但是這四種法治概念觀都主張官員必須受到法律的拘束，這意味著官員不能「藉法而治」，而必須「依法而治」。因此，Fallon 不會主張官員不須受到法律拘束。其次，第三種與第四種法治觀所強調的是法官對於國家權力行使的審查，因此相較於歷史理念型與形式理念型，法律程序學派與實質理念型更強調政治道德的重要性。從這個角度來看，「法律形式」的重要性是次要的。

³⁹ 有很多理論家探討法治的觀念，為何需要以 Fallon 為對象。例如 Tamanaha (2004) 在《論法治》一書區分了形式觀點與實質觀點也是一個可以採取的方式。由於本文並不是一篇專門探討法治的專論，但是仍需要針對研究對象素材之選擇進行說明。筆者認為 Tamanaha 的區分非常細緻，例如其區分不僅建立在形式 (formal) 與實質 (substantive) 之上，並進一步有薄的程度到厚的程度，例如所謂的藉法而治就是形式觀點中較薄的觀念，而在實質觀點中較薄的法治是堅持個人權利的保障。Tamanaha 的分類有其清晰性，但是在論述上較大的缺點是，不容易看出來不同類型的法治觀點的規範性側重點何在，因此容易讓人誤解為法律沒有本身的內在價值。相較之下，Fallon 明確地使用概念觀並且直接說明不同法治概念觀對於法律的正當性來源、法治的價值基礎以及影響進行具體說明。在這個意義上，運用 Fallon 的區分與說明更有助我們建立一個較為價值偏重的理解框架。

個案中的適用。換言之，歷史理念型關心的問題是，法律是否由正當法律權威所制定，進一步認為，法官應該受到法所支配 (ruled by law)，而非在個案裁判中創造法律 (Fallon, 1997: 11)。此外，這個概念觀的法律正當性 (legal legitimacy) 是跟過去正當立法權威的事實相連結，法律的正當性並不在於其內容，而是在於法律本身是由具有正當立法權威的立法者所創造出來的這個事實之上。歷史理念型的法治觀主張，立法過程滿足嚴格的民主課責的規範是法律具有正當權威的重要來源。⁴⁰

形式理念型認為法治的理念是，法律作為一種「規則」(rule) 是具有重要意義的。法「治」(“rule” of law) 要求依法而治理，因此「規則」是法律重要的形式。「法律作為一種規則」(law as a form of rule) 優先於法律的適用，且是法律判斷妥當性的重要條件 (Fallon, 1997: 14)。形式理念型與歷史理念型最大的差異在於，歷史理念型著重法律規範的制定是與過去正當立法權威相聯繫，但是形式理念型強調法律作為規則的形式。⁴¹ 形式理念型另一個重要的實質設定是，它將人視為理性規畫者且可以極大化自己所想要達成的目標，因此人可以運用法律規則與他人合作或者達成自己的目的。規則的主要任務就是指引行動，一旦行動指引的功能完整，行動者就可以按照規則的指示達成自己的目標。需要注意的是，一般可能認為形式跟價值無關，然而這個想法在形式理念型的法治觀中是錯誤的。形式理念型的學者對於法律的形式與其所促成之價值有不同

⁴⁰ Fallon 將歷史理念型的法治觀跟憲法解釋中的原意主義相連接。雖然這可能有所爭論，但卻也精準地指出原意主義的法治觀念可能具有的核心要素。關於美國憲法解釋的原意主義與活生生憲法之爭是一個非常大的課題。關於美國憲法解釋理論的原意主義的法治理念的初步分析，請參考陳弘儒 (2021)。

⁴¹ 關於規則與規範的差異，Shapiro (2011) 曾指出規範可以有抽象與具體，但是規則具有普遍性。

的著重，例如 Fallon 指出，Scalia 認為法律作為一種規則其意義具有可讀性與清晰性 (clarity of laws) 等特點，狄驥 (Albert Venn Dicey)、海耶克 (Friedrich August von Hayek) 與羅爾斯則認為明確的規則是作為保障特定價值的必要條件，而 Lon Fuller (1997: 15-16) 則將規則與法的實效性相連結，雖然上述學者的法治理論有所差異，但是他們都著重法律作為一種規則形式的概念，可以促成某種善或價值的實現，因此所謂的形式不是無涉價值的。⁴²

相較於前兩種法治觀，法律程序理念型認為，法律若要符合法治的理念，有四個關鍵要件需要滿足：(一) 法律的制定與適用必須符合程序公平。(二) 法與合理性具有內在關聯。(三) 法律權威的來源與個案的法律判斷，這兩者間的關聯必須被合理闡述 (reasoned elaboration)。(四) 司法審查是確保立法行為跟行政決策之程序合理性與理性思辨的重要條件 (Fallon, 1997: 18)。可以見到這種法律程序理念型法律的正當性源自於更為實質的價值理念——程序公平性，並且賦予司法權更寬廣的裁判權力。

在法律程序理念型中，司法權不僅僅是要**合理地闡述**過去法律權威 (立法者) 與權利保障的實質關聯性，司法權更是行政權與立法權之權力正當行使的重要守護者 (Fallon, 1997: 18)。因此，法律

⁴² Fallon 在說明完形式理念型的意義之後，處理了兩種反對形式理念型的批判：法律唯實論與語言的文本開放性結構的批判。但這與本文問題意識不相關，因此筆者不進一步加以說明。筆者認為，Fallon 在此指出的諸多學者是否真的可以如此理解仍必須加以探討。例如，莊世同曾指出海耶克區分了立法之法與自由之法，進而指出英國的普通法傳統是一種基於「自發性秩序發展成為自由之法的範例」，透過法官造法的方式找出適應時空變遷的普遍性法則，在這個詮釋下，法官似乎是可以有較大的造法空間。但是，莊世同文章的脈絡是在探討海耶克對於自由之法及立法之法的區分，因此不必然意味著海耶克會支持法官一般而言有較大的造法空間。在立法之法下，法官的角色可能是奠基於憲政主義下調整或是修正立法的內容。關於這部分請參考莊世同 (2017: 673-675)。

程序理念型與法律程序學派 (legal process school) 有很大關係。⁴³ 法律程序理念型與形式理念型的主要差異在於，對於行政權以及司法權的權力行使條件不一樣。形式理念型強調法律作為一種先存在的規則，也因此行政權無法僅以其政策觀點行使國家權力且恣意地忽視規則的重要性，同時司法權也不能在欠缺法律規則的情況下進行裁判。然而，法律程序理念型認為法律的規範性來源在於其決策的妥適性，而在思考決策妥適性必須判斷哪一個機關才是作為特定決策之適合的決策機關。因此，法律程序學派理念型會著重思考行政、立法與司法之不同功能與意義，不以法律規則必須由立法通過為前提。此外，法律程序學派賦予司法權更大的實質審查權限，因為它認為合理闡述是法律 (司法) 程序的本質。⁴⁴

實質理念型認為，法治的規範性來自於法的可理解性 (intelligibility)，這種可理解性奠基在法律具有被合理所肯認的目的。當法律的內容被人們所理解時，將會產生道德忠誠，使得人們知悉法律到底要做什麼。Fallon (1997: 21-22) 自己認為，實質理念型本身其實是非常紛雜的概念觀，但是不同的理論都會認為，如果法律未滿足一定的道德正確性的話，那麼法律就牴觸了法治的理念。換言之，

⁴³ 法律程序學派主要是由 Henry Hart 與 Albert Sachs 所共同發展的。這個學派的發展具有具體美國法律發展的歷史，著重由政府不同部門的功能特色，而發展出動態的法律解釋理論，以及法律概念觀。其中最關鍵的想法是，合理闡述這個觀念，其認為司法權應該顯明法律素材中規範性命題。關於合理闡述的基本內容，可以進一步參考 Fallon (1994: 966)。關於法律程序學派的中文簡短介紹，請參考陳起行 (2002: 1-27)。

⁴⁴ Fallon (1997: 21) 認為，法律程序理念型被批評為，無法解決法律的不確定性問題，因為即使大家對於司法權的運作以及程序價值有共識，但是仍無法解決價值歧見的問題。Fallon 認為這的確是一個重要問題，但卻是一個獨立的問題 (separate question)，他舉了一個反證，他說美國現今有很多關於正義的不同意見與深刻歧見存在，但是法律體系仍然運作的還可以。更精確地說，傳統法律程序所要求的合理闡述仍舊在美國法律體系持續運作。

實質理念型的法治觀明確地擁抱了某種特定的道德價值，作為判定法律內容的正確性之判準。Fallon 舉例指出，例如 Ronald Dworkin、Richard Epstein 及 Frank Michelman 的理論都可算是實質理念型。實質理念型有兩個特色：第一，透過法的可理解性，實質理念型的法治觀加強了法律主體的守法動機，因為他對於法律的理解增加了對於法律的忠誠。第二，實質理念型強調法律與實質道德的可接受性的關聯性。⁴⁵

我們可以將上述的四種法治觀以表 1 呈現出來。

⁴⁵ 審查人提出，Scott Shapiro 的實踐差異理論可能會屬於 Fallon 的哪一種法治觀呢？非常感謝審查人此一提問，筆者原先並沒有如此思考，原本透過 Shapiro 的分析是要說明法律影響行動的方式（決策模型與侷限模型），經由審查人提醒，筆者發現法實證主義的法治觀確實是一個有待探索的課題。因此以下筆者的說明是非常初步的，試圖透過 Shapiro 的一般性法理論理解他的主張最適合被放置到 Fallon 哪一種法治觀。

一般認為，法實證主義較傾向法官應該受到實證法的拘束，或者至少在法律規則的核心事例之中是如此。但是 Shapiro (1998: 44, 2011: 254-256) 主張，適當理解下的法實證主義，不一定會認為法院僅有較為限制性的角色，因為法官是一種地域性的制度性角色，它的角色並無法一概而論。實證法是否限制法官的法律推理與內容取決在各個地域的法律。因此，Shapiro 應該不反對自身的理論與實質理念型法治觀相結合，給予法官較大的論證空間，允許法官使用更多的道德之理據。但是筆者認為，從實踐差異命題來看，最適宜說明 Shapiro 理論的法治觀應該是形式理念型，因為形式理念型著重法律具有可預測性以及行動協調的功能，這些目標都與 Shapiro 在實踐差異命題或是法律的計畫理論 (the Planning Theory of Law) 相吻合。此外，Shapiro 在探討實踐差異的文章中，有個論爭脈絡是在處理柔性法實證主義的不足之處，在這個脈絡中 Shapiro 認為，如果立法者通過一條具有道德內涵的法律（例如，法律規定應該遵守允諾），但因為法律本身具有被廢止的可能，因此我們可以設想當一條具有道德內涵的法律被廢止之後，這個事實會影響法官或行動者的實踐推理，因此具有實踐差異。從這個角度來看，Shapiro 可能會堅持即使法官在判決中使用到道德理據，但是在關鍵之處指引法官行動的規範仍不是道德規範，而是法律規範。因此，Shapiro 可能不會將自身允諾到實質法治觀的立場，然而即使是法律程序理念型，他依舊會堅持那些可能被廢止的實證法才是具有創造實踐差異的法律。上述是關於 Shapiro 對於柔性法實證主義的批判。

表 1 四種法治觀的比較

概念觀	正當性來源	拘束國家權力程度			法治的規範性價值來源
		行政	立法	司法	
歷史 理念型	民主立法程序	強	弱	強	(薄) 意義下的民主
形式 理念型	法律作為 規則的形式	強	弱	強	法律可預測性 以及行動協調功能
法律程序 理念型	法律判斷之 合理性	弱	強	弱	法律判決的理性論證
實質 理念型	法律判斷之 道德正確性	強	強	弱	法的道德性質

資料來源：作者自行整理。

從上表中可以看見四種法治觀的彼此差異。此外，我們可以看到民主正當性的堅持可以屬於歷史理念型的法治觀，這與 Tamanaha 在形式觀點的法治觀有異曲同工之妙。Tamanaha (2004: 101) 指出形式觀點的法治觀有一種是強調民主與合法性，另一種是形式合法性 (formal legality)。所謂民主與合法性是指法律設立過程是經由人民同意而產生的，但是 Tamanaha 提醒我們此種結合合法性與民主的形式法治觀點，不意味著法律的內容都是正當的，因此法律有可能經由民主程序產生，但是對於法律安定性或是可預測性 (這些形式法治價值) 具有破壞效果。

二、依法而治的主體意義作為法治的核心價值

上述四種法治概念觀各有不同的價值取向，對於法律的意義也有所不同，但是我們可以從一個關鍵點出發：即使法治的許多內容僅具有形式意義或工具價值，但是法治之所以有價值並不一定僅基

於**工具性的理由**。例如，法治要求法律具有普遍性、清晰性，以及原則上不溯及既往等等，但這不意味著法治對於政治社群的成員僅具有形式價值，**相反地落實法治理念的社會往往可以建立一種互惠的連結，減少統治者與被統治者的權力不對等**。⁴⁶ 因此，在關注法治主要是針對國家權力運作的限制時，同時也必須要注意到法治也要求受法律影響的公民們應該尊重法律以及遵守法律。結合這兩者，我們可以問下述基本問題：**為何我們有理由接受法治作為一個重要的價值？**

莊世同從 Tamanaha 將其界定在形式意義的法治觀的反省出發，進一步探討海耶克的消極自由觀時指出：「海耶克對不自由的闡釋相當清楚：唯有在完全依照他人意志而行動的情況下，才是受強制而不自由的。反過來說，一個人只要還存有依自身目的行動的空間或餘地，即使他可選擇的選項極少或不確定，也不能說受強制而不自由」(2017: 229, 684, 678)。雖然莊世同文章著重對於海耶克的法治觀進行批判性探討，論證海耶克的「新羅馬理論的消極自由觀」之發展，但是筆者想要透過莊世同之研究指出，海耶克的法治觀其實也關注於促進個人自主與尊嚴。⁴⁷

因此，筆者更進一步要指出，**法治的核心內涵就是維持個人的自主道德判斷**。法治要求社群成員必須尊重法律以及遵守法律，但是法治是以維持社群成員的個人自主判斷為其前提條件，用一種具象的說法，可以稱之為道德自主性的判斷空間。⁴⁸ 在這個判斷空間

⁴⁶ 這樣的想法來自於 Jeremy Waldron 的觀點。請參考 Waldron (2020)。

⁴⁷ 如莊世同提出「海耶克提出的三種法治屬性，與 Fuller 的主張相當接近，同時兩人都同意，藉由讓人們事先知悉已頒布的法律，進而得以規畫安排自己事務，法治確實可以促進個人自主和尊嚴」(2017: 678)。

⁴⁸ Shapiro 在〈權威〉(“Authority”)一文中，針對 Wolff 對於自主性與權威之吊詭提出的分析中，也使用理據的空間來說明道德自主性的意涵，並且運用獨立於內容與阻斷

之中，社群成員（或是行動者）將法律的內容、法律的概念、自身的實踐考量以及其他價值進行評判與價值排序，即便最終行動者做出的行為是一樣的（例如都是符合法律的行動），但是**行動者仍保有做出在判斷空間中價值位階以及保留改變其價值次序的可能性**。在這個意義上，可以說個人擁有自主行使道德判斷的能力。

在法哲學中，對於理據判斷空間的討論往往在權威與自主性的弔詭脈絡下開展，例如 Shapiro 從檢討 Robert Paul Wolff 的論點開始探討權威的意義。但是，我們可以從另外一個脈絡看到道德自主性之判斷空間的重要性，Fuller 在此有關鍵的啟發。Fuller (1969: 46-51, 63-79) 在《法律的道德性》(*The Morality of Law*) 一書中闡述了他的法治理論。⁴⁹ 他基本上認為，法律必須要有**普遍性 (generality of law)**，其根本的意義在於，**法律必定是規則，法律若要規制人們的行動就必須以規則的特性而來**，且法律必須要被**頒布 (promulgation)**。法律，原則上還不能是**回溯性法律 (retroactive law)**，或稱之為**溯及既往**。⁵⁰ 此外，法律也必須具有**清晰性**。Fuller 認為，清晰性是合法性原則最本質性的要件，但是**法律具有清晰性**

性理據 (content-independent and peremptory reason; CIP reason) 來說明權威與自主性的矛盾可能性衝突。筆者在此並未要進入法哲學中理據論的討論，但是筆者的道德自主性的判斷空間跟 Shapiro 談到的理據空間 (space of reasons) 很相近 (Shapiro, 2002: 390)。

⁴⁹ 必須要先說明的是，Fuller 本身在書中沒有特別著重使用法治 (rule of law) 這個詞，相反地他使用的是「法律的內在道德」(the inner morality of law)。雖然 Fuller 在《法律的道德性》一書中較少使用法治一詞，但是究其實質他的八項法律內在道德原則就是法治的基本內涵。關於內在道德一詞的說明，請參考 (Fuller, 1969: 42)。

⁵⁰ 在《法律的道德性》一書中，Fuller 認為，溯及既往的法律不是必然有害的。理由很簡單，法律的內在道德有很多要件，法律作為一種指引行動的事業可能會在達成某一項要件時，犧牲另外一項。也因此法律的內在道德會受到嚴重的破壞，而當此種情形出現時，溯及既往的法律就是作為矯正措施而不可或缺的。

不表示法律不能使用概括用語，而是若要使用模糊的概念最好是要使用符合在常人生活中所建立起的判斷常識。法律也必須**避免矛盾 (contradictions in the laws)**。法律也**不得要求做出「不可能」之事 (impossibility)**。**穩定性 (constancy)** 要求法律不應該過於頻繁地更動。最後，法律必須要有**官方行動的一致性 (congruence between official action and declared rule)**。⁵¹

Fuller 的八項法律的內在道德往往被認為不具有內存道德價值的要素。的確 Fuller 也認為，兩種相反內涵的法律規範可能都滿足他所提的八項法律內在道德的要求。然而，**筆者認為關鍵不是八個要件是什麼 (what)，而是 Fuller 如何 (how) 定位這八個要素。** Fuller 一再強調法律是一種「將人類行為為臣屬於規則治理之下的事業」(1969: 153, 162)，因此法律內在道德是中立的，具備內在道德的法律可以用來追求多樣的目標。⁵² 但是，Fuller 在《法律的道德性》一書中，最關鍵的價值主張是：**法律的內在道德雖然可以用來支持不同道德立場的主張，但是內在道德在一個關鍵議題上是具有實質的目標：法律的內在道德的要求是用來達成更廣泛的人生目標。**這意味著擁有法律內在道德的法律體系背後有個實質的法律主體的圖像，**這些內在道德的實質目的就在於協助受法律所規制的人民可以朝向這個圖像前進，我稱之為「依法而治的主體意義」(meaning of subject under the rule of law)。**

Fuller 以下這段話值得我們進一步分析：

⁵¹ Congruence 在國家教育研究院翻譯詞典中譯為「一致性」(國家教育研究院，2012)，筆者有時稱之為「幅合」。

⁵² Fuller 對於內在道德中立性的表述可能會讓人誤以為他的內在道德是工具價值而已，因為可以用來服務於不同的正義觀或是道德主張。但是，這種觀點很可能誤解了 Fuller 的表述，因為他同時認為有些實質道德目標是會減損法律的內在道德的 (1969:153)。

現在來談談最重要的觀點：普遍來看，對於法律道德性的要求之遵從可以服務於更寬廣的人生目標。這個觀點是基於，隱含在法律內在道德性之中的人性觀點。我持續地觀察到，法律的道德性在廣泛道德議題上可說是中立的。[但] 它對於人性的觀點無法是中立的。若欲將人類的行為臣屬在規則的治理之下，這必然是允諾以下觀點：人是——或者終將成為——負責任的行動者，其有能力理解與遵守規則，以及得以對自身錯誤負責。(1969: 162)

Fuller 批評的對象是將法律視為可以改變人類行動的媒介，從而僅將人視為一種針對外在刺激給予反應的想法。他提到了心理學家 B. F. Skinner 的經典觀點，Fuller 反對個人的自由意志與責任可以僅憑刺激與反應而表現出來。⁵³ Fuller 對於 Skinner 的評價是否正確當然可以進一步檢視，但這並未取消他的法治觀具有內在價值的可能性，因為他所提到的八項法律的內在道德就是要讓人得以成為具有能動性 (active)、負責任 (responsible)，且可針對自身錯誤進行回覆 (answerable) 的法律主體觀念。要實現這個目標就不能夠僅將法律當成一種對於行動調控的工具，換言之不能將法律視為一種刺激，而人之行動是反應，因此不能僅將法律視為一種可作用在人之行為的機制。如果行動者要成為具有能動且負責任的法律主體，那麼法律體系必須給予行動者在思索法律對其行動影響上具有充足的自主判斷空間。在這個意義上個人的自主判斷空間與法治是具有內在聯繫，這也是為何法律具有協調社群成員的功能。

調控學將訊息與行動或決策連繫起來，而理解 (understanding) 對於行動者而言並非必要。Fuller 的理論提醒我們，法律協調社群

⁵³ Shoshana Zuboff (2019: 376-397) 也在《監控資本主義時代》(*The Age of Surveillance Capitalism*) 一書中，指出 Skinner 的刺激-反應觀點將在機器控制力量下獲得很大程度的勝利。

成員的行動不能僅止於法律調控學 (cybernetics) 面向而已，我們必須要看到行動者**如何理解**法律而做出相對應的行動脈絡。Mireille Hildebrandt 在這點看得很清楚，她說：「調控學一詞源自希臘文中的『導航』、『引導』與『治理』，這些意涵也因此將資訊理論連結到法與管制的課題。可以想見，偵測訊息的能力——在現代這些能力被理解為預測某些潛藏的行為模式或價值——在調控一群人的行為上有極大的潛力」(2016: 18)。將法律視為調控社會成員之手段的觀點是一種外部觀點，不需要納入行動者主觀上對於法律內容是否理解，而是將法律視為刺激，行為視為反應進行管制。但如果我們認為法治中的主體意涵是有意義的，就不可能僅採取此種外部觀點的管制方式，相反地，必須重視行動者的道德自主判斷空間，將行動者對於法律的理解納入其行動決策的理據模型之中。

如果上述對於依法而治的主體意涵是可接受的話，下一個問題是，什麼是行動者的道德自主判斷空間。筆者使用「道德」一詞在於這是涉及行動者應該如何做的實踐判斷，也因此我將省略道德一詞的解釋。自主判斷空間的「空間」可以視為隱喻式用法，因此關鍵問題是何謂「自主判斷」？雖然筆者在此無法針對這個概念進行詳細說明，但是所謂「自主判斷」的意涵基本上是指，行動者治理自身的那些因素，例如思慮、慾望、條件或特質並不僅是由外而強加於行動者身上的，而是來自於行動者自身真實的一部份。⁵⁴ 我們可以給予自主判斷空間提出一個操作型界定：**自主判斷空間是指行動者對於涉及到法律的理據或考量之重量與優先次序具有相當決定可能，換言之行動者對於涉及到法律的理據何者較為重要或是何者具有優先性是由自身所決定的，而非由外在對象所強加的。**在這

⁵⁴ 這部分的定義與討論請參考 Christman (2020)

個意義上，筆者將自主判斷空間視為一種對於行動判斷的理據排序與衡量過程。

法治要求法律必須讓行動者在涉及法律行動之選擇上具有自主判斷空間，這意味著行動者在行動選擇上保有理據的優先次序與輕重的衡量空間。我們曾經在貳、二、(二) 之中勾勒出遵守法律的因果結構圖，指出八個影響行動者的變項，我們也可以將這些變項視為行動者遵守法律與否的理由。

在理據的次序與輕重上，一般而言法律較少給予一體適用 (one-size-fit-all) 的答案，它讓行動者自身去考量理據的優先性以及對於其行動的終局重要性，行動者思考的是「我應該如何行動的判斷」，因此法律主體的意義便在此顯示出來：行動者會「聽到」許多理據所給出的資訊，因此行動者是處於「衝突」的位置，所有理據的箭頭都指向行動者。若要讓行動者有獨立自主的判斷空間時，這意味著行動者必須自行判斷諸多理據的對於自身的因果影響力。行動者自身如何判斷理據對其決策的影響程度，原則上我們是無法知悉的，除非進行扎實的實證調查。

從圖 2 可以見到捍衛法治就是捍衛一種主體法學的觀點。法治原則上是讓何種理據置於優先考量、理據有多重要等交給行動者決定。這表示，法治要求國家在運用手段進行介入行動者行動決策時，仍必須維持不同理據到行動者的因果影響力。⁵⁵

⁵⁵ 簡單來說，捍衛法治就是捍衛一種主體法學的觀點。關於主體法學的觀點，可以參考顏厥安，他談到主體法學最關鍵之處在於主體的性質時指出：「我將此一主體性質描寫為：一個擁有思考認知能力的抽象虛空主體」(2001: 4-5)。

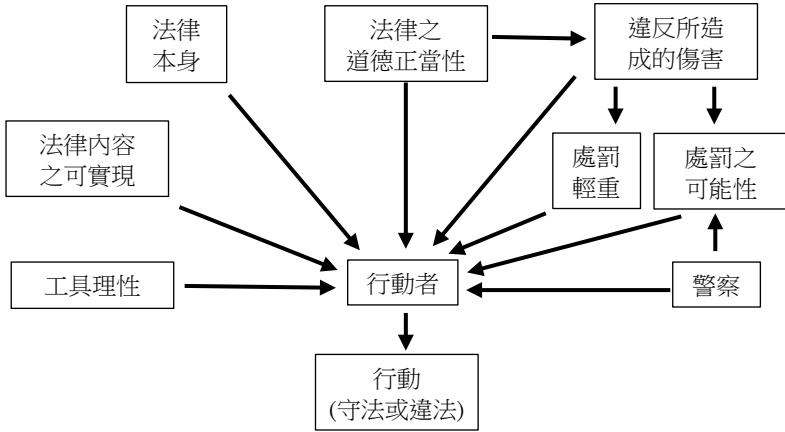


圖 2 依法而治的主體意涵

我們在貳、二、(二) 之中運用因果圖表示了採用決策模型跟侷限模型的 ALI 對於要介入的變項會有所差異。而有種反對意見認為，就算法治是捍衛主體之治，但是為了維持法治，對個人的自主判斷空間而反對 ALI 可能是一種錯誤，因為我們不會反對駕駛人自行裝設車用的速限警示系統（提醒速限與告知偵測位置）。⁵⁶

針對此種反對意見，筆者認為「行動者（一般公民）自行採納速限警示」跟「國家採納超速偵測系統」的因果圖是不一樣的。當行動者自行使用速限警示系統時，他是在「處罰可能性」到「行動者」之間，增加了一個中介變項（警示系統）。這個中介變項讓行動者對於速限之控制取決在警示系統，截斷了處罰可能性對於行動者的直接影響。此外，它也截斷了警察透過處罰可能性到行動者的影

⁵⁶ 這個反對意見是審查人所提出的，非常感謝審查人的這個反對意見，讓筆者有機會透過因果圖釐清「自己裝設速限提醒」與「非自己裝設的差異」。

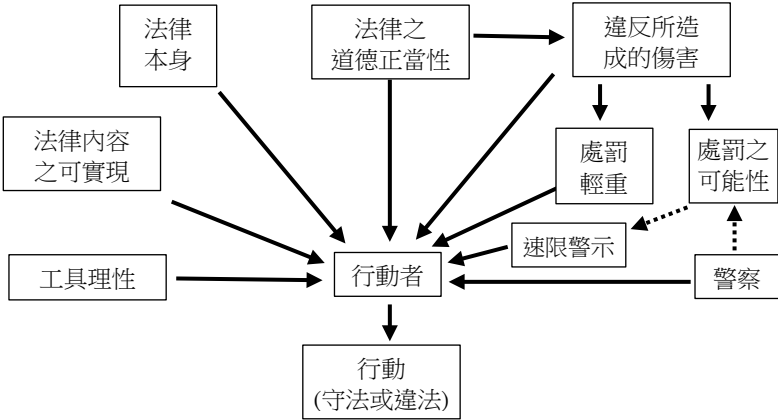


圖 3 自行裝設速限警示系統因果圖

響。⁵⁷ 例如警示系統說：「您的車速 110 公里，沒有超過限速」，行動者不需在意「警察是否存在」這個事實，行動者也不受到因為看見警察，所以可能受到處罰的影響。圖 3 表示自行裝設警示系統的因果圖，虛線表示原本因果影響力的減弱或消失。

然而，若是國家採納自動測速系統，這個時候國家所干預的方式不是在處罰可能性與行動者中間增加一個中介變項，而是增加一個變項藉以提高處罰可能性對於行動者的影響。國家希望透過測速系統的存在，提高處罰可能性，藉著提高處罰可能性，影響行動者的行動，其因果結構圖呈現為圖 4 (國家企圖干預的路徑為灰色粗線)。

因此，行動者自行裝設速限警示系統跟國家裝設自動測速系統，兩者的因果結構不一樣。前者是行動者自行改變要遵守法律與

⁵⁷ 具體例子就是，當駕駛人一裝上警示系統時，他可能就不需要在意警察存在或不存在。

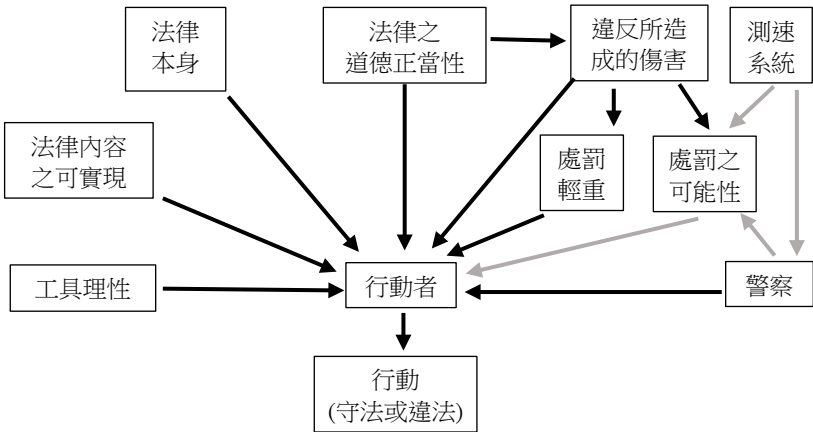


圖 4 國家裝設自動測速系統

否的因果結構圖，從而行動者願意讓其遵守法律與否的行動依賴於速限警示系統，而截斷了處罰可能性與警察對於行動者的影響。但後者是國家加強了處罰的可能性。此外，筆者也不否認行動者自行裝設速限警示系統也是一種判斷權限的移轉，但是行動者原本就可以將判斷權限移轉出去。甚至可以說，基於對於個人自主判斷空間的尊重，行動者可以自主地移轉判斷權限。

最後，我們需要仔細區分跟上圖類似的介入方式，但是完全不一樣的因果圖。假設國家所設置的不僅是自動測速系統，其所設置的是一套全自動的執法系統（辨識車牌、測速、開罰單、違規通知），該套系統的測速方式不是定點測速，而是區間測速，因為是機器不會累，所以是 24 小時運作。那麼這種介入方式跟使用自動測速是完全不一樣的。此種的因果圖會以圖 5 方式展現出來（國家干預路徑為灰色粗線）。

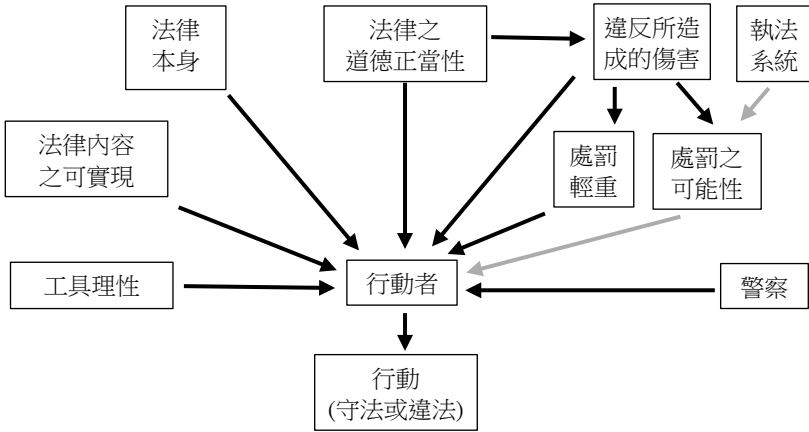


圖 5 全自動測速執法系統

由於是全自動執法系統，因此它與警察之間沒有聯繫。既然警察不開單，因此警察對於處罰之可能性也沒有聯繫，對於處罰可能性有影響的就是執法系統的運作。而行動者所認知到的因果路徑是直接從「執法系統→處罰之可能性→行動者」這一條路徑。警察對於行動者雖然有影響，但可能影響不如執法系統透過處罰可能性對於行動者的影響來得大，因為行動者知道警察不開單。⁵⁸ 這個因果圖呈現出，如果國家要透過執法系統來調節一般公民的行為的話，那麼她只需要更動「執法系統」跟「處罰可能性」這兩個變項就好。例如她可以更強化執法系統的能力，一天 24 小時執法、全路段偵測超速與否等等，這個時候即使其他變項的影響力不變，但是「執

⁵⁸ 這個說法乍看有點反直覺，因為警察為何不開單呢？這個例子是設想全自動的執法系統，但是日常生活中我們也有類似體驗。例如，警察抓闖紅燈時，眼前如果有人被抓的話，警察對於後面的車就沒有太多嚇阻力，因為警察只有一個人，一次僅能臨檢一輛車。

法系統→處罰可能性→行動者」的因果影響力會大幅上升。在這個意義上，行動者對於各式理據的衡量會受到影響。

肆、對於 ALI 的規範性反省

一、對於權限分配原則的無差別原則的批判

從上述分析可以見到，法治與 ALI 具有內在緊張關係。這個緊張關係存在於，AI 的設計原理就是以取代人類判斷或是任務執行為核心，當 AI 運用在法律之中，就是要將人的主體判斷給排除出去，這個發展是與法治相牴觸的。法治雖然要取消人的恣意統治，但不意味著要將人的判斷給完全排除掉，法治是要將人的判斷臣屬在法律的統治之中，它要讓統治行為的恣意性 (arbitrary) 降到最低，但是不取消人類作為判斷主體的可能性。法治要求執行法律事務的行動者 (法官、律師、檢察官等) 必須善用自身對法律的確信與認知，實踐自身的法律責任與要求。因此，以取代人類判斷的 ALI 恰恰與法治相違背。

然而，相容論者可能提出以下理由反駁，並認為 ALI 仍舊與法治相容不悖。以下我將逐一討論可能的反駁理由。

相容論者可能先有以下主張：「好，替代型的 ALI 確實不應該應用在法律之中。那麼在設計理念上使用輔助型的 ALI 就可以了。換言之，AI 讓人類具有特定事物的最終判斷權限，甚至人類可以有將系統關閉的可能，以展現其最終判斷權限主體的意義。」因此，讓我們從輔助型的 AI 作為前提出發，再一次思考法治與輔助型的 ALI 是否相容。

相容論可能先提出有限度的相容論證 (argument of limited compatibility)，主張「實質理念型」與「程序理念型」的法治觀確

實與 ALI 有緊張關係，但是只要抱持這些概念觀的人將 ALI 視為輔助性質，法律人是做出判斷權限的主體，兩者仍然可以**有限度地相容**。這個論證首先不否認 ALI 可能對於法律內既有的基本權利與價值的保障產生威脅，但是它認為即使如此，輔助式的 ALI 可以讓這個威脅降低，因為輔助式 ALI 仍將最終判斷權限交給法律人來做，而不是機器來做。

這種有限度的相容論證實際上是誤解了實質理念型與程序理念型所提出的關鍵問題。程序理念型著重的是法律的可測試性 (contestability)，透過司法論證程序經由雙方當事人思辨法律究竟要求什麼以及法律的權利義務之分配根據何在以達成法律判斷的合理性，這是合理闡述的重點。但是，ALI 的黑盒子性質以及設計機制本身具有的阻礙，往往阻礙了此類程序正義的實現，最著名的例子就是 Loomis 一案。⁵⁹ 即便是讓法律人擔負起最終判斷的主體，但是可能仍無法理解 AI 是「如何」做出特定決策。特別是，這個「如何」還有技術性的疑慮 (例如可解釋性的問題) 以及法律機制的阻礙 (例如既有的營業秘密保障機制等)。⁶⁰ 此外，資料驅動的 AI 蘊含的資料偏見 (bias) 問題也有待克服，隨著偏見問題所產生的歧視 (discrimination) 如何處理也仍然未知。如果著眼於系統應用後的大規模影響，那麼更需要謹慎小心，而不能僅是提出人亦有偏見與歧視等論據加以駁斥。⁶¹

相容論者可能提出另一個論證，我稱之為最低限度論證 (argument of the minimal compatibility)。這個最低限度論證主張，可以採取形式理念型或歷史理念型的法治觀，如此一來輔助型 ALI 就可

⁵⁹ 請參考 Liu et al. (2019), Whittaker et al. (2018)。

⁶⁰ 關於機器學習中的可解釋性問題，請參考 Gilpin et al. (2019), Lipton (2017)。

⁶¹ 請參考顏厥安 (2018) 關於 AI 影響的討論。

以跟法治相容。這個論證很有吸引力，因為它認為，較厚的法治觀本身就會有許多爭議出現，例如隱私與程序公平性等都是厚的概念 (thick concept)，需要結合經驗與價值的判斷。如果連法治觀內部都有爭議的話，那麼要求 ALI 符合此類法治觀是過於苛求的。因此，我們可以採用最低限度的要求，只要 ALI 符合形式理念型或歷史理念型就可以了。

形式理念型法治觀強調，法律作為一種事前公布的規則，其清晰性與公開性有助於行動者理性地擬定自身計畫，並且透過規則實現自身的目標。認為 ALI 與形式理念型相容的理由在於，ALI 在執法方面是幾乎全面地適用，也因此法律的落實上比由人而來的落實更普遍，沒有受到人類情緒的影響。且由於法律效果的普遍落實，更能讓行動者協調彼此的行動，因為他們會知道一旦違法就會受到處罰的機率大幅提昇。就歷史理念型的法治觀，強調法律必須由民主立法程序所產生，因此 ALI 只要將那些經由立法者產生的法律轉化成演算法或是程式碼就可以符合歷史理念型的要求。

我將在下一節仔細討論形式理念型與 ALI 的議題，在此先省思歷史理念型與 ALI 的關係。

根據 Hildebrandt 的區分，ALI 在概念上可以區分為兩種：採取傳統程式碼的 IFTTT (if this then that) 以及運用機器學習的資料驅動 (2018: 1-11)。Hildebrandt 將前者以決策數作為範例，但是在概念上也包含了著重法律推理或法律論證的計算機模型。設計者希望可以將法律規範轉換成程式碼，藉由著程式編寫的運作來執行法律任務，專家系統一般而言都是以這種方式進行的。與 IFTTT 相對的 AI 則是，奠基於資料 (data) 的收集，希望機器可以針對某一個問題給出二元區分，在設計系統時並不著重程式設計，而是著重資料、演算法模型以及模型的運作成效。就前者所涉及的問題是，法

律是否真的可以被程式編碼化？法律規範所固有的問題是否會移轉到程式碼去？這個問題的處理目前暫無答案，雖然人工智慧學者皆較為樂觀的認為，法律可以被程式碼化。⁶² 但是 Hildebrandt 認為，法律的文本具有特殊的規範性意義，程式編碼並無法展現出這種法律獨特的規範性，因為程式編碼是一種關於「世界的什麼」資訊生產過程，而不是「世界會是什麼」的法律生產過程。**根據筆者的初步研究，法律規範的部分內容的確有可以被編碼化，但是法律規範很大程度上是文字作為表達媒介，因此文字所具有歧異性與模糊性如何被處理仍有待討論。**這需要人工智慧學者與法理學者共同合作，思索法律形上學等基礎問題，才有初步的答案。

對於以資料驅動為主的 AI，問題更為險峻。因為，系統並不是依靠法律規則而產生「判斷」(或區分) 的，**系統是產生法律規則所要的區分，但不是依據法律規則而來。**例如，系統產出的結果是再犯率高低、勝訴機率高低或是特定類型之案件結果。但是此類結果不是因為系統適用法律規則，而是因為系統是透過大量資料的「學習」而來。因此，人工智慧系統在運作上所適用的那個對象到底是不是「法律規範」本身是一個極大的爭議，必須要清楚地區分出適用法律與找出適用法律後的行為模式是兩個不同層次的問題，不應該混為一談。

此外，歷史理念型的法治觀還向 ALI 提出一個關鍵的問題：承認經由民主立法程序的法律具有拘束力是因為立法者本身的正當性，他們是多元利益、經由民主程序選舉出來的社會成員的代議士。雖然不一定是專家，但是作為法律設計者的立法者具有將多元意見帶入法規制訂的可能。這一點在 ALI 上顯然非常欠缺，我們可以

⁶² 例如，Bart Verheij (2020) 就明確指出，人工智慧可以是法律。

說將 ALI 僅視為執行法律的系統會錯誤地掩蓋了系統設計者作為後設專業知識 (meta-expertise) 的爭議。⁶³ 他們本身並未受到民主正當程序的監督，到底是如何將法律轉化為程式或是演算法也不清楚。甚至，在將法律轉化為程式的過程中所產生的落差亦將由系統的使用者承擔。⁶⁴ 從這個角度來看，歷史理念型與 ALI 的緊張關係並不亞於其他三種法治觀。

透過上述的檢討，我們已經可以見到 ALI 是如何與法治的理念產生緊張關係。不論是有限度的相容論或是最低限度的相容論在實質理念型、程序理念型以及歷史理念型的法治觀下皆不成立。然而，相容論者還可能會主張「好吧，我們在實質價值、法律程序保障以及法律本體論是有不相容的可能。但是，讓我們著重 ALI 如何促成行動協調以及確保法律的落實。至少 ALI 可以與形式理念型的法治觀相容吧。」

最後這個相容論的主張值得我們認真對待，因為它將 ALI 與法律的關鍵特色聯繫在一起：ALI 可以促進行動協調。這將我們帶入了法律與行動之間的關係之中。下節筆者將奠基於前面的分析，說明 ALI 如何會與形式理念型法治觀相容與相抵觸。

二、對於行動協調功能的反省

在討論形式理念型的法治觀與 ALI 是否相容的議題前，首先必須要明確定位，若形式理念型與 ALI 相容時，到底是指什麼對象相容？精確來說，是指 ALI 可以塑造行動者的行動協調與行動模式，

⁶³ 後設專業知識一詞是來自於 Pasquale (2020a) 在中研院主題演講的使用。

⁶⁴ Lawrence Solum (2019) 在〈人造的智慧法律〉(“Artificially Intelligent Law”) 一文中就以思想試驗方式說明這種自我產生規則與執行規則的人工智慧系統會面臨到的法律授權 (delegation) 問題。

而這種塑造與形式理念型所要達成的法律的穩定性與可預測性的價值是一致的。法律的穩定性與可預測性是多元社群成員彼此行動協調的必要條件，換言之，行動者在與他人進行交往時，他可以預期他人也會依循法律而行事，或是若有爭議時依法律而解決，藉以行動者可以調整自身策略、安排要實現的目標或是運用法律達成自身利益最大化。

因此，相容論可以主張，若著重法律的應然與行動者的實然之間的關係的話，ALI 可以協助法律內容的落實，當行動者預測到法律內容是會被落實的，他可以調整自身的行動模式與策略，因此藉由法律規範於當前個例的可適用性 (applicability) 讓行動者可以提高對他人的行動預測 (prediction of others actions)，進而協調社會成員彼此的行動產生穩定性。例如，當法律規定駕駛車輛應行駛於道路之右側，這個規定之存在與執行確保了用路人得以預期他人之行動，並且透過此一預期協調自身的行動反應。相容論可以堅持一種很薄的法治觀 (a thin conception of rule of law)，其希望法律體系具有高度穩定性與可預測性，而採用 ALI 會與此價值相容。

的確，當法律若沒有被落實時，法律作為協調行動的功能會喪失，也因此即使法律的應然與實然關係複雜，但是實然層面的關注是重要的。此外，雖然從應然推導不出來實然，但是實然的行動實踐往往可以作為應然規範存在的徵兆。相容論在此的論證涉及到了，法律與行動之間的複雜關係的重要面向。對相容論而言，ALI 就是一種用來管制社群成員之行動的機制。因此，我們可以將 ALI 視為法律的執行者。⁶⁵ 那麼問題變成了：ALI 可以有哪幾種方式改變

⁶⁵ 由於之前我已經分析了實質理念型、程序理念型與歷史理念型跟 ALI 的緊張關係，因此此處我就先擱下 ALI 不具有內在觀點的這個討論，因為我們可能可以更退一步思考：假設僅將 ALI 視為一種對於行動者行動管制的機制，那麼這仍會跟法治理念

社群成員的行動結構？筆者在貳、二、(一) 之中已指出兩種造成實踐差異的說明模型：決策模型與侷限模型。以下我將透過這兩種模型說明法律人工智慧對於行動改變的關鍵差異。

侷限模型的成立在於，行動者在接受規則時同時自我允諾 (precommitment) 會遵循規則，透過「現在的我」來侷限「未來的自我」的實踐選項的範圍，然而當未來的自我要實現規則的內容的時候，行動主體不是思辨規則的內容，而是實施規則，進而使規則成為行動的指引方針。可以見到，侷限模型之所以能適當說明規則如何創造實踐差異，在於它說明了行動者在不同時間上如何透過規則維持自身對於諸多不同行動選項的偏好，以及對於規則的思辨與對於規則的實施之間的差異。若從侷限模型的角度來看，**現在的我與未來的自我對於規則與行動之間的關係是構成了法律主體的重要關鍵，行動者自願地 (volitional) 實施規則，減少未來其他行動選項。**雖然 Shapiro 指出「如何減少行動選項」仍需要進一步討論，但是侷限模型並未取消行動者對於規則的思辨與實施的差異。

侷限模型是建立在規範內容與規範實踐的落差可能性之上，「自我」運作在現在與未來對於規範與行動指引的落差之中，這是一種實踐縫隙 (practical gap)。但是，高效率執法系統正是希望取消這種實踐縫隙。⁶⁶ 由於 ALI 希望提高法律的實效程度 (實然)，因

有衝突嗎？如果我們可以說明，衝突仍存在的話，那麼 ALI 與法治的緊張關係這個主張就具有相當合理性。

⁶⁶ 根據顏厥安 (2006: 71-74) 的研究，規範的實踐縫隙主要可以區分為三個部分：第一，實效縫隙，這意味著一個有效規範並不能保證規範對象的遵守或是服從，即使規範對象肯認其規範內容之正當性或道德妥當性。第二，行動指引縫隙，這表示一般性規範並未告訴行動者在具體情境下應該遵循的方針。第三，態度縫隙意味著規範的適用者或是接納者對於規範到底抱持何種態度，其實是不易察覺的。

此目標便是減少實踐縫隙，它必然也要降低行動者對於具體情境的考量可能性，要求行動者減少或不理會「法律規範是否適用於當前處境」。簡言之，ALI 要做的就是行動者實施法律之內容，而減少思辨法律之內容。

因此，侷限模型的 ALI 目標就是降低行動者在具體情境，思辨法律如何指引行動的判斷空間，減少「現在的自我」與「未來的自我」的互動可能性。而運用高效率執法機制就是減少思辨法律的一種方式。高效率執法機制在意的是，行動者就是要做出符合規則的行為 (behavior)。技術的進步已讓「執法機制」變得多樣且具有效率，且在對於行為之調控上具有比以往更細微的操作。主體是否可能在諸多調控下而展現出來是一個大課題。以往法學不會覺得這是個大課題是因為行為調控機制往往是作為法學思辨對象而進入問題視野的，但是當調控機制也進入法律內部時，那麼我們如何面對便是一個重要課題。筆者必須承認，在此尚未有詳細的看法。但是，這篇文章的分析至少给出了一些基本方向與討論可能。

如果採取侷限模型的 ALI 跟法治原則上會有緊張關係，那麼決策模型會是較好的選項嗎？決策模型的 ALI 的確是一個可以考量的方向。決策模型的 ALI 思考的是，ALI 如何改變行動者偏好的問題。一般來說，ALI 可以使用兩種一般性方式改變行動者偏好。第一，ALI 增加行動者遵守法律的誘因。第二，ALI 降低其不遵守法律的誘因。從筆者之前初步描繪的行動主體決策理據的結構來看，問題變成了：ALI 要向主體傳遞的理據類型以及內容是什麼？以及此類理據類型如何改變跟影響行動者的自主判斷的理據結構？

以測速系統為例，我們可以將 ALI 設計為，一旦有違法的狀況就逕行舉發，讓駕駛人收到交通違規罰單，或是將 ALI 設計為提高駕駛人遵守速限之誘因，例如遵守限速者可以參與樂透抽獎等等

(Sorrel, 2010)。雖然無法一概而論，但是，較佳的方式或許是，ALI 原則上必須向行動主體說明法律內容之道德理據的方向前進，因為這是行動主體的一階理據考量，是關於行為應該與不應該做的實質理據。例如，工具理性、法律內容之可實現性、法律之道德正當性、違反法律所造成的傷害等等。當主體可以充分思考一階理據時，也可推論他對於法律之要求越來越具有法的忠誠確信，這種 ALI 的設計反而會加強了行動者對於法律的信念。

但如果 ALI 無法朝向上述的方向設計，ALI 至少也可用來強調法律本身的規範性，例如重申速限規定等等。如果設計者覺得這樣還不夠，那麼 ALI 有可能就要改變行動者對於處罰、處罰可能性跟警察的偏好程度。通常 ALI 會朝向增加「處罰可能性」對於行動者的影響。

筆者主張，如果 ALI 的設計是以告訴行動者「一旦違法，就很可能受到處罰」為主要改變偏好的方式，那麼我們必須小心該系統的應用。因為，這種方式是要行動者將「處罰」，由可能性 (likelihood) 提升到高度可能性 (highly likelihood)，甚至在現今科技下是可以變成確定性 (certainty) 的層次。這種增強處罰之可能性對於行動者的因果影響力，雖然看似不影響行動者思考其他理據的可能以及因果影響力，但是這仍然會影響行動者進行理據的排序。如果我們將各項理據的重要性假定為理據對於行動者的因果影響力，那麼很明顯地提升「處罰可能性→行動者」的因果影響力，就是會降低其他理據對於行動者的因果影響力。

的確，我們還不知道這種以提高執法效率而改變因果影響力的介入方式會對於行動者的自主判斷空間有多大的影響。但是，我們知道提高「處罰可能性→行動者」的因果影響力，會影響行動者對於其他因素的排序。也因此，從這個角度來看，高效率執法的 ALI

是有疑慮的。它會讓行動者在面對具體情境中產生兩難：要嘛面對違法的制裁效果，要嘛遵守法律。這個兩難在人類執法的情境中，不太需要考量，但是機器執法的情境中卻可能時常出現。⁶⁷

此外，如果 ALI 不是僅提升處罰可能性，而是將處罰可能性變成處罰的確定性，那必須要小心「一旦違法，就會開罰」的訊息很可能會改變了行動者遵守法律的因果結構。⁶⁸ 或許有人會說，行動者可以在事後提出申訴，因此不必然受到處罰。但是，這仍沒有處理「確實開罰」本身是作為凌駕其他理據的可能。且行動者提出申訴反而是預設了確切執行的理據的存在以及效果。

或者可能有人會指出，即便國家透過「一旦違法，就會開罰」要影醒行動者的行動，但是行動者也具有自主判斷的空間去選擇不理會 ALI 所給出的這個訊息，因為沒有人真正強迫行動者如此做。這種觀點將強制限縮到很小的範圍，認為唯有物理性手段的強制才是強制。可是，在日常生活中我們都可以見到當惡害的通知與實現幾乎是確定時，這種方式對於行動者心理強制作用不亞於物理性強制。當然，物理性強制還是不一樣的，因為物理性強制可能不在意行動者的心理狀態。⁶⁹

⁶⁷ 我們再舉一個日常例子說明。甲開車遇到紅燈停下來，此時車後有一輛救護車鳴笛，甲的車道是唯一救護車可通過的車道。此時，如果是人類執法，甲前面有警察，這個時候甲闖紅燈，將車子停在道路中央，也不會受罰。但是，若是此時甲的車道有闖紅燈照相偵測，甲就肯定會收到罰單，但是，甲事後可以舉證，撤銷罰單。若我們將時間點定在「甲是否要闖紅燈時」，真人執法跟機器執法在制裁上所產生的差異就會出現。

⁶⁸ 審查人指出「必定受到處罰」的用語不妥。筆者後來也這麼認為，因此修正為「就會開罰」。這意味著行動者會收到違法事實的罰單，至於這個罰單最終是不是合法的罰單本文暫且不論。

⁶⁹ 非常感謝審查人細緻地指出這一個區分，其實不是惡害之通知跟物理強制有相同之效果，而是惡害之通知所產生的心理強制作用跟物理強制的效果相類似。

最後，可能還有一種意見認為，假如警察的執法是基於決策模型，那跟基於決策模型的 ALI 在改變行動者的偏好上會有相當差異嗎？⁷⁰ 針對這個問題，筆者認為在分析上可能不會有太大差異，因為，關鍵仍在於警方要介入行動者因果結構圖的哪一個變項。若是有無限的警力可以使用，那麼理論上會跟 ALI 沒有差異，但實際上警力是有限資源，因此可能因果影響力會有差異。但是有個差異會存在，就是在辨識這種事實是法律事實上，ALI 與警方的判斷會有不同。⁷¹

從上述的分析可以見到，如果 ALI 要與形式理念型相容的話，那麼它至少必須是以改變行動者偏好的方式而被設計的，且此偏好之改變原則上必須以提供正向誘因的方式，去影響行動者對於法律之規範內容的一階理據的判斷，而不是透過制裁之確切執行之方式達成。當然，在特殊例外的情況下可以允許 ALI 以制裁之確切執行作為改變行動者偏好的主要方式，但是由於這種方式與法律主體的道德價值相衝突，它必須是一種例外且有限度的使用才行。

伍、結論：思索法律人與智慧系統的未來

現在隨處可見 ALI 在許多法律任務的履行、功能的運作以及權限的分配上佔有一席之地。在這樣的發展下，或許會有個疑問：從法治的角度上來看，到底要如何畫定人與機器的互動關係？根據本文的分析，可以獲得下述結論：

第一，替代型 ALI 會根本地與法治理念相牴觸，因此不應該接受。因此，法律人工智慧系統不論是在設計、研發或應用上，都應

⁷⁰ 感謝審查人提出這一個意見，筆者之前並未審視這種情況。

⁷¹ 請參考本文註 22 的例子。

該採取輔助人類智慧行使的方向前進，讓人類保有最終判斷權限。這個結論與 Frank Pasquale (2020b: 3-7) 在《機器人新法則：人工智慧時代下捍衛人類專長》(*New Laws of Robotics: Defending Human Expertise in the Age of AI*) 一書中是一致的。

第二，在輔助型的 ALI 的前提下，仍必須小心其與不同法治觀的緊張關係。如果要透過 ALI 來執行法律，那麼最適當的方式是利用 AI 協助行動者對於法律之內容進行一階理據的釐清或是考量，而不是以提高執法效率的機制進行，提高執法效率的構想企圖高度擬和法律的實然與應然，在這個過程中主體的自主判斷空間會受到處罰可能性到行動者之間因果影響力提升的影響，筆者並非完全反對運用 ALI 提升執法效率，但我們必須要小心這個運用對於一般公民在思索遵守法律的各式理據間的緊張關係，筆者運用因果圖的主要目的就是希望細緻區分 ALI 的不同運用方式、ALI 對於不同理據的干預以及不同干預的類型會對於法治理念的維持有相當影響。

第三，並非任何法律事務的輔助履行也是全然沒有問題。我們仍需要注意到有哪些法律事務判斷之領域可以透過 AI 來輔助本身是一個法律問題，更是政治道德層面的問題。作為法哲學層次的探討，筆者認為善用思想試驗 (thought experiment) 是一個具體可行的方向。當前在思考人工智慧時所提到的思想試驗通常指涉 Philippa Foot (1967) 或 Judith Jarvis Thomson (1985) 所提的電車難題 (trolley problem)。電車難題固然有其討論價值，但是筆者認為，或許可以更深入應用思想試驗的特色找出到底哪些法律事務的領域會涉及到較為根本的法律課題。例如，Lawrence Solum (2019: 58-60) 在〈人造的智慧法律〉(“Artificially Intelligent Law”) 一文中便應用了交通管理 AI 的思想試驗，指出雖然自主性的 ALI 尚未具體實現，但是從思想試驗的方式開始會發現到此類系統在開發與

應用過程中逐漸與法律既有的基本原則產生衝突。例如，我們可能會將 ALI 應用在氣候變遷 (climate policy) 的管制體系之中、監獄管理 (prison regulation)、或是反恐政策 (terrorism) 的擬定上等等。在上述脈絡下發展的 ALI 無疑地會在法律的授權上產生許多爭議，此類爭議除了是政治爭議之外，也會涉及到道德上或是政治道德意義下的爭議，例如跟正義與正當性產生衝突。

第四，或許可以更深入一點看現今 AI 的技術運用是否真的可以稱之為「智慧」？本文是將 ALI 當成既定的探討對象，然而以找出資料間相關性的機器學習技術是否真的稱得上智慧卻是個重要問題。筆者的想法是，以關聯性為主的工程設計無法帶我們更上一階地理解智慧這個概念。我們必須要思考世界如何可以透過因果結構再現於機器內部，從而建構出可以處理關聯 (association)、介入 (intervention) 與反事實 (counterfactual) 等問題的機器。現今以關聯性、需要大量收集資料的機器學習的「智慧」系統也構成了對於法治的相當威脅，因為這樣的系統希望介入人類決策的領域，但是介入人類決策本身涉及到兩個層次的問題：第一，系統所希望介入的那個人類決策之因果結構想像是什麼？第二，系統自身內部對於資料的因果想像是什麼？本文只很初步處理了第一個層次的問題。而第二層次的問題，很大程度就是機器的可解釋性的問題，但不僅限於此。本文囿於篇幅與問題意識，無法處理第二層次的問題。然而，如果要更堅實地捍衛法治的價值就必須面對第二層次的問題。⁷²

從上述的幾點來看，本文的分析是個起步，如果我們認為法治的核心是維持個人道德的自主判斷空間，那麼我們應該要開始思考

⁷²非常感謝審查人提出關聯性機器學習的問題以及此類機器學習系統對於法治的威脅，筆者受此觀點啟發甚大。

在輔助型的 ALI 發展下，究竟哪些事務是適合透過 AI 來輔助判斷的？我們也需要發展不同領域對於 ALI 的核心規範性問題是什麼？上述課題雖然無法進一步在本文中獲得處理，但是筆者相信善用思想試驗能夠提供相當助益，唯這必須留待他文為之。

參考文獻

- 王鵬翔 (2012)。〈獨立於內容的理由與法律的規範性〉，《中研院法學期刊》，11: 203-247。(Wang, P.-H. [2012]. Content-independent reasons and the normativity of law. *Academia Sinica Law Journal*, 11: 203-247.)
- 王鵬翔、張永健 (2019)。〈論經濟分析在法學方法之運用〉，《臺大法學論叢》，48, 3: 791-871。(Wang, P.-H., & Chang, Y.-C. [2019]. On the use of economic analysis in legal methodology. *National Taiwan University Law Journal*, 48, 3: 791-871.) [https://doi.org/10.6199/NTULJ.201909_48\(3\).0001](https://doi.org/10.6199/NTULJ.201909_48(3).0001)
- 交通部鐵道局 (2019)。〈臺鐵豐原站智慧型影像監控系統試辦案〉。<https://www.rb.gov.tw/news.php?lmenuid=11&smenuid=0&tmenuid=0&inford=2060> (Railway Bureau, MOTC. [2019]. *Digital monitoring system in TRA Fengyuan Station.*)
- 艾莎·威克福斯 (2021)。《另類事實：關於知識和它的敵人》(葉品岑譯)。春山。(原著 2017 年出版) (Wikforss, A. [2021]. *Alternativa fakta: Om kunskapen och dess fiender* [P.-C. Ye, Trans.]. Spring Hill. [Original work published 2017])
- 吳從周 (2019)。〈初探 AI 的民事責任〉，劉靜怡 (編)，《人工智慧相關法律議題芻議》，頁 89-119。元照。(Wu, C.-J. [2019]. Initial inquiries into civil liabilities from the AI perspective. In J.-Y. Liu [Ed.], *On the legal issues of AI: A primer* [pp. 89-119]. Angle.)
- 李開復、王詠剛 (2017)。《人工智慧來了》。天下文化。(Li, K.-F., & Wang, Y.-G. [2017]. *Upcoming artificial intelligence*. Commonwealth.)
- 李榮耕 (2019)。〈初探刑事程序法的人工智慧應用〉，劉靜怡 (編)，《人工智慧相關法律議題芻議》，頁 120-152。元照。(Li, Y.-G. [2019]. Initial inquiries about the application of AI in criminal procedural law. In J.-Y. Liu [Ed.], *On the legal issues of AI: A primer* [pp. 120-152]. Angle.)
- 東吳大學 (2018)。〈東吳 AI 中心攜手法扶會、SAS 以「人工智慧」提升「法律扶助」效益〉。<https://www-ch.scu.edu.tw/october/news/>

9267 (Soochow University. [2018]. *Center for applied artificial intelligence research works with legal aid foundation of Taiwan and SAS to promote legal aid utilities via artificial intelligence.*)

林昀嫻、王道維 (2020)。〈AI 親權判決預測系統：法律、社會、技術與應用面向〉，《台灣人工智慧行動網》。<https://ai.iias.sinica.edu.tw/ai-child-custody-decision-prediction> (Lin, Y.-H., & Wang, D.-W. [2020]. Child custody via AI prediction system: Application in terms of law, society and technology. *Taiwan Artificial Intelligence Wise Agent Network.*)

林勤富、李怡俐 (2020)。〈人工智慧時代下的國際人權法：規範與制度的韌性探索與再建構〉，李建良 (編)，《法律思維與制度的智慧轉型》，頁 413-465。元照。(Lin, C.-F., & Li, Y.-L. [2020]. International human rights law under the era of AI. In J.-L. Li [Ed.], *The transformation of legal system and legal reasoning under the influence of AI evolution* [pp. 413-465]. Angle.)

法務部行政執行署桃園分署 (2020)。〈桃園分署與桃園市府機關合作以車追人〉。<https://www.tpk.moj.gov.tw/media/20253243/05-%E6%A1%83%E5%9C%92%E5%88%86%E7%BD%B2%E4%BB%A5%E8%BB%8A%E8%BF%BD%E4%BA%BA-%E5%85%A9%E5%B9%B4%E8%BF%BD134%E5%8F%B0-%E6%94%B6%E5%9B%9E484%E8%90%AC.pdf> (Taoyuan Branch, Administrative Enforcement Agency, Ministry of Justice. [2020]. *Taoyuan Branch, Administrative Enforcement Agency, Ministry of Justice works with local municipal offices to track owners via vehicles. Ministry of Justice News.*)

哈特 (2010)。《法律的概念》(許家馨、李冠宜譯)。商周。(原著 1961 年出版) (Hart, H. L. A. [2010]. *The concept of law* [C.-H. Hsu & K.-I. Li, Trans.]. Business Weekly Publications. [Original work published 1961])

國立清華大學 (n.d.)。〈AI 輔助親權判決預測網頁〉。<https://custodyprediction.herokuapp.com> (National Tsing Hua University. [n.d.]. *AI's prediction of legal judgments for child custody.*)

- 張永健 (2020)。《社科民法釋義學》(初版)。新學林。(Chang, Y.-C. [2020]. *Rechtsdogmatik of civil law* [1st ed.]. New Sharing.)
- 莊世同 (2017)。〈論海耶克的自由主義法治觀：一個批判性的探索〉，《臺大法學論叢》，46, 3: 665-702。(Jhuang, S.-T. [2017]. On Hayek's liberal conception of the rule of law: A critical survey. *National Taiwan University Law Journal*, 46, 3: 665-702.) <https://doi.org/10.6199/NTULJ.2017.46.03.01>
- 陳弘儒 (2021)。〈論原意主義的憲法解釋方法論：對於中立化與語意規範性的初步反省〉，《中研院法學期刊》，29: 135-224。(Chen, H.-J. [2021]. On original public meaning Originalism: A critique of its neutralization and the source of semantic normativity. *Academia Sinica Law Journal*, 29: 135-224.)
- 陳起行 (2002)。〈美國法理學發展概述，1870-1970〉，《政大法學評論》，69: 1-27。(Chen, C.-H. [2002]. A sketch of American jurisprudence, 1870-1970. *Chengchi Law Review*, 69: 1-27.)
- 新北市政府警察局交通大隊 (2021)。〈區間平均速率科技執法設置地點〉，<https://www.traffic.police.ntpc.gov.tw/cp-918-44982-27.html> (Traffic Police Corp, New Taipei City Police Department. [2021]. *Spots for average speed technological enforcement*.)
- 國家教育研究院 (2012)。〈國家教育研究院雙語詞彙、學術名詞暨辭書資訊網“congruence”〉。<https://terms.naer.edu.tw/search/?q=Congruence&field=ti&op=AND&group=&num=10> (National Academy for Education and Research [2012]. *Congruence*.)
- 漢娜·弗萊 (2019)。《打開演算法黑箱：反噬的 AI、走鐘的運算，當演算法出了錯，人類還能控制它嗎？》(林志懋譯)。臉譜。(原著 2018 年出版) (Fry, H. [2019]. *Hello world: How to be human in the age of the machine* [Z.-M. Lin, Trans.]. Faces. [Original work published 2018])
- 顏厥安 (2001)。〈道德、政治與法律：由對現代性法學之觀察反思緊急命令體制〉，《臺大法學論叢》，30, 6: 1-28。(Yen, C.-A. [2001]. Moral, politics and law—A reflection on president's right to emergency decree based on the observation of jurisprudence of modernity. *National Taiwan University Law Journal*, 30, 6: 1-28.) <https://doi.org/10.6199/NTULJ.2001.30.06.01>

- 顏厥安 (2006)。〈規範縫隙初探〉，楊日然教授紀念論文集編輯委員會 (編)，《法律的分析與解釋：楊日然教授紀念論文集》，頁 61-89。元照。(Yen, C.-A. [2006]. A preliminary to normative gap. In Editorial Board for proceedings in memory of professor R.-R. Yang [Ed.], *Analysis and interpretation of law: Proceedings in memory of professor R.-R. Yang* [pp. 61-89]. Angle.)
- 顏厥安 (2018)。〈人之苦難，機器恩典必看顧安慰：人工智慧、心靈與演算法社會〉，劉靜怡 (編)，《人工智慧相關法律議題芻議》，頁 47-85。元照。(Yen, C.-A. [2018]. All watched over by machines of loving grace: AI, mind and algorithm society. In J.-Y. Liu [Ed.], *On the legal issues of AI: A primer* [pp. 47-85]. Angle.)
- Anderson, S. (2011). Coercion. *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/win2017/entries/coercion/>
- Bix, B. (2004). *A dictionary of legal theory*. Oxford University Press.
- Campanella, T. (2017, July 10). Robert Moses and his racist parkway, explained. *Bloomberg*. <https://www.bloomberg.com/news/articles/2017-07-09/robert-moses-and-his-racist-parkway-explained>
- Caro, R. A. (1975). *The power broker: Robert Moses and the fall of New York*. Vintage.
- Christman, J. (2020). Autonomy in moral and political philosophy. In *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/entries/autonomy-moral>
- Executive Office of the President. (2016). *Preparing for the future of artificial intelligence*. National Science and Technology Council Committee on Technology. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf
- Fallon, R. (1994). Reflections on the Hart and Wechsler paradigm. *Vanderbilt Law Review*, 47, 4: 953-991.
- Fallon, R. (1997). "The rule of law" as a concept in constitutional discourse. *Columbia Law Review*, 97, 1: 1-56. <https://doi.org/10.2307/1123446>

- Farnsworth, W. (2007). *The legal analyst: A toolkit for thinking about the law*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226238364.001.0001>
- Foot, P. (1967). The problem of abortion and the doctrine of the double effect. *Oxford Review*, 5: 5-15.
- Fuller, L. (1969). *The morality of law* (Rev. ed.). Yale University Press.
- Gateshead Council. (2013). *Mobile CCTV enforcement—Frequently asked questions*. <https://www.gateshead.gov.uk/media/3701/Mobile-CCTV-Enforcement-Frequently-Asked-Questions/pdf/CCTV-FAQ.pdf?m=636438393766370000>
- Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2019, February 3). *Explaining explanations: An overview of interpretability of machine learning*. Cornell University. <https://arxiv.org/abs/1806.00069>
- Hart, H. L. A. (1994). *The concept of law* (2nd ed.). Oxford University Press.
- Hauer, T. (2018). Society and the second age of machines: Algorithms versus ethics. *Society*, 55: 100-106. <https://doi.org/10.1007/s12115-018-0221-6>
- Hildebrandt, M. (2016). Law as information in the era of data-driven agency. *The Modern Law Review*, 79, 1: 1-30. <https://doi.org/10.1111/1468-2230.12165>
- Hildebrandt, M. (2018, August 6). Algorithmic regulation and the rule of law. *The Royal Society*, A376: 20170355. <https://doi.org/10.1098/rsta.2017.0355>
- Kelsen, H. (1991). *General theory of norms* (M. Hartney, Trans.). Oxford University Press.
- Kurzweil, R. (2012). *How to create a mind: The secret of human thought revealed*. Viking.
- LawGeex. (2018, February 26). AI vs. lawyers: The ultimate showdown. *The LawGeex Blog*. <https://blog.lawgeex.com/ai-more-accurate-than-lawyers>
- Lipton, Z. (2017). *The mythos of model interpretability*. Cornell University. <https://arxiv.org/abs/1606.03490>

- Liu, H.-W., Lin, C.-F., & Chen, Y.-J. (2019). Beyond State v. Loomis: Artificial intelligence, government algorithmization and accountability. *International Journal of Law and Information Technology*, 27, 2: 122-141. <https://doi.org/10.1093/ijlit/eaz001>
- Palombella, G. (2010). The rule of law as an institutional ideal. In L. Morlino & G. Palombella (Eds.), *Rule of law and democracy: Inquiries into internal and external issues* (pp. 1-37). Brill. <https://doi.org/10.1163/ej.9789004181694.i-215.6>
- Pasquale, F. (2019). A rule of persons, not machines: The limits of legal automation. *The George Washington Law Review*, 87, 1: 1-55.
- Pasquale, F. (2020a, November). *Battle of the experts: Artificial intelligence, meta-expertise, and professional performance*. Paper presented at the conference of the Institutum Iurisprudentiae, Academia Sinica on the 8th Law, Science and Technology AI and Professionalism, Taipei.
- Pasquale, F. (2020b). *New laws of robotics: Defending human expertise in the age of AI*. Harvard University Press. <https://doi.org/10.4159/9780674250062>
- Pearl, J. M., Glymour, M., & Jewell, N. (2016). *Causal inference in statistics: A primer*. Wiley.
- Rawls, J. (1971). *A theory of justice*. Harvard University Press.
- Robeyns, I., & Byskov, M. F. (2020). The capability approach. In *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/win2020/entries/capability-approach>
- Russell, S. (2019). *Human compatible: Artificial intelligence and the problem of control*. Viking.
- Russell, S., & Norvig, P. (2010). *Artificial intelligence: A modern approach* (3rd ed.). Pearson Education.
- SAE International J3016 (2014). *Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems J3016_201401*. https://www.sae.org/standards/content/j3016_201401/
- Schneider, S. (2019). *Artificial you: AI and the future of your mind*. Princeton University Press. <https://doi.org/10.1515/9780691197777>

- Shapiro, S. (1998). The difference that rules make. In B. Bix (Ed.), *Analyzing law: New essays in legal theory* (pp. 33-64). Oxford University Press.
- Shapiro, S. (2002). Authority. In J. Coleman, K. E. Himma, & S. Shapiro (Eds.), *The Oxford handbook of jurisprudence and philosophy of law* (pp. 382-439). Oxford University Press.
- Shapiro, S. (2011). *Legality*. Harvard University Press.
- Solum, L. (2019). Artificially intelligent law. *BioLaw Journal-Rivista di BioDiritto*, 1: 53-62. <https://doi.org/10.2139/ssrn.3337696>
- Sorrel, C. (2010, December 6). Swedish speed-camera pays drivers to slow down. *WIRED*. <https://www.wired.com/2010/12/swedish-speed-camera-pays-drivers-to-slow-down/>
- Tamanaha, B. (2004). *On the rule of law: History, politics, theory*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511812378>
- Tamanaha, B. (2007). *A concise guide to the rule of law*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1012051
- Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence* (1st ed.). Knopf.
- Thomson, J. J. (1985). The trolley problem. *The Yale Law Journal*, 94, 6: 1395-1415. <https://doi.org/10.2307/796133>
- Tsai, C.-H. (2020). Artificial wisdom: A philosophical framework. *AI & Society*, 35: 937-944. <https://doi.org/10.1007/s00146-020-00949-5>
- Verheij, B. (2020). Artificial intelligence as law. *Artificial Intelligence and Law*, 28: 181-206. <https://doi.org/10.1007/s10506-020-09266-0>
- Waldron, J. (2020). The rule of law. In *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/rule-of-law>
- Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., West, S. M., Richardson, R., Schultz, J., & Schwartz, O. (2018, December). *AI now report 2018*. https://ainowinstitute.org/AI_Now_2018_Report.pdf

- Yen, A. (2015). On democratic resistance and the rule of law. *Asian Legal Philosophy*, 2, 1: 53-80. https://www.academia.edu/24249856/On_Democratic_Resistance_and_the_Rule_of_Law
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs. <https://doi.org/10.1080/23738871.2019.1637914>

Rule by Algorithms? A Preliminary Study on the Tension Between the Rule of Law and AI

Hung-Ju Chen

Institute of European and American Studies, Academia Sinica

E-mail: marxchen@gate.sinica.edu.tw

Abstract

Artificial legal intelligent systems have been extensively developed to perform various legal tasks. However, rare discussions focus on the question of whether such development might conflict with the idea of the rule of law. In this article I argue that there is an internal tension between ALIs and the rule of law based on two reasons. First, the idea of a legal subject in the rule of law is conceived as an active, responsible, and answerable agent who has the authority for making a decision. The fact that ALIs become more capable of performing legal tasks eventually leads to the possibility of shifting the decision-making authority away from human. The rule of law draws a baseline for developing ALIs, that is, ALIs should be supplementary, not substitute for human. Secondly, the highly efficient mechanism of law enforcement, one of the core features in ALIs, severely decreases an individual's autonomous sphere of moral reasoning for ALIs seek for an assurance that an agent would comply with law through the certainty of legal sanctions. As a result of adopting this mechanism, ALIs aim at influencing an individual's autonomy of ranking reasons. Based on the two arguments, I demonstrate there are two principle on the regulation of ALIs. The first principle prohibits the development of substitute ALI. The second principle prohibits the widely adoption of highly efficient mechanism of law enforcement. This article examines the fundamental issue of the rule of law in the age of AI from the perspective of legal philosophy.

Key Words: artificial legal intelligent system, the rule of law, the practical difference thesis, the shift of decision-making authority, computer functionalism