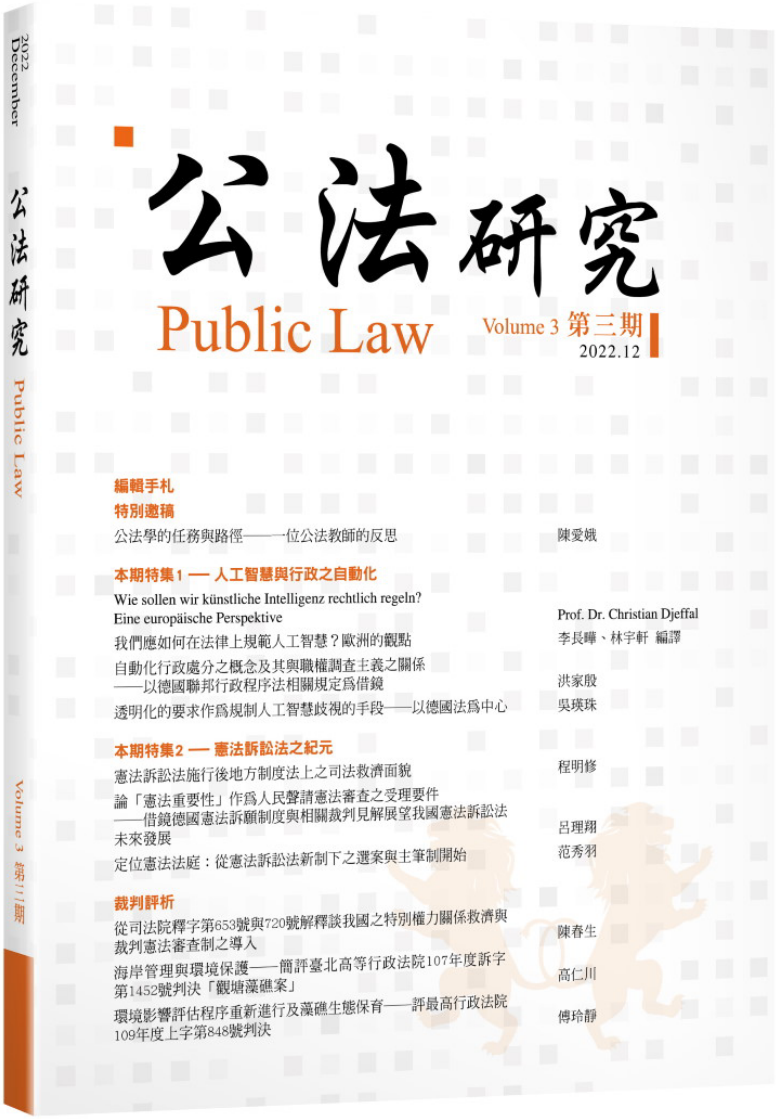

Djeffal, Christian (2022). 我們應如何在法律上規範人工智慧? (Wie sollen wir künstliche Intelligenz regeln?): 歐洲的觀點, In: 公法研究 第三期, Ausgabe 12, S. 33-44.

我們應如何在法律上規範人工智慧？

歐洲的觀點

作者：Prof. Dr. Christian Djeffal *
譯者：李長曄**、林宇軒***

I. **人工智慧管制的歐洲進程**

II. **相關問題**

A. 在「創新」與「安全」之間的管制方向

B. 人工智慧的定義

C. 合致性評價

D. 革新之協助與引導

III. **展望**

* Christian Djeffal 博士為慕尼黑理工大學，法律、經濟與科技教授。
** 逢甲大學土地管理學系助理教授。
***開南大學法律系助理教授。

I. 人工智慧管制的歐洲進程

人工智慧系統在社會的不同領域，尤其是近十年來，已儼然成為了創新驅動者。其乘載了越來越多的任務，同時也越來越深入社會生活。這般發展仍然持續進行中，藉由新的創意，以及運用科技加以實現，許多過往被認為，對於機器來說，是無法想像的事，如今都成為了可能。然而，這些新的可能性卻也時常意謂著新的風險。因此也產生了一個問題，即：是否，以及在具體情況下，應如何透過法律，就人工智慧系統進行管制。有別於「網際網路」，人工智慧並不涉及「強制互聯的技術」，但是，兩者之間法規範的交流，在此也是非常重要，無論是相互學習，還是確保進一步可能的適用，或者，至少確保體系的互相操作性。秉持這樣的精神，本文將引介目前在歐洲的相關討論，尤其是讓更多的讀者，能夠了解相關論述，以及目前仍懸而未決的問題，此外，並促使跨國的討論。因此，重點不應在那些，於撰寫本文的當下，仍未形成的結論，毋寧是在於，以法律進行管制的過程中，所要處理的問題與困難。

在數位化開始蔓延的極初期階段，便納入人工智慧面向的討論。在 1978 年時，法國第一部的資料保護法即有一條規定，係有關自動裁決系統，以及在公行政與司法當中的裁決輔助[1]。此項規定，激發了許多個資保護法之中，有關自動裁決的規定[2]。自從人工智慧在政治的最高層級引發興致之後，從 2017 年開始，便有許多不同的「人工智慧策略」出爐[3]。同樣的，歐盟有關人工智慧的管制進程，乃是源自於一個

---

1 Art. 2 Act No. 78-17 on Information Technology, Data Files and Civil Liberties, Journal Officiel de la Republique Française 1978, 227.

2 *Djeffal*, The Normative Potential of the European Rule on Automated Decisions, ZaöRV 2020, 847 (851 ff.).

3 於此特別參見 *Mazur*, Beijing Wants A.I. to Be Made in China by 2030, New York Times, 2017 （日期：20.07.2017，網址：https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html）; *Vincent*, Putin says the nation that leads in AI ‘will be the ruler of the world’, The Verge, 2017（日期：04.09.2017，網址：https://www.theverge.

策略進程。其會員國於 2018 年初，發表一份聯合聲明[4]；執委會也以一份「人工智慧策略」緊接在後，其中將人工智慧系統的管制，列為當中所包含的一項措施[5]。新任歐盟執委會的主席——烏蘇拉．馮德萊恩——上任時，即將人工智慧的管制列為「最優先」，並在其當初於歐洲議會競選的政見中，有過下列的表示：

「在本人任期一開始的 100 天內，為了歐盟內的一致性，敝人將就人工智慧對人類和倫理的影響，提出一份立法建議。同時並研究我們該如何利用大數據進行創新，以利於我們的社會及企業」[6]。

事實上，人工智慧的管制已經是歐盟執委會五項主要優先事務當中的一項。2020 年 2 月 19 日，歐盟執委會發布了一份白皮書，允許各方利益團體就此一進程表示意見[7]。2021 年 4 月 21 日，執委會即公布其關於人工智慧法（PAIA）的提案[8]。該草案目前正在歐洲議會以及歐盟理事會進行討論。在現階段的討論框架內，甚至之外，都出現了許

---

com/2017/9/4/16251226/russia-ai-putin-rule-the-world）; *May*, PM's speech at Davos 2018: 25 January, 2018（日期：25.01.2018，網址：https://www.gov.uk/government/speeches/pms-speech-at-davos-2018-25-january）; *Macron*, Sorbonne speech of Emmanuel Macron, 2017（日期：26.09.2017，網址：http://international.blogs.ouest-france.fr/archive/2017/09/29/macron-sorbonne-verbatim-europe-18583.html）.

4 Declaration on Cooperation in Artificial Intelligence, 2018（日期：13.04.2018，網址：http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50951）.

5 *Europäische Kommission*, Artificial Intelligence for Europe, COM(2018) 237 final, 2018,（日期：25.04.2018，網址：https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN）.

6 *Europäische Kommission/von der Leyen,* Eine Union, die mehr erreichen will. Meine Agenda für Europa: Politische Leitlinien für die künftige Europäische Kommission 2019-2024, 2019, S. 16（日期：09.10.2019，網址：https://ec.europa.eu/info/sites/default/files/political-guidelines-next-commission_de.pdf）.

7 *Europäische Kommission*, Weissbuch. Zur Künstlichen Intelligenz – ein europäisches Konzept für Exzellenz und Vertrauen, COM(2020) 65 final, 2019（日期：19.02.2020，網址：https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf）.

8 *Europäische Kommission*, COM(2021) 206 final, 2021/0106 (COD), 2021（日期：21.04.2021，網址：https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0019.02/DOC_1&format=PDF）.

多有關的問題，而這些問題都將在本文當中加以介紹。

II. 相關問題

A. 在「創新」與「安全」之間的管制方向

人工智慧的管制實際上，是發生在兩個相對立利益的磁場當中。一方面，人工智慧也應得到來自國家的資助，無論是在基礎方面，抑或是在社會中的創新性定錨方面。另一方面，也應該保護社會，免受該科技使用的負面影響。在此緊張的範圍內，必須選擇一個盡可能合乎比例的方式，以盡可能地兼顧雙邊利益。執委會於其提案中，試圖去形塑一個，適用於所有人工智慧運用的基本框架。其遵循所謂「以風險為基礎」的方法（einem sog. risikobasierten Ansatz），憑藉各種人工智慧的運用，會產生對法益及合法利益不同程度的風險，來決定各種不同程度的要求。第二編（第 5 條）係規定，在人工智慧領域當中，某些被禁止的特定行為（Verbotene Praktiken）；相較於此，第三編（第 6-51 條）則是有關「高風險的人工智慧系統」（Hochrisiko-KI-System），並就其制定了一個詳細的管制體系。第四編（第 52 條）則是對特定的系統，規定了透明度義務（Transparenzpflichten）。第九編（第 69 條）則是針對其他編所無法涵蓋的人工智慧系統，規定了行為準則（Verhaltenskodizes）。由此，可得出下列整體概況：

插圖 1：根據風險高低之分類

風險的分配，有部分是在其中被明確的規定，例如在第 5 條的情況下，其規定了禁止的行為。在其他情況下，該法則是可以藉由行政部門來加以調適。例如，由第 7 條之規定可知，界定高風險運用的附件三，得經由執委會來加以更新。歐盟藉此希望架構一個，能夠適時擴展到新的風險領域的法規框架。為此，這部人工智慧法草案，必須釐清其他法規的適用性，並根據第 2 條第 2 項之規定，定期賦予其他法規優先性。此種管制方向，也讓人想起「民法的大法典化」（großen Kodifikationen des Zivilrechts），因為其制定了一般性規則，然後再透過特別法規對其進行修改。相對於此，仍有其他的管制取向，在不同的國家發揮其作用。巴西一份攸關人工智慧管制的立法建議，即是先設立一般性原則，然後再將這些一般性原則，整合至「與特定的人工智慧應用領域更為相關的」特別法當中[9]。在美國，總統的學術顧問們則

[9] *Câmara dos Deputados,* Marco Legal da Inteligência Artificial, PL 21/2020, 2020（日期：03.02.2022，網址：https://www25.senado.leg.br/web/atividade/materias/-/materia/151547）.

是提出一項程序（Prozess），其係出自於權利法案框架下的主觀權利；然後將此程序轉化為普通法。（本文作者特別註解：就此仍尚未有草案）[10]加拿大最近的一項立法草案，大體上也是遵循「以風險為基礎」的人工智慧管制方法[11]。「以原則為基礎」的管制方法（prinzipienbasierte Ansätze）則（似乎）能夠（könnten）加強對「科技的某些特定危險」的認知，而不用立即去設置管制障礙（regulatorische Hemmnisse）。不過，「以原則為基礎」的管制方法總是會與特定內容產生關聯。這裡所產生的問題就在於，人工智慧作為一種新興科技，亦會造成以前所未知的新問題與風險。一個以風險為基礎，且與程序（Prozess）更相關的管制方法，則能夠掌握與處理這些新的風險。然而，程序（Prozess）同時也意謂著行政官僚的阻礙（bürokratische Hemmnisse）。

一個著眼於管制與創新的特殊問題係在於：禁止的問題。倘若某些人工智慧的運用形式是被禁止的話，乃是一種特別嚴格的管制模式，幾乎可說是在該領域，完全阻絕創新的空間。相反地，在某些應用情形——例如有意識地去操縱人——，問題則是：在什麼情況下，仍應忍受、接受「對於使用者所會產生的風險」。準此，第 5 條所規定的禁令，仍飽受爭議，相較於民間團體希望進一步擴張禁止的範疇，工業團體則傾向限縮禁令的範圍。在一份由人工智慧倫理問題專家團體所提出的報告當中，並未含有任何的禁令，僅提及某些具爭議的人工智慧運用領域[12]。此立場後來遭受執委會的一位成員猛烈地批評[13]。基

---

10 *Lander/Nelson*, Americans Need a Bill of Rights for an AI-Powered World, https://www.whitehouse.gov/ostp/news-updates/2021/10/22/icymi-wired-opinion-americans-need-a-bill-of-rights-for-an-ai-powered-world/ （日期：22.10.2022）.

11 *House of Commons of Canada*, Bill C-27 (First Reading), 2022（日期：16.06.2022，網址：https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading）.

12 *Unabhängige Expertengruppe für Künstliche Intelligenz*, Ethik-Leitlinien für eine vertrauenswürdige KI, 2018, S. 44.

13 *Metzinger*, Ethics washing made in Europe, Tagesspiegel, 2019（日期：08.04.2019，網址：https://www.tagesspiegel.de/politik/eu-guidelines-ethics-washing-made-in-europe/241

於如此的論辯，使「嚴格禁令」的問題性，得以繼續被探究。

B. 人工智慧的定義

人工智慧的精準定義，乃是人工智慧管制的一個關鍵問題。原因係在於，人工智慧的應用領域，也就是其管制的邊界。應用領域又取決於要管制的技術或技術狀態[14]。定義人工智慧的難點在於，人工智慧並非一個單一的技術，毋寧是涉及一個「探究能夠獨立解決複雜問題的技術系統」的研究領域[15]。該概念本身可以追溯到 1955 年一份非常公開的「第三方資助申請」（Drittmittelantrag），該申請同樣公開探詢解決技術問題的方法[16]。因此有許多不同的技術都被認定為人工智慧。不同於早期，人工智慧也等同於專家系統，現今則是集中在討論機械學習，以及所謂的「深度類神經網路」。這也就引發了一個問題，應如何在法律中定義人工智慧。在草案第 3 條第 1 款當中，有關技術上的定義實質上是參考了附件一。其將人工智慧定義如下：

「a) 機器學習的方法，使用包括深度學習在內的多種方法，進行受監督的，不受監督的，以及強化的學習；

b) 以邏輯與知識為基礎的方法，包括知識呈現、歸納（邏輯）編程、知識基礎、推理及演繹機器、（符號式）推理及專家系統；

c) 統計方法、貝葉斯（Bayesche）估計、搜尋及優化方法。」

引人注目之處在於該定義非常廣泛，囊括了機器學習以外的不同

95496.html）.

14 *Gasser/Almeida*, A Layered Model for AI Governance, IEEE Internet Computer. 2017, 58.

15 *Djeffal*, Normative Guidelines for Artificial Intelligence, in *Wischmeyer/Rademacher* (Hrsg), Regulating Artificial Intelligence (2020) 277-293 (277ff. mit weiteren Nachweisen).

16 *McCarthy/Minsky/Shannon/Rochester*, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, 1955（日期：04.04.1996，網址：http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html）.

技術。該立法提案在例如歐洲議會之處，引起其大爭議，也就不足為奇了。準此，歐洲議會的法律事務委員會提議，將 b) 以及 c) 點完全刪除，並在 a) 點當中，加入「以及進化的人工智慧等的優化方法」[17]。此定義因而將比執委會的定義更為限縮，並特別強調人工智慧的新形態。

要在廣義或狹義的人工智慧定義之間作抉擇，並非一件容易的事。廣義的定義會使得適用範圍擴張至那些，早就在使用的「較為一般的系統」（„einfachere“ System）。對這些系統而言，這將意謂著一個高度的行政阻礙。反之，這樣的系統也可能會帶來風險。問題在於，是否也要將這些風險納入適用範圍之內。如果要以這種方式來重新描述緊張關係的話，真正的問題將在於，管制是否應該總體性的保護免受「某些一定程度的風險」（gewissen technischen Risiken）；或者是只有在極度複雜的技術所帶來的特殊風險——例如深度類神經網路的風險——，進入到既存的危險之內時，才提供保護。因為在這些狀況下，演算法並未被明確編程，而是透過「訓練數據」（Trainingsdata）加以編程，因此這些系統都不夠透明，且難以預測它們在某些情況下的舉措。

C. 合致性評價

人工智慧法草案所涉及的是對於特殊規範，特別是關於高風險人工智慧系統的處理。在科技管制的議題上，除了義務本身，如何以及由誰來確認義務是否被遵守亦屬於關鍵的問題。就此歐洲在科技管制的方向上其實已經有了相當大程度的推移。最初審查權限很大程度上

---

[17] *Europäisches Parlament/Committee on Legal Affairs,* Draft Opinion, 2021/0106(COD), 2022, Amendment 285-287（日期：02.03.2022，網址：https://www.europarl.europa.eu/doceo/document/JURI-PA-719827_EN.pdf）.

是由國家所掌控，然而在所謂的共同管制（Ko-Regulierung）或者所謂的「受管制的自我管制」（regulierten Selbstregulierung）的概念下，審查權限被大範圍的向經濟部門移轉[18]。其背後的理念是，由國家制定一個框架，而私部門有能力透過其知識與機遇去尋得適當的解決方案。就此，可以透過如企業對於其義務的自我審查，抑或是透過其他私部門如監察團體來完成。

人工智慧法草案基本上規範了一種不同的模式，大致上可分成兩種類型：一者由人工智慧系統的供應者，在內部控制的框架下（附件六）進行自我要件審查，或者由公告的單位來進行審查（附件七）。這些單位皆是符合草案第 33 條所規範之特定要件，並被授權得以審查人工智慧系統的組織。根據草案第 43 條繁瑣之規定，其進一步強調，人工智慧法清晰明確地將重點放在供應者自我管制的部分。相對於此，經認證之單位只有在例外的情形下，方會承擔審查之任務。根據草案第 43 條第 1 項，在生物測量監控的情形，當受監控者明確表示願由「受認證單位」審查，或是尚缺乏共通標準的時候，可以被視為是例外的情形。同條第 3 項則規範了另一種例外情形，亦即是有其他法規可以適用，且該法規規定了由「受認證單位」審查之情形。

D. 革新之協助與引導

草案第 53 條以下也包含了一項有趣的人工智慧真實實驗室（Reallaboren）附加條款（Sandboxes 沙盒）。就此可以理解為一種管制技術，在一定的要件下，將管制障礙加以排除，藉此使革新變為可能。這是為了減少人工智慧發展阻力所制定的技術規範。草案第 53 條對於真實實驗室作了基本描述，並明確表示，第五節之責任規範就此

[18] *Schulz/Held*, Regulierte Selbstregulierung als Form modernen Regierens. Endbericht Mai 2002, 2002.

不受影響。第 54 條則是對於一些特殊情況作了緩和的規定。就此對於所謂的「歐盟一般資料保護規則」（Datenschutzgrundverordnung）上之「目的拘束原則」（Zweckbindungsgrundsatz）作了一些鬆綁。根據該規則第 5 條第 1b 項，個人相關之資訊須因「確定、清晰且合法之目的」方得提取，並且不得以與上述目的不相符合之方式進一步加以運用。草案中之鬆綁對於人工智慧之發展扮演著相當重要的角色。至今為止，對於資訊之提取僅在極少數的案例中是以用於人工智慧演算法的發展為目的。故此，許多個資並不允許用於此方面。第 54 條對於這樣的原則作了例外規範，然而僅僅在特定的情形方得適用。第 54 條規定了一連串的要求，唯有滿足了這些要求，方能討論是否有例外規定適用的空間。有趣的是，僅有特定的幾項發展目標被納入了例外情況的適用範圍之內：

「a) 為維護以下其中一項或多項領域中，極端重要的公共利益，發展『創新的人工智慧系統』：

i) 為了防止、偵查、發現或追訴犯罪行為或刑罰之執行，其中也包含了在主管機關的控制與職權責任範圍內，為了公共安全而對危險之預防與防衛，就此，應當要以各會員國之法規或是歐盟法規為基礎，處理個資。

ii) 公共安全以及公眾健康，其中包含了對疾病的防止、克服以及治療。

iii) 高度的環境保護程度，以及對環境品質的促進。」

因為有這些特定優先目標的設置，在管制手段上就得以利用真實實驗室的吸引力，藉此來引導創新。這樣的狀況也反映了一個現象，亦即公部門得透過不同的措施來對科技的發展產生影響。這些影響不僅是可以透過設置管制上的障礙，也能夠透過支持性的措施來達成。

在這些措施的框架內，到底應當考量那些目標，自然也經常是爭議之處。對於真實實驗室還能在那些部分給予放寬，第 53 到 55 條採取較開放的態度。這些都將會是進一步立法程序當中的關鍵問題。原因在於，在處理減少管制上阻礙以促進科技的議題上，永遠會涉及到一個問題，即：如何能在不發生損害的情形下實現科技帶來的合法權益。除了目的拘束原則外，就此問題並沒有其他明確的規範。因此，未來可以在人工智慧領域容許那些例外情況，以及這些例外情況的容許條件為何，尚需進一步的討論。

III. 展望

以上提及的四個問題，自非是在歐盟的立法程序中所討論的全部。然而從所談及的領域中可以對這個在歐洲相當重大迫切的議題有一個初步認識。這是從一個歐洲人的視角出發所作的討論，自然也僅僅是作為一個比較法上對話而踏出第一步。就這個議題我們亦可自問，人工智慧的議題會在哪一些論壇上受到關注。在聯合國的不同部門，例如網際網路管理論壇，對於此議題即有相當深入的討論，同樣的情況也發生在國際電信聯盟。事實上，目前也已經存在了一個初步的全球共識。聯合國教科文組織的 193 個會員國對於人工智慧倫理之倡議也達成了一致的見解[19]。特別是對於管制以及不同法規之間的相容性，還需要在意見交換上有更多的努力。法學上的，以及特別是比較法上的研究，為此就具有特別的重要性。在這樣的意義之下，希望能透過這篇文章對於促進相關議題的理解作出一些貢獻。

---

19 *UNESCO*, Recommendation on the Ethics of Artificial Intelligence, SHS/BIO/PI/2021/1, 2021（日期：23.11.2021，網址：https://unesdoc.unesco.org/ark:/48223/pf0000381137）.