《安全與情報研究》第七卷第二期　　中華民國 113 年 7 月
Security and Intelligence Studies, Vol. 7, No. 2　　July 2024

人工智慧對國家安全的影響與因應

趙萃文

東吳大學法律學系兼任助理教授

摘 要

人工智慧 (AI) 被應用在越來越多的領域，包括戰爭武器、國安情資，乃至選舉。生成式 AI 技術的發展，更提高了人們工作與生活的效率。惟 AI 技術亦可擾亂資訊判讀，影響人們思維；威權政體更用來鞏固政權。AI 所衍生的安全、法律與道德爭議，已成為大眾關注之議題。對國家安全而言，AI 自然需要健康有序之政策；要如何妥適建立監理防護機制，以確保能安全確當應用，即顯得格外重要。本文以 AI 技術帶來之國安挑戰切入，就其已產生或可預見危害為研究基調，因應數位發展，思考相對應的治理規範或手段，參考歐盟立法經驗，梳理 AI 科技創新下安全治理之政策。為建立兼顧開發創新與安全維護之發展方策，本文亦提出相應框架性規管策略，以作為對國安衝擊的因應。

關鍵詞：人工智慧、生成式人工智慧、資訊隱私、言論自由、假訊息

The Impact of Artificial Intelligence on National Security and the Countermeasures

Tsuey-Wen Chao

Adjunct Assistant Professor, School of Law, Soochow University

Abstract

Artificial intelligence (AI) is being used in more and more industrial fields, including war weapons, national security intelligence, and elections. The development of generative AI (Generative Artificial Intelligence) technology has improved the efficiency of people's work and life. However, AI technology can also disrupt the interpretation of information and influence people's mindset. Authoritarian regimes often use AI to consolidate their power. The security, legal and moral issues brought by AI have raised public concerns. For national security reason, AI has to be governed. It is important to establish proper supervision protection mechanism to ensure the safe and proper use of AI. This article considers the national security challenges brought by AI technology, and analyzes its existing or foreseeable harms. In response to digital development, the paper then discusses the governance norms or means by referring to the EU experience. In order to establish a regulatory policy for AI, while maintaining development innovation and security maintenance, this paper also proposes some regulatory strategies to mitigate AI's impacts on national security.

Keywords: Artificial intelligence (AI), Generative AI, Information Privacy, Free Speech, Fake News

壹、前言

人工智慧（Artificial Intelligence，以下簡稱 AI）為人類近年來發展出的變革性科技，對社會造成廣泛且巨大影響，如同當年網路被發明一樣。AI 被應用於製造業、金融、交通、醫療保健、能源與食品等關鍵基礎建設等越來越多產業領域，讓科技融入生活，帶來提升生活品質的美好藍圖；然而在 AI 的技術能力日臻進化下，無數以假亂真之 AI 圖片和影像在社群媒體流竄，在無法及時查證下，社會大眾很容易在「有圖有真相」直覺下，受到假訊息影響，對現行法制亦造成衝擊。

要確保 AI 發展與安全，同時有賴其他新興技術之穩定，如網路安全、智慧財產權與數據安全等，才能確保新興科技安全，避免產生更大的風險，導致國家社會安全發生漏洞。本文以 AI 技術對國家與社會的影響切入，就其已產生或可預見危害為研究基調，因應數位發展思考相對應之管理規範或手段，參考歐盟立法經驗，梳理安全治理之政策取捨與規範調適。諸如：AI 之發展是否應受監管，監管策略該如何妥適拿捏，以建構兼顧開發創新與安全維護之發展策略，提出框架性規管策略之建議，以作為對國安衝擊之因應。

貳、AI 對國家與社會的影響

隨著 AI 應用拓及各個領域，且逐漸普及於基礎設施及人類日常生活，而其技術內容與網路安全、智慧財產權、數據安全等息息相關。近來在 AI 的發展中，更出現了生成式人工智慧（Generative Artificial Intelligence，以下簡稱生成式 AI）的技術，

衍生若干問題。生成式 AI 則係指一種特定演算法，以 ChatGPT 為例，在於將輸入之文字資料轉換成另一些文字資料輸出，可以是提供使用者諮詢的具體內容，如對話、翻譯、查詢、延伸、歸納或解答等各種不同的關係。其餘之生成式 AI，可能以圖像轉圖像，文字與圖像之間互換，或更多如文字、影像、影片、語音、行動等各類資料模態間之轉換生成。[1]

生成式 AI 技術之運用，涵蓋政府、企業乃至個人創作等各領域。由於生成式 AI 技術之可擴展特性，可應用於各種數據查整與分析，並隨著資料量增加而提高生成效果，能夠生成高解析度之人臉和風景圖像，而模擬仿真之應用，更提供了創新方法來解決現實世界中的複雜問題。[2]

在中國大陸，AI 是一系列國家法令的一環，發展和應用 AI 旨在作為達成其政策目標的基礎，如控制中國境內產生的所有資料，並監控境內居民及各種戰略合作夥伴國的國民資料，其中一項法令，要求所有外國公司要將大陸人民的資料儲存在中國境內的伺服器裡，方便讓政府公安機構可隨時讀取。中國大陸也在輸出監控科技，並在過程中蒐集資料，用以強化其全球影響力，而全球能源互聯網更是習近平主席的國家戰略，旨在設立世界上第一個全球電網，由中國管理，以重新包裝形塑為稱霸全球的超級大國。[3]

---

1 王道維、邱筱涵，〈當 ChatGPT 來敲法官的門 — 淺談 AI 應用於司法審判的原則與方式〉，《當代法律》，第 18 期，2023 年 6 月，頁 54-55。

2 董慧明，〈生成式 AI 技術發展對國家安全的影響與挑戰〉，《清流雙月刊》，第 48 期，2023 年 11 月，頁 4-6。

3 艾美．韋伯著，黃庭敏譯，《AI 未來賽局》（新北市：八旗文化，2020

雖然 AI 有潛在的風險，然各國政策、法律與監管制度對 AI 之限制仍少，且幾未對其發展造成任何阻礙，但近年各國紛紛擬定相關法律規範，以降低 AI 發展的負面影響。以 AI 操控認知作戰為例，其帶來之威脅不容小覷，例如，俄烏戰爭期間，莫斯科透過社群媒體及網路輿情操作，以 AI 大量投送具特定目的訊息，試圖影響國際社會的觀點和看法，並透過操縱訊息，支持其在烏克蘭之地緣政治目標。此一操作，不僅削弱了民主政治，更對烏克蘭主權和安全構成威脅。

一、AI 與國安及治安息息相關

情報蒐集往往要整理與分析大量資訊，又得花費時間撰寫研析意見；生成式 AI 之豐富功能，有助於提升情報蒐集分析效率與品質。以公開來源情報蒐集為例，欲從宛如大海之資訊中，分析出具重要價值訊息，可導入生成式 AI 技術，快速分析、自動識別並提取關鍵資訊，讓情報單位能更快速精確地識別潛在威脅和線索機會，並藉由分析多種文本、圖片、影像與聲音等數據來源，提供相關事件、行為及趨勢之輔助判斷。目前以色列國安單位已採用生成式 AI 作為情報工具，打擊潛在國家安全威脅，甚至協助政府執法單位打擊犯罪。[4] 惟情報實乃需要嚴謹看待之資訊，其精確性不容有所差池，內容務須經過情報專業人員篩濾；對來路不明、無法保證的資訊，如果經反復檢核過濾仍無法解除其精確性疑慮時，卻提供決策者使用，恐充滿無法預測之風險。

年 3 月），頁 20、21。

[4] 董慧明，〈生成式 AI 技術發展對國家安全的影響與挑戰〉，《清流雙月刊》，第 48 期，2023 年 11月，頁 7-8。

資訊科技與資訊社會也助長了假訊息問題，許多公共政策被片面傳遞又難以查證；假訊息的氾濫，固然會對社會造成諸多不利影響，惟過度之管制手段，亦會對言論自由造成限縮甚至侵害。在許多類假訊息已經被刑罰化的我國，重懲之下是否對言論自由產生寒蟬效應，務須慎思。資訊科技同樣提供了吾人治理假訊息之「監理工具」，如《傳染病防治法》第 63 條散布疫情假消息罪，公權力得否運用此類技術偵查？以 AI 監管假訊息，其本身對於社會及人民之言論自由等基本權，是否亦產生一定程度風險？[5]

因為未來的戰爭將會用電腦程式碼來決定輸贏，而不是徒手肉搏戰，軍隊可以應用 AI 科技破壞他國的經濟穩定，不用靠摧毀農村和城鎮中心來「贏得」戰爭，從這個角度來看，中國大陸透過購買美國公司的股票，以避開美國的監管，並獲得美方可能是應用於軍事方面的技術的敏感 AI 科技，人民解放軍也大力投資一系列 AI 相關計畫和科技，[6] 其推展 AI 的全面發展，已領先西方國家，情況自是讓美國等民主國家感受到危險。[7]

科技必然會成為法律治理之工具，AI 運用將無可避免地進入法律治理系統，成為法律執行之主要工具；[8] 目前以 AI 作為治理假訊息工具之態樣，係以人工進行事實查核，再結合 AI、機器人

[5] 張文貞，〈2018 年憲法發展回顧〉，《臺大法學論叢》，第 48 卷特刊，2019 年 11 月，頁 1534。

[6] Phil Stewart, “China Racing for AI Military Edge over US: Report,” *Reuters*, November 27,2017, <https://www.reuters.com/article/idUSKBN1DS0GN/> (last visited on Jan. 12, 2024).

[7] 艾美．韋伯著，黃庭敏譯，《AI 未來賽局》（新北市：八旗文化，2020 年 3 月），頁 110、111、143。

[8] 李崇僖，〈人工智慧時代之資料治理與法制初探〉，《臺灣科技法學叢刊》，第 2 期，2021 年 7 月，頁 83。

與演算法，讓檢舉過程更為簡便，查核結果更有效益。而將 AI 真正應用在識別假訊息，爾來頗受國際社會關注之深度偽造 Deep Fake 技術，除以法律規範技術發展準則外，引進或研發相對應之偵測技術，讓管制者能有同等武器去對抗被濫用之新興科技與 AI，亦有論者提倡。[9]

因此，有論者指出假訊息監理科技要依據其自身風險程度不同，而受到科技治理或倫理規範之限縮；但科技治理對於假訊息管制有其重要性，「高風險監理科技」被運用在對抗「高法益侵害」之假訊息，建立以風險為基礎的分級管理法制框架，毋寧較能兼顧法治與發展的衡平，[10] 如將關涉疫情、國安的境外資訊戰、選舉不實資訊等「特別有害」之假訊息，歸類為高法益侵害，而予以較重刑責時，運用高風險之偵測、下架假訊息之 AI 或演算法來監理這類假訊息，甚至有讓政府針對特定領域假訊息利用這類科技加以監理，可具備更堅強的正當性基礎。[11]

在詐欺方面，有論者稱 ChatGPT 可能讓電信詐騙集團進化至下一階段之自動化，真正對受害者行騙的電話手將由機器人取代。值得注意的是，在防範社會安全事件，包括經濟安全事件、重大刑事案件及群體性事件等部分，因涉公共安全與社會秩序之維護，特別是群體性電信詐騙事件，已成為社會動亂根源，影響

[9] 余和謙，〈人工智慧之治理—以深度偽造為例〉，《科技法律透析》，第 31 卷第 8 期，2019 年 8 月，頁 70-71。

[10] 陳陽升，〈從法治原則探索人工智慧之應用界限〉，《臺灣法律人》，第 26 期，2023 年 8 月，頁 97-106。

[11] 李岳軒、林志潔，〈臺灣假訊息管制的未來展望—以規範與科技的互動為核心〉，《高大法學論叢》，第 18 卷第 2 期，2023 年 3 月，頁 57－108

社會安定與國家安全。[12]

為根除電信詐騙造成之社會問題，主管機關除研議與大型平臺業者及電信公司加強合作，建立反詐騙聯合通報機制，即時下架假訊息、追查犯嫌，亦可與來電辨識服務商合作，於顯示來電號碼時揭露可疑來源，將 AI 技術應用於防範詐騙，藉由 AI 與 AI 之對決，降低民眾受騙機率，[13] 亦可獲致 AI 革新帶來之科技紅利。

由於 Al 進步速度已超越人類想像，所產生之假新聞、假影音正在左右選舉結果影響政治;如無人機搭載自主武器造成殺戮，類此致命自主武器犯錯時，要由誰來負責？恐將會產生「問責缺口」(accountability gap)，針對美國在興都庫什攻打塔利班和蓋達組織之軍事作戰，網站《攔截》(The Intercept) 在 2016 年 11 月的調查指出，遭到無人機襲擊死亡者，每 10 人就有近 9 人並非預定目標，這仍是靠人類來發號施令的時侯，機器容或會太快行動，使人類無從介入並防止出錯；其實，武器的行為並非靠製造商來寫程式，武器係自行學習，而它會學到什麼，則端賴於它看到了什麼資料。因此，設若吾人透過立法來對致命自主武器之使用限制，AI 反而可能會拯救而不是奪走性命，如清理地雷區，協助提供人道援助，減少平民傷亡，及保護軍人免於受害等。[14]

---

[12] 朱蓓蕾，〈中國大陸應急管理體系之改革：風險、治理與挑戰〉，《安全與情報研究》，第 6 卷第 1 期，2023 年 1 月，頁 1-47。

[13] 劉奕成、葉柏廷，〈ChatGPT 問世五年內金融服務業的 AI 戰略〉，《當代法律》，第 18 期，2023 年 6 月，頁 95-96。

[14] 托比・沃爾許著，戴至中譯，《2062 人工智慧創造的世界》（臺北市：經濟新潮社，2019 年 10 月），頁 150、151、270。

美、中兩國刻正競相發展生成式 AI，並試圖應用於軍事層面，一旦時機成熟，可想像在戰場上生成足以混淆敵軍之圖像。AI 作為一種工具， 可以是國家用來強化自身安全的利器，亦可以是弱化他國安全之兵器。基此，AI 時代之國家安全，對 AI 之範圍及管理，必須認識到其社會複雜性之風險，不能單從已發生損害事件之經驗法則角度，而必須同時藉由「事件已發生、可能發生，甚至從未發生」的風險級別角度，來設計管理及預防因應機制，避免低（錯）估狀況發生。目前可行方法，係建置包括我國內部資料和與其他相似情況國家外部資料的 AI 數據資料庫，亦即讓我國 AI 系統在處理國家安全事務時，擁有全面性臺灣觀點，也有可交叉比對異同性之國際觀點。[15]

二、AI 極易遭到濫用

AI 發展若不受控，可能輸出有害或不實之結果，及洩漏機敏資料等；因此所帶來之風險，需要被進一步管理，否則一切唯利是圖，將可能肇致「失控的 AI 列車」，終至腐蝕國家根基，禍福莫辨。聯合國秘書長古特瑞斯 (Antonio Guterres) 指出，由於 AI 被濫用在網路攻擊、深偽、散播虛假訊息與仇恨言論，恐對全球和平與各國安全造成重大影響。他以社交軟體為例，原先用以增進人與人互動的工具和平臺，現今卻成為操控選舉、散播陰謀、煽動仇恨與暴力的地方。除 AI 系統存有潛在的不穩定或故障外，AI 結合核子武器、生化武器或機器人更是令人擔憂，不僅肇至既有政府組織架構政策推動上之窒礙，恐對全球和平與各國安全造

[15] 譚偉恩，〈AI 對國安的衝擊：勿高估中國威脅，莫低估未知風險〉，《清流雙月刊》，第 48 期，2023 年 11 月，頁 10-15。

成重大危害。面對當前之變局，必須從全球角度去重新思考因應戰略，故聯合國有必要推動建立新國際準則、簽訂新國際條約，建立相應之全球機構。[16]

三、AI 引發的言論自由保障議題

AI 問世，可謂人類邁向更美好未來的關鍵指標。然而，從 AI 改變訊息傳播模式和加速擴散的特性而論，一旦遭不當利用，精準定位社群媒體平臺，則會成為惡意行為者製造和傳散爭議訊息，及操縱目標受眾感知的不法工具。因生成式 AI 而發酵、擴散，甚至危及國家安全；例如劍橋分析公司協助川普打贏美國總統大選，讓英國脫歐等，即便先進的歐美亦難倖免，2023 年 6 月世界新聞媒體年會召開「生成式 AI 是媒體救星還是殺手？」座談會，當中有關生成式 AI 錯誤率過高、假訊息爆量議題更是關注焦點。[17] 美國 2022 年公布之《人工智慧權利法案》，內容對言論自由與國家安全的界線有所觸及，顯然 AI 造成錯誤與訊息擴散衍生之失控效應，已成為人類必須嚴肅面對的課題。

中國大陸對我進行認知作戰，可溯自胡錦濤時代，即曾要求對臺宣傳要達到「入島、入戶、入腦」。中共對我的資訊戰，隨著小紅書、TikTok 等社群平臺興起，更大肆對臺行「入腦」作戰，已獲致卓越成果。資訊戰對民主自由造成之危害，就是直接改變人的想法、侵蝕人心。而中共對我行認

[16] 鄭旭高，〈人類與 AI：共榮或毀滅？〉，《清流雙月刊》，第 48 期，2023 年 11 月，頁 22-26。

[17] 宋啟成，〈如何防患生成式 AI 失控？以第一次波灣戰爭凸顯的作戰型態為例〉，《清流雙月刊》，第 48 期，2023 年 11 月，頁 16-21。

知作戰最終目標，就是讓人民失去對政府的信心，民主體制就被挑戰，加上網路演算法、網路同溫層效應等，身陷謠言者將難以自拔，[18] 假訊息氾濫已造成國安隱患。

由於中共對我長期統戰與認知作戰之手段，均透過網路攻擊和訊息操作，試圖影響國際社會對我支持，並進一步挑撥國人情感，達成影響社會輿論及政府政策的雙重目的；如透過深偽技術進行換臉、輸入自訂提示詞生成特定圖像、影片，或運用演算法精確推薦自動化訊息，以情感化互動方式影響目標對象心緒，皆係認知作戰常見手法；當「有圖有真相」的說法被顛覆，分辨虛實也不再像過去那樣可靠，更凸顯謹慎使用 AI 之重要性。

雖然生成式 AI 結合大型網路平臺已成最新趨勢，但我國相關法制仍未盡完足，面臨錯、假訊息攻擊又遠較歐美嚴峻。期望立法者能模仿歐盟《數位服務法》(Digital Services Act, DSA) 之精神，結合當前國安情勢，以開放多元態度，強化反滲透法及國安法規之修調。此外，在立法防範同時，政府應帶頭與業界合作，集思廣益找出可能窒礙及解方；簡言之，AI 平臺自然需要健康有序之政策與法律，包括管理組織、資訊蒐集運用、資料安全、虛偽資訊、平臺監管及法遵倫理等，如何對其有適當的法律規制，或係非常困難之議題，但實為重中之重。[19]

由於假訊息會使人做出錯誤決定，侵害個人權利，破壞社會互信，使公共利益遭受侵害；網路平臺的特性使得假訊息在網路

[18] 吳柏軒、楊媛婷、謝君臨，〈假訊息入島入腦數發部、NCC 相互卸責〉，《自由時報》，2024 年 2 月 17 日，A3 版。

[19] 趙萃文，〈在AI取代人力的那一天—談AI技術革新下的法律挑戰〉，《清流雙月刊》，第 48 期，2023 年 11 月，頁 33-38。

上對個人權利與公共利益的破壞更為巨大，也讓政府有了管制之動機與壓力。[20] 美國政府已於 2022 年 10 月發布《人工智慧權利法案》(AI Bill of Rights)，希望藉由協助指導科技業者設計、開發與部署 AI 及其他自動化系統，藉此保護美國公民權利與安全。[21] 美國參議院聽證會，OpenAI CEO 奧特曼在會中表達生成式 AI 可能被誤用，導致假訊息及犯罪攀升，主張成立聯邦專責機構，制定使用標準及業者法遵義務。

瑞典哥德堡大學 2021 年報告指出，臺灣是接受境外假消息侵擾最頻繁地區，在社會及媒體兩個領域被中國大陸影響最深。[22] 美國在臺協會孫曉雅處長指稱，假訊息介入選舉之情況層出不窮，臺灣和美國都深受其害，唯有有效地反制假訊息之威脅，才能捍衛民主、捍衛國家安全。[23] 觀諸《國家安全法》第 4 條明定，國家安全之維護，應及於中華民國領域內網際空間及其實體空間；《電信法》第 22 條但書亦規定，電信事業對於電信之內容顯有危害國家安全或妨害治安者，得拒絕或停止其傳遞。依此，網路安全管理，乃至對假訊息散布，主管機關可主動通知電信業者為必要處置，電信業者亦可依主管機關告知或公告逕為處置；因此，現階段關鍵不在如何立法，而在如何執行。

[20] 許育典、李霽恆，〈網路平臺上假訊息的管制問題〉，《國立中正大學法學集刊》，第 75 期，2022 年 4 月，頁 215。

[21] 茅毅編譯，〈賀錦麗找科技巨擘研商 AI 風險〉，《聯合報》，2023 年 5 月 6 日，A10 版。

[22] 鍾麗華，〈社媒遭中國滲透臺灣排名世界第一〉，《自由時報》，2022 年 3 月 25 日，A04 版。

[23] 楊堯茹，〈孫曉雅：假訊息介選臺美都深受其害〉，《自由時報》，2023 年 10 月 18 日，A2 版。

參、AI 治理與規管的必要性

隨著 AI 技術在全球治理之廣泛應用，數據隱私及個人自由、AI 與人權、機器倫理與 AI 的責任等問題日益凸顯。[24] 愈來愈多部門將 AI 用於與國家安全有關領域，例如利用 AI 監控關鍵基礎建設等，為確保 AI 技術不會傷害人類，允宜加強推動 AI 系統整合，以確保能緊急啟動備援通訊系統，並納入高度數位韌性之設計規範，確保各項目在實際應用時具備最佳穩定性，俾政府單位能夠正常運作，避免依賴以 AI 進行決策之結果出現偏差，俾維持社會基本生活，[25] 達致「友好人工智慧」之系統安全所追求目標，AI 之倫理、信任及治理機制即顯重要。

要善用 AI 必須要有利他普世善念與價值，創造對話、共鳴與感動，確保 AI 發展能從人類福祉出發，達至有利於整體發展的社會環境。要達成上述目標，有賴以正當程序保障公民權利，提升透明度要求，測試評估的展開，說明理由之引入等，藉由正當程序原則之約制，避免公民權利遭受威脅；[26] 而面對網路平臺之法律問題，如機會平等、禁止歧視、資料可攜性等規範需求與保護機制，亦均應建構。

---

[24] 有關 AI 在全球治理的道德與法律挑戰，參見洪錦魁，《AI 和 ChatGPT 人類和機器共生的未來》，（臺北市：深智數位，2023 年 4 月），頁 19-21 ～ 19-26。

[25] 盧天麟、呂宜諠，〈發展自主衛星產業，強化數位韌性與通訊安全（上、下）〉，《工商時報》，2023 年 2 月 21、22 日，A06、07 版。

[26] Michael Moran, *The British Regulatory State: High Modernism and Hyper-Innovation* (New York: Oxford University Press, 2003), p. 154.

一、AI 之風險管理應從資料治理開始

經濟學人雜誌指出，資料係 AI 時代之新石油，[27] AI 之發展得仰賴足夠資料量來進行篩選與判斷，資料既是 AI 技術數位平臺之驅動元素，是資料之保護自應為 AI 之內在基因。AI 之創新產品或服務，需要良好的資料治理環境，資料之可信與取得資料之手段，均將影響後續技術和服務運作之正當性與合理性。[28] 惟資料來源多而分散，在取用及輸出過程中，應有適切之管控治理，以利其合法取得高品質、多樣化資料，爰 AI 之風險管控即應從資料治理開始。

（一）建立良好的資料治理環境

AI 之創新產品或服務，需要良好的資料治理環境，以利其合法取得高品質、多樣化資料。資料治理法制化之目標在於建立一個安全、透明且效率的資料使用和共享環境。要防範認知作戰，主要作為即是重新審視資訊安全之重要性與防護措施，此於眾多國際衝突與戰爭中，屢屢獲得印證，爰此，AI 系統應有防止資訊及網路安全遭到破壞之防禦機制；而個人資料之保護亦屬資料治理一環，[29] 在保護資訊隱私權及個人資料自決權之前提下，建立

---

[27] "The world's most valuable resource is no longer oil, but data," *The Economist*, May 6, 2017, <https://myppt.cc/Z4piA0/> (last visited July 30, 2023).

[28] 有關不法取得數位資料的刑法回應，參見蕭宏宜，〈網路釣魚的刑事爭議問題—以資訊取得為中心〉，收錄於蕭宏宜，《科技發展與刑事立法》（臺北市：一品文化，111 年 1 月），頁 308-325。另參見顧振豪，〈生成式人工智慧與法律的和諧共舞〉，《當代法律》，第 18 期，2023 年 6 月，頁 51-52。

[29] 由於資料當事人之更正權、刪除權等重要權利，均奠基於知悉其個資被

完善特種個資保護規範，以及各種明確化之公共利益為目的所為資料蒐集、處理及利用之具體化機制，有必要為 AI 量身定製一套個資蒐集規制，[30] 從而促進資料創新應用和價值實現，同時保護個人隱私和資料安全。

（二）推展網路守門人概念

由於生成式 AI 需要高速運算能力及巨量資料，技術及資源均掌握在大型業者手裡；渠等大多贊成管制，加諸業者法遵義務，等同提高營運門檻，形成穩定寡占。凡此，本該為人們創造更多競爭與更好選擇，但事實是讓科技巨頭們擁有巨大影響力。[31] 隨著近代監視社會之到來，網路守門人此一概念，為歐盟數位市場法和歐洲法院判決所採納，大型網路平臺掌握資訊控制權，其縱令並非傳統公共投資等必要設施，但亦接近人民必須使用之基礎設施，自應承擔更高義務。[32] 賦予企業體自我管理義務 (policing duties)，責令法人制定法遵行為準則與企業倫理計畫，主動揭露不法行為，形塑並執行有效內部自律，以預防或減低企業不法或犯罪行為。[33]

何人應用？被如何應用？方能有效行使權利之前提上，為免資料當事人權益遭受侵害，使其得以知悉其個資接收者確切名稱，用是能落實資料當事人的近用權。

[30] 李建良，〈元宇宙與法秩序〉，《臺灣法律人》，第 8 期，2022 年 2 月，頁 38-40。

[31] 張瑞雄，〈當企業比政府更有權〉，《中國時報》，2023 年 9 月 23，A10 版。

[32] 楊智傑，〈網路選舉宣傳揭露資助者、外國勢力與言論自由〉，《憲政時代》，第 46 卷第 4 期，2023 年 1 月，頁 529。

[33] 溫祖德，〈法人犯罪量刑與法令遵循〉，《刑事政策與犯罪防治研究》，第 24 期，2020 年 4 月，頁 115-155。

二、安全與信任為 AI 效能之基石

全球智慧化時代快速到來，但在借助 AI 系統同時，AI 的應用存在安全、歧視和隱私風險，[34] 自應納入法治軌道。為建立合法、安全與可信賴之的 AI 系統，歐盟執委會人工智慧高級專家小組 2019 年提出《可信賴的人工智慧倫理準則》(Ethics Guidelines for Trustworthy AI)，來打造可信任生態系，落實在 AI 整個生命週期的持續稽核，藉由遵守人性、法治國、自由權、平等及不歧視等各項倫理原則，建立可信賴 AI 基礎，並透過「以人為本」的作法，達致讓 AI 在合法之外，猶能符合倫理及穩健之高品質要求，[35] 以促進 AI 之開發與創新。該準則要求 AI 須遵守行善、不為惡、保護人類、公平與公開透明等倫理原則，具體描述可信賴的 AI 包含七大面向：即人類自主性和監控、技術穩健性和安全性、隱私和資料治理、透明度、保持多樣性、不歧視和公平、社會和環境福祉及問責制。[36]

此外，歐盟於 2023 年 3 月分別通過《人工智慧法案》(Artificial Intelligence Act)，希望降低因 AI 技術快速發展產生的族群偏見、侵犯隱私與其他各種風險，鼓勵各國藉由立法規範及管制，負責任且包容地發展 AI，避免不適當或惡意的設計、發展、

[34] 有關人工智慧治理的風險及相應的制度規範之道，可參閱宋華琳、孟李冕，〈人工智慧在行政治理中的作用及其法律控制〉，收錄於臺灣行政法學會主編，《跨越傳統與現代科技之行政管制與執行法制》（臺北市：元照出版社，2019 年 9 月），頁 269-294。

[35] 林昱梅，〈預防原則於人工智慧之適用—歐盟人工智慧規則草案及其基本原理之觀察〉，《臺灣法律人》，第 12 期，2022 年 6 月，頁 1-17。

[36] European Commission, Ethics Guidelines for Trustworthy AI, <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477>.

部署和使用，削弱對人權、自由與生命的保護，共同提升 AI 系統的安全及可信度。

（一）重視 AI 之風險

ChatGPT 介入人們的生活愈趨廣泛，已然成為生活內容的新常態。然此大數據科技對用戶電子郵件、電話通訊及網聊天紀錄進行大規模挖掘、疊加與整合，將那些散落於數位世界中，零碎的、本無意義的巨量數據予以系統化使用，讀出非常驚人之個人身分特徵，人民之隱私權即可能遭受侵害，對資訊安全造成相當程度侵蝕，現有法律護衛資訊隱私之保障功能，恐將漸次喪失。[37] 號稱 AI 教父之辛頓 (Geoffrey Hinton) 振臂疾呼，呼籲各國政府應該認真看待 AI 之風險。[38]

ChatGPT 引發隱私、資安、假訊息等疑慮，義大利前曾宣布禁止 ChatGPT，指控 OpenAI 在缺乏法律基礎下，大量收集儲存個資，來訓練 ChatGPT 之演算法，西方國家禁用主要係考量「隱私」，惟從「資安」角度觀察，公務員以公務設備使用 ChatGPT，在對話過程中上傳機密或敏感資料，恐成為我國資安漏洞。就數位治理與法制重構觀之，非政府機關乃至民間機構如持有人民個資，亦宜本諸維管辦法之授權依據與拘束力，就其「資訊受託人」之善管義務、忠實義務、保密義務等擔保責任，為更細緻多元、理性完善之規制，建立可信任人工智慧基礎，以達致

[37] 林昕璇，〈論大規模政府監控之資訊隱私保障—評析美國聯邦法院相關判決〉，《臺灣民主季刊》，第 17 卷第 2 期，2020 年 6 月，頁 58、59。

[38] 參見〈人工智慧的美麗與憂愁〉，《工商時報》，2023 年 7 月 5 日，社論。

護衛資訊隱私之保障功能。

（二）建立可信任之人工智慧基礎

資訊權作為憲法基本權利，其保障包括人與事之範圍。目前用於 AI 創作之主流 AI 技術，所運用之原理與大數據及資料分析有關，在創作歷程中所扮演之角色，大抵為人類之輔助工具，雖然能力較傳統工具強，但真正主導者是人，在將人類作為權利主體之基礎下，或未挑戰現行認定內涵，僅制定管控 AI 之規範即可。但各國大抵未擬定相關監管政策或規制[39]，未能形成一套完整之數位法制規範。[40]

學者馬可尼在《新聞製作者：AI 與新聞學的未來》一書中指出：AI 帶來的其實是個「人機協作」之世界，[41] 由 AI 執行例行性之優化任務，人類則是負責執行需要創意和策略思維的工作。[42] 現下 AI 系統從如何取得及輸入個案數計、風險群組之建立與正確性之評估等事項，均取決於人的價值判斷及主觀決定；具價值判斷或感知因素之決定，AI 至多僅係輔助性

[39] 許力儒、莊弘鈺，〈人工智慧創作之著作權適格與歸屬—法律與技術之綜合觀點〉，《萬國法律》，第 241 期，2022 年 2 月，頁 28-30。

[40] 歐陸法學者 Giovanni De Gregorio 曾將歐盟國家與社會中，數位科技轉型與法制規範互動之發展歷程，區分為數位自由主義、司法積極主義、數位憲政主義等三個階段，參見劉汗曦，〈從數位憲政與數位信任看我國健保資料庫的爭議與使用〉，《月旦法學》，第 331 期，2022 年 12 月，頁 49-51。

[41] 游梓翔，〈AI 衝擊傳播業，你是 80 還是 20 ？〉，《聯合報》，2023 年 6 月 26 日，A10 版。

[42] 李開復，《AI 新世界》（臺北市：遠見天下文化，2019 年 5 月），2 版，頁 361。

質。[43] 因是，目前之法制基礎是否足以解決 AI 所衍生之法律爭議，且能夠有效地執行，是否有制修規範之必要？至關重大。

如何設計新型態之監管和治理措施，發展出新一代價值秩序，關鍵即在數位信任之建構與維持，及寬嚴適中的規範要求。信任就是「願意承擔風險投入行動，來追求所期待的他方行為與結果」；而數位信任指涉的應該是，「在數位平臺上，基於對於特定人或機構行為結果的期待，而願意承擔風險來提供資訊或進行交易」。數位信任的重心應放在四個柱石：倫理與責任、隱私與控制、透明與問責、安全與可信。凡此，均可作為制度設計參據，確保相關立法及機制符合公開透明與民主課責。

三、AI 的風險與因應

民主社會之開放性質，雖促進資訊之自由流通與觀點之多樣性，在一定程度上，強化了國家之溝通及社會之凝聚力；但高度量化時代，數據假象充斥，統計、圖表、懶人包，常是理性裝扮的謊言，點讚、分享、演算法，助長「類事實」瘋傳成禍，為敵對勢力提供潛在機會，使之能加以利用，進行認知作戰。因此，處在假新聞、不實或誤導資訊充斥的深度偽造時代，數據識讀成為最重要之思辨素養，亦是抗制數據假象之基礎。[44]

---

[43] 李榮耕，〈刑事程序中人工智慧於風險評估上的應用〉，《政大法學評論》，第 168 期，2022 年 3 月，頁 178、179。

[44] 卡爾．伯格斯特姆、杰文．威斯特著，穆思婕、沈聿德譯，《數據的假象》（臺北市：天下雜誌，2022 年 12 月），頁 8-13。

根據微軟近期發布的網路威脅報告，揭露中共認知作戰之策略與操作，亦即以國家之力量對外國選民進行網路影響行動，此類秘密之數位行動涵蓋各類型錯假訊息在不同平臺散播，特別是應用生成式 AI 工具，其策略展現比以往更先進更細膩。微軟指出，中國大陸自 2023 年初就開始試驗 AI 生成之新聞主播，今年 1 月我國總統選舉期間，各種複雜細緻之 AI 內容遽增，包括 AI 創造之假影音片段在網路中散播，經由數位跡證之辨識技術，檢視出主要出自於中國大陸虛假訊息網路之手，微軟將其命名為「網路垃圾偽軍」。由於今年是全球大選年，包括美國總統大選在內，西方情報安全官員也公開表示，愈來愈擔憂各國選舉將充斥 AI 工具製作之誤導性影片或其他數位內容，藉以影響選民投票行為，改變選舉結果。[45]

中共對我統戰傳播之重要管道，主要藉由「海峽導報」等官媒，及「坐井觀天」等代理人隱身帳號，鎖定國內特定族群傳播具認知影響之謬誤影音，假言論自由之名行傷害民主之實；其操弄手法包括：否定臺灣主體、貶低臺灣盟友、分化臺灣民眾，說好中國故事及形塑符合中國利益的華語世界觀等；技術上則持續更新，結合時事「以臺批臺」，全面介入公共議題，特別是在生成式 AI 技術上，產製出更多元、細緻且判讀難度極高之跨媒體宣傳內容，對高度仰賴社群的臺灣民眾，將淹沒在龐大錯假訊息

[45] 我國立法院於 2023 年三讀通過《刑法》增訂第 28 章之 1《妨害性隱私及不實性影像罪章》及《犯罪被害人權益保障法》，對製作或散布不實影音檔案者量刑；修正《總統副總統選舉罷免法》（第 90 條）和《公職人員選舉罷免法》（第 104 條），明定意圖使候選人當選或不當選而散布、播送候選人之深度偽造影音，最重處 7 年以下有期徒刑。

之中，破壞國人對於民主制度的信任。

從近期的俄烏戰爭及以巴戰爭中，發現深偽技術的普遍運用，在國家衝突中，充斥許多假造影像，成為製造仇恨之敵對工具，在網路推波助瀾下，加深雙方誤解，嗣後縱經公正媒體解析認定確實造假，但對於真相而言早就於事無補。智慧科技從網路虛擬進入真實世界，遭不法轉為操弄認知之用，禍害程度之高，不僅將戰亂帶至更深層之恐懼，更成為擴大衝突的藉口，泯滅人性，淪為暗黑科技。[46] 因此，民主國家更需要推動跨國資訊合作，遏止類似錯假訊息傳散，以確保言論自由與國家安全。[47]

（一）科技倫理應與 AI 原則互補

法律秩序之建構僅居於表層，德行倫理 (virtue ethics) 則可以相當程度反映出深層文化之內涵，毋寧是一種人際間信賴之核心價值；AI 發展要經宏久遠，無法光靠世俗國家律法所保護，更須加上價值觀等「倫理」信念之維繫，方屬穩定發展之要件。因此，德行倫理應該與 AI 原則規範互補，共同建構完善之 AI 基本規範。倫理準則之位階秩序與體系脈絡要精準掌握，唯有將德行倫理要求內化為倫理規範，成為個人行為與態度之一部分，同步積累，才是數位環境永續發展之最佳實踐。

[46] 林臺森，〈強化識讀能力，嚴防中共深偽入侵〉，《青年日報》，113 年 4 月 16 日，14 版。

[47] 蘇紫雲，〈中共操弄認知作戰，「垃圾偽軍」為禍全球〉，《青年日報》，113 年 4 月 15 日，11 版。

（二）自我管理及政府介入並行

AI 技術可加速假訊息之生成與傳播，目前各國在對於 AI 監管與立法模式尚無共識之情況下，我國現行主要以《社會秩序維護法》第 63 條第 1 項第 5 款「散佈謠言，足以影響公共之安寧者」處罰，有必要加重處罰，或明確生成式 AI 假訊息之態樣；亦可參考歐盟《數位服務法》第 9、10 條規定，原則上平臺業者自我管理下，政府若發現平臺上資訊明顯違法時，仍然可以介入，以命令方式要求平臺將違法內容資訊取下；另於法院裁決或行政機關依法責令中介服務者提供相關協助，亦必須盡快配合。楊智傑教授亦稱，若真要制定一個法律要求網路服務業者對假訊息散布負責，必須對所需要處理的言論類型清楚定義，才能對網路平臺要求較為嚴格之措施與責任，[48] 可謂相仿的信號傳遞。同時 DSA 針對用戶數超過 4500 萬人之超大型平臺業者，設計諸多行政檢查義務，如透明報告義務、具公信力之檢舉人機制，均值得我國仿效推行。

（三）參考歐盟與美國之立法經驗

歐盟 AI 法草案雖在幾個關鍵議題上爭論仍多，如關於不可接受風險、高風險 AI 規定詳盡，但卻明文排除可能更嚴重侵害歐盟基本價值觀及人民基本權利之軍事用 AI，有論者以為應有必要思考於 AI 法或國際法層次為規範。再者針對即時生物辨識，歐洲議會主張全面禁止即時人臉辨識系統，保留非即時人臉辨識

[48] 楊智傑，〈歐盟與德法網路平臺假訊息責任立法比較與表意自由之保護〉，《憲政時代》，第 45 卷第 1 期，2019 年 7 月，頁 45。

針對恐怖攻擊、失蹤兒童等，部長理事會則主張將執法和邊境用 AI 脫離高風險管控，惟執委會激烈反對；此外由於科技中立一般不會納入管制，歐盟 AI 法案不僅美國企業如 Google、OpenAI 強烈反彈揚言退出歐洲市場，2023 年 7 月數百間歐洲大企業向歐盟聯名反對納管 AI 基礎模型，主張回歸本諸風險而管理。[49] 據報導，歐盟在 2023 年 7 月更派出官員到亞洲國家（包括日、韓、印度及新加坡等），說服採取類似歐盟之 AI 管制立法，但反應相當冷淡。大多數國家採取先觀察 AI 科技發展再進行規管的態度，部分國家則傾向採取較為寬鬆彈性的管制。[50]

美國 2022 年 10 月公布之《人工智慧權利法案》，旨在保護民眾個人數據不被 AI 演算法濫用，雖還不具法律拘束力，亦尚未達到歐盟具里程碑意義之隱私法規標準，但至少為此一規範尚處模糊的領域帶來一個清晰、可遵循之框架。本案強調：(1) 安全有效的系統；(2) 演算法歧視保護；(3) 數據隱私；(4) 通知與解釋；(5) 人類替代方案、考量與取消等五個風險案例與保護原則，

[49] iThome, "Open letter to the representatives of the European Commission, the European Council and the European Parliament," <https://www.igizmo.it/wp-content/uploads/2023/06/Open-Letter-EU-AI-Act-and-Signatories.pdf> (last visited Jan. 12, 2024); Michelle Toh, "Serious concerns: Top companies raise alarm over Europe's proposed AI law," <https://edition.cnn.com/2023/06/30/tech/eu-companies-risks-ai-law-intl-hnk/index.html> (last visited on Jan. 12, 2024).

[50] Reuters, "EU wants AI Act to be global benchmark, but Asian countries are not convinced," Jan. 12, 2024, <https://www.scmp.com/tech/tech-trends/article/3228050/eu-wants-ai-act-be-global-benchmark-asiancountries-are-not-convinced?module=perpetual_scroll_0&pgtype=article&campaign=3228050> (last visited on Jan. 12, 2024).

用以指導自動化系統的設計、使用和部署，保護美國民眾免受威脅，但僅作為企業及立法者之指引，不具法律實質效益。

四、創新與管制兼顧之政策

人類生活正在朝數位模式轉型，各種數位創新除了風潮趨勢外，最主要原因在於政府以政策、法規及預算在後支撐推動。如果 AI 是一項和醫藥發展一樣重要的科技，值得在我國落實其養成或訓練，並且有充分理由相信其對我國家發展充滿機遇，那麼監管就確實有其必要，需要一個明確的推進與管理規制。因是 AI 智慧科技既已成為網路犯罪集團之攻擊手法，AI 平臺需要健康有序之治理，包括管理組織、資料收集、資料運用安全、虛偽資訊、平臺監管與法遵及倫理之處罰等，[51] 兼顧開發資料價值之同時，又能保障個人之人格權益，以確保其安全、效能與品質，[52] 這毋寧是 AI 基礎法制架構之立法難局。

近來歐洲法院亦常援引《歐盟基本權利憲章》(Charter of Fundamental Rights of the European Union) 對於歐盟人民之保障，於數件數位平臺有關之判決書，賦予符合當代思潮與人權保障新定義，兼亦促進數位生態的發展。[53]

---

[51] 陳春山，〈元宇宙產業及智財法制前瞻〉，《萬國法律》，第 246 期，2022 年 12 月，頁 74-78。

[52] 歐盟之《資料治理法》建置了一套值得信任的資料治理規範，確保受保護之權利、個人資料與秘密性不受侵害，亦為資料經濟及人工智慧之發展奠立穩定紮實的法律框架。參見林昱梅，〈歐盟資料治理法制新趨勢，以公部門及資料利他組織之資料再利用為中心〉，《臺灣法律人》，第 25 期，2023 年 7 月，頁 96。

[53] 張文貞，〈跨國憲政主義的合縱與連橫 — 歐洲人權法院及內國憲法法院關係初探〉，《月旦法學》，第 151 期，2007 年 11 月，頁 57-70。

（一）避免科技成為威脅

鑒因 AI 可能提升資安及國安風險，人類應該思考 AI 之研究、設計和軟體部署都應優先考慮，改善人類生活品質願望這個「以人為本」的基本價值觀，以避免科技邁向邪惡。馬克尼 (Kevin Macnish) 指出，近來對民主最大挑戰取決於吾人是否能發展出新的能力與保護方式，俾期對大數據之拓展與利用，能夠在適當管理下進行，[54] 即在提醒人類需要掌握控制權，由人類來掌控數據，而不是被數據掌控。另 AI 產品製造人需要達到之產品安全性標準，應對齊目下之專業知識、技術規則、科技水準，均符合可能且可期待的水平；讓法律有效性及實效性兼具，讓法規範建構與科技能更有效連結。[55] 但應注意的是，僅只遵守法律面相關規範或相關標準並不意味已符合上述要求，蓋這些規範和標準亦可能已過時，甚或是出現設定當時尚未慮及的危險。[56]

（二）AI 嚴禁發展事項

國際紅十字會在 2016 年關於「自動化武器：殺人與摧毀之決定是人類的責任」聲明中，即強調「有意義的人類控制」，應指對於自動化武器「所執行之任務、攻擊目標、行動環境、地理

[54] 見 Kevin Macnish, Big Data and Democracy, Edinburgh, UK, Edinburgh University Press Books, p.1(2020). 又參見陳起行，〈初探大數據與人工智慧對未來立法影響〉，《月旦法學》，第 333 期，2023 年 2 月，頁 81-82。

[55] 陳起行，〈論人工智慧時代演算法爭議諮詢及調解〉，《月旦法學》，第 325 期，2022 年 6 月，頁 109-111。

[56] 魏伶娟，〈人工智慧浪潮對民事責任建構的挑戰—以智慧醫療器材之應用為例〉，《中正財經法學》，第 25 期，111 年 7 月，頁 24。

空間與行動時機，都必須施以嚴格的行動限制，促使人類保有監督武器體系以及在必要時予以關機的能力。」毋寧就是公共良知、人性本質之原則條件[57]，這嚴格的人類決策與介入，無非要三思後行，謀而後動。

歐洲議會於 2023 年 6 月中通過之《人工智慧草案》」中，將潛意識操縱、社會信用評分、大規模遠程人臉辨別列為 AI 領域嚴禁發展事項。就是從功能上為生成式 AI 設限，即便失控亦可將損害降至最低限度，甚而消弭於無形。[58]

美國及韓國皆特別重視資訊及國安之風險管理，針對重要基礎設施及軍民兩用模型制定特別規範，如美國政府發布之《第 14110 號行政命令：關於安全、可靠且可信賴地開發和使用人工智慧》(E.O. 14110: Safe, Secure, and Trustworthy Development and Use of AI)，[59] 責令軍民兩用模型之廠商有義務在美國商務部註冊登記。國家安全顧問蘇利文 (Jake Sullivan) 表示，美國有意建立一項全球協議，以明確規範 AI 永遠不能在沒有真人參與決策下動用核武，[60] 此確為不可忽略之國安思維。2023 年 11 月 1 日在

[57] 黃居正，〈與人工智慧相關的國際法議題—從國際人道法到生命體法〉，收錄於劉靜怡主編《人工智慧相關法律議題》（臺北市：元照出版社，2019 年 3 月），頁 233、234。

[58] 宋啟成，〈如何防患生成式 AI 失控？以第一次波灣戰爭凸顯的作戰型態為例〉，《清流雙月刊》，第 48 期，2023 年 11 月，頁 233、234。

[59] E.O. 14110, “Safe, Secure, and Trustworthy Development and Use of AI,” October 30, 2023.

[60] 中央社，〈白宮國安顧問：美會嘗試拉中國進限武談判〉，2023 年 6 月 3 日。其實，美國推動《人工智慧政治宣言》雖不具實質約束力，卻是一套基於人類安全的原則；其宗旨在於無論 AI 管理何種武器攻擊鏈路、決策指揮程序，只要是攸關人命生死的判斷，仍應由人類下達最終決定，

倫敦召開全球首場 AI 安全峰會，英、美、中國大陸等 28 個政府代表及歐盟簽署通過針對管制 AI 之《布萊切利宣言》，就因應 AI 安全議題之急迫性達成共識，各國咸認最強大 AI 模型功能，可能伴隨嚴重甚且具毀滅性之潛在風險，應強化全球合作。[61]

肆、AI 治理與規管的建議

AI 技術影響全球，協助人們用以打造一個更人性化的世界；但缺乏規管之 AI，讓人類的能動性、尊嚴及人權，均處於嚴重風險。AI 要經宏久遠，就必須加以規管，但要如何具體落實於法律規定，允宜進行與社會大眾之溝通，以利形成共識，讓監理制度確實達到保障 AI 使用者之實績，歐盟法制已形成相當穩固之標準，當可援引做為我國借鏡。

一、完善監理規範建立新的國際準則

可信任的 AI 是確保人工智慧被善用之關鍵，但目前針對 AI 技術，各國及國際組織大多以政策文件宣示倫理原則，目前僅歐盟、加拿大、巴西及中國大陸制定全面性監管專法，[62] 缺乏可執行之全球安全標準。在解決策略上，各國專家一致認同必須建立

以確保 AI 自主性武器系統不會殘害生命。

[61] AI 自主性武器、輔助軍事系統等執行任務之實際運用與規範，不僅作戰方式改變，亦關涉軍事科技發展，國際政治與國家安全，國際社會對武裝衝突之認知變革，及《國際人道法》相關規制，牽動多方動態平衡，亟待人類社群在軍事 AI 研發運用上，制定兼顧必要性、正當性及合法性的有效監管機制。

[62] 有關生成式 AI 之監管方向，可參見周晨蕙，〈生成式 AI 監管方向初探—美、歐、日之路徑分析〉，《科技法律透析》，第 35 卷第 10 期，2023 年 10 月，頁 63-70。

跨部會、跨領域、跨界別之監管委員會，負責制定相關法規、標準、指引和監督機制，以規範生成式 AI 之設計、開發及應用。

2024 年歐洲理事會公布《AI、人權、民主與法治綱要公約》(Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law) 草案，當 AI 之決策或應用違反人權與基本自由時，締約國應提供符合國際法及其國內法之程序保障。[63] 唯有透過加強治理，包括提升偵測、鑑別、追蹤和查核技術水準，建立公開透明資訊平臺，並隨時公布其相關資訊和風險評估，方能提升民眾「數位韌性」，以對抗如資訊迷霧般之外在環境。

二、規範建構與科技研發作有效連結

國家之任務是保障人民自由發展，制度之建構必須能讓人民在心靈上產生對制度之信賴。AI 技術發展快速，但其效益應該是追求經世濟民，讓人類永續和平發展，使人類生活永遠幸福安樂；從而 AI 的一切規制，必須有用、有功能，才能取得價值。因是，對提問或指令所作出的回應可能有偏差甚或不正確，利用人必須有判斷能力，不得任意利用。[64] 爰此，視不同運用需求及風險程度，回歸各法規主管機關，以不同政策或措施作為監理工具，才能適切兼顧規範建構與科技研發有效連結。

[63] 張腕純，〈歐洲理事會公布之「AI、人權、民主與法治綱要公約」草案〉，《科技法律透析》，第 35 卷第 10 期，2023 年 10 月，頁 18、19。

[64] 范姜真媺，〈生成 AI ChatGPT 之運用與個人資料保護〉，《月旦法學》，第 341 期，2023 年 10 月，頁 34。

AI 立法之基本原則，宜儘量將各種產品或服務均納管，但對風險較小產品應給予較寬鬆對待。基於風險評估，AI 研發監管思維應著重確保軟體製造商在「開發過程」各階段是否遵守規定，而不是放在最終產品本身，且對高風險產品而言，在開發過程有外部專家監督，熟悉整個開發過程，會比開發完成後才由外部人員檢視產品服務技術來得更有效益，期能在鼓勵科技創新之同時，達到管制科技風險之政策效果。

三、歐盟立法兼顧創新與控管可為借鏡

歐盟 AI 法立法所建立的風險等級，依不同等級課予業者不同義務，兼顧創新及控管，已打造可資借鏡之典範，亦較契合我國國情。由於目前大型 AI 公司幾全由美、中兩大強權獨占，在臺美經貿往來頻密下，我國未來 AI 立法應綜合考量產業發展與政策目標，選擇最適合監管模式，[65] 相較歐盟之強力監管，或可張弛有度，適度減輕業者行政義務，對中小企業更應考慮輔導及補助，庶免扼殺我國 AI 新創發展，並維持臺美經貿良好關係。[66]

隨著 AI 科技持續演進，本文僅係初步探究嘗試，試圖勾勒 AI 對國家安全之衝擊與因應概貌。AI 無法取代人，其至多僅能部分取代人，開發人機協同增強智慧，加強人機協作交互能力，讓 AI 之發展，在「安全可控、為我所用」下去開展；以各種措

[65] 行政院於 2023 年 5 月邀集國科會等部會共同研擬《人工智慧基本法》草案，嗣因考量 AI 技術發展快速，應用面太廣而延後立法時程，但行政院業於同年 8 月 31 日通過「公務機關使用生成式 AI 參考指引」。

[66] 趙萃文，〈我國 AI 基本法立法宜緩不宜急〉，《自由時報》，2023 年 8 月 3 日，A14 版。

施來確保，讓 AI 所帶來的好處是有全體共享，而不只是科技之擁有者。當然，其中仍存在諸多問題有待吾人繼續觀察 AI 之發展，開啟各方對話，並結合實踐做進一步探討。

四、強化媒體識讀因應認知作戰威脅

臉書執行長祖克柏稱：為因應假新聞，解決方式之一就是 AI；臉書要過濾內容，唯一希望就是靠智慧演算法，AI 或許會有助於識別假新聞，但也可能讓問題惡化，因演算法就能產出假新聞，且隨著產出假新聞之演算法變得更聰明，要分辨真、假新聞就會愈來愈困難；終究來說，「真相」將是這場戰爭之受害者。[67] 因此，要防止錯、假訊息危害，避免受認知作戰影響和操控，澄清總是緩不濟急，受眾應強化媒體識讀素養，主動檢視、查核、深究資訊餵給，而非被動全盤接收，讓謠言止於智者；此外，累積不同領域知識與經驗，建立一個可以信賴之專家網路，據以交換資訊破解假象，打造一個動態而堅實之腦細網路，用以維護民主社會的安定和健康發展，保護國家免於受到外部勢力之影響。

五、建構我國 AI 的治理準則

隨著 AI 的廣泛普及，例如自駕系統與金融系統等，一旦遭遇網攻或故障，將因其普及度而導致損害如同癌症蔓延至整個系統中。因此，要確保 AI 發展與安全，如何降低脆弱性，藉由負責、公平、可靠與可治理的監控 AI，完善其風險管理，政府對 AI 監管措施將十分重要，應以循序漸進方式，納入國家基礎建設範圍，

[67] 托比．沃爾許著，戴至中譯，《2062 人工智慧創造的世界》（臺北市：經濟新潮社，2019 年 10 月），頁 234。

避免科技發展失控，同時有賴其他新興技術的穩定，例如網路安全、智慧財產權與數據安全等，避免導致國家關鍵基礎設施安全漏洞。

無論是 AI 之惡意使用或認知作戰威脅，皆是利用個人和社會的認知脆弱性來達到目的。因此，因應 AI 治理之需求，政府、媒體和科技公司除應盡其責任，構建公私民協力抵禦威脅之安全機制，亦須著重行為法律規範之制定，及時更新法律制度，以及在使用和開發 AI 過程中，遵循對人權、個人隱私之尊重和保護，以促進社會公正等更高道德標準，打造值得信賴之 AI 環境。另在降低認知作戰危害方面，更要注重事實即時查核澄清、資訊倫理與理性思辨，強化全民心防韌性。

我國在晶片到系統平臺供應鏈擁有領先地位，更應該關注人文領域的資源，發展符合多元、民主價值的 Al 模型，讓臺灣成為多元和包容的文化中心。NVIDIA 執行長黃仁勳說臺灣是 AI 最重要的國家，但重要性指的是晶片生產的硬體優勢，至於後端商機更大的，如繪圖、廣告行銷、程式設計等軟體部分，屬於文化的層面，還需努力拓展商機。

針對 AI 安全治理與隱私保護，目前數位發展部已採取下列措施，協助業者在確保 AI 安全與個資保護的前提下發展 AI：

（一）依據《個人資料保護法》於 112 年 10 月訂定《數位經濟相關產業個人資料檔案安全維護管理辦法》，要求業者訂定個人資料檔案安全維護計畫及業務終止後個人資料處理方法；並針對數位經濟相關產業，如資訊服務業者，發布個資參考指引及安維計畫範本。

（二）2024 年 1 月推出《隱私強化技術指引》，推廣以技術方式降低直接利用原始資料所衍生之風險，同時保有資料可用性。

（三）2023 年 12 月 6 日臺北科技大學舉辦「AI 產品與系統評測中心」啟動活動，推動我國 AI 評測制度與可信任 AI 環境發展，降低 AI 科技發展帶來之潛在負面影響。並於 2024 年 3 月提出《人工智慧 (AI) 產品與系統評測參考指引（草案）》，作為推動我國 AI 評測環境發展與保障運用 AI 產品與系統安全性之依據，讓各產業開發或使用 AI 產品與系統時有一可遵守依循文件，實質提升 AI 應用安全及可信任程度。

伍、結論

AI 為人們提供便利與效率，堪稱歷史上最令人驚嘆的科技革命之一，且前景可期，企業應制定 AI 策略，學生應普遍學習 AI，資訊從業人員尤須致力研發 AI，立法委員也應思考如何立法，讓我國在建構 AI 國際準則上，扮演正面、領先的角色。

AI 產品之設計、發展、使用與評估，必須在實踐 AI 效益同時，考量潛在風險，並確保整體社會共享這些效益。更且 AI 應用到戰爭、金融、交通、醫療等各產業時，亦需要符合各該項領域法規之制約，已有證據顯示，愈來愈多部門將 AI 用於與國家安全有關領域，有必要審視 AI 對和平與安全的影響，AI 自主程度也被認為是對基礎設施的潛在威脅，一旦交由 AI 決定，不排除可能產生人類無法接受的錯誤；另就目前被用於戰爭的自主性武器，更應進一步評估影響及管制，以確保 AI 符合人類要求與

預期，保留人類主動權。

AI 真正價值在創造，不在破壞；負責任的 AI 科技發展，需要法律、規範與政策協助配合。因此，保持對 AI 治理的動態性，建構安全的 AI 產業環境，盡量以不阻礙創新之方式對 AI 進行監理，可透過制度面與技術面框架來統整，依 AI 實際功能類型來分類風險，據以進行區分管制與風險監測機制，並同時強調負責任治理 AI 的重要性，才是永久之謀。（投稿：2024 年 1 月 21 日；修訂：2024 年 4 月 28 日；接受：2024 年 5 月 20 日）

參考文獻

一、中文部分

（一）專書

卡爾．伯格斯特姆、杰文．威斯特著，穆思婕、沈聿德譯，2022 年 12 月。《數據的假象》。臺北市：天下雜誌。

托比．沃爾許著，戴至中譯，2019 年 10 月。《2062 人工智慧創造的世界》。臺北市：經濟新潮社。

艾美．韋伯著，黃庭敏譯，2020 年 3 月。《AI 未來賽局》。新北市：八旗文化。

李開復，2019 年 5 月。《AI 新世界》。臺北市：遠見天下文化，2 版。

洪錦魁，2023 年 4 月。《AI 和 ChatGPT 人類和機器共生的未來》。臺北市：深智數位。

張麗卿主編，2021 年 6 月。《人工智慧與法律挑戰》。臺北市：元照出版公司。

陳銘雄、楊哲銘、李崇僖，2019 年 9 月。《人工智慧與相關法律議題》。臺北市：元照出版公司。

臺灣行政法學會主編，2019 年 9 月。《跨越傳統與現代科技之行政管制與執行法制》。臺北市：元照出版公司。

劉靜怡主編，2019 年 3 月。《人工智慧相關法律議題芻論》。臺北市：元照出版公司。

蕭宏宜，2022 年 1 月。《科技發展與刑事立法》。臺北市：一品文化出版社。

（二）期刊論文

王道維、邱筱涵，2023 年 6 月。〈當 ChatGPT 來敲法官的門—淺談 AI 應用於司法審判的原則與方式〉，《當代法律》，第 18 期，頁 54-55。

朱蓓蕾，2023 年 1 月。〈中國大陸應急管理體系之改革：風險、治理與挑戰〉，《安全與情報研究》，第 6 卷第 1 期，頁 1-47。

余和謙，2019 年 8 月。〈人工智慧之治理—以深度偽造為例〉，《科技法律透析》，第 31 卷第 8 期，頁 52-72。

李岳軒、林志潔，2023 年 3 月。〈臺灣假訊息管制的未來展望—以規範與科技的互動為核心〉，《高大法學論叢》，第 18 卷第 2 期，頁 57-108。

李建良，2022 年 2 月。〈元宇宙與法秩序〉，《臺灣法律人》，第 8 期，頁 38-40。

李建良，2022 年 8 月。〈智慧駕駛的德法學思辨〉，《月旦法學》，第 327 期，頁 94-114。

李崇僖，2021 年 7 月。〈人工智慧時代之資料治理與法制初探〉，《臺灣科技法學叢刊》，第 2 期，頁 117-186。

李崇僖，2022 年 2 月。〈從產業革命反思人工智慧專利議題〉，《萬國法律》，第 241 期，頁 3-10。

李榮耕，2022 年 3 月。〈刑事程序中人工智慧於風險評估上的應用〉，《政大法學評論》，第 168 期，頁 117-186。

周晨蕙，2023 年 10 月。〈生成式 AI 監管方向初探—美、歐、日之路徑分析〉，《科技法律透析》，第 35 卷第 10 期，頁

63-70。

林昕璇，2020 年 6 月。〈論大規模政府監控之資訊隱私保障－評析美國聯邦法院相關判決〉，《臺灣民主季刊》，第 17 卷第 2 期，頁 48-59。

林昱梅，2022 年 6 月。〈預防原則於人工智慧之適用—歐盟人工智慧規則草案及其基本原理之觀察〉，《臺灣法律人》，第 12 期，頁 1-17。

林昱梅，2023 年 7 月。〈歐盟資料治理法制新趨勢，以公部門及資料利他組織之資料再利用為中心〉，《臺灣法律人》，第 25 期，頁 75-96。

林發立，2023 年 6 月。〈內容提供產業及創作人如何看待 AI －從巨觀到微觀〉，《當代法律》，第 18 期，頁 60-67。

范姜真媺，2023 年 10 月。〈生成 AI ChatGPT 之運用與個人資料保護〉，《月旦法學》，第 341 期，頁 26-35。

張文貞，2007 年 11 月。〈跨國憲政主義的合縱與連橫—歐洲人權法院及內國憲法法院關係初探〉，《月旦法學》，第 151 期，頁 57-70。

張文貞，2019 年 11 月。〈2018 年憲法發展回顧〉，《臺大法學論叢》，第 48 卷特刊，頁 1503-1545。

張腕純，2023 年 10 月。〈歐洲理事會公布之「AI、人權、民主與法治綱要公約」草案〉，《科技法律透析》，第 35 卷第 10 期，頁 17-19。

許力儒、莊弘鈺，2022 年 2 月。〈人工智慧創作之著作權適格與歸屬—法律與技術之綜合觀點〉，《萬國法律》，第 241 期，頁 28-30。

許育典、李霽恆，2022 年 4 月。〈網路平臺上假訊息的管制問題〉，《國立中正大學法學集刊》，第 75 期，頁 167-227。

許恒達，2019 年 2 月。〈人工智慧與司法與談意見（三）〉，《檢察新論》，第 25 期，頁 42-47。

許嘉芳，2023 年 5 月。〈淺談 NIST「人工智慧風險管理框架」〉，《科技法律透析》，第 35 卷第 5 期，頁 10-18。

陳春山， 2022 年 12 月。〈元宇宙產業及智財法制前瞻〉，《萬國法律》，第 246 期，頁 74-78。

陳家駿，2022 年 3 月。〈「元宇宙」科技之法律議題初探〉，《月旦法學》，第 322 期，頁 209-223。

陳起行，2022 年 6 月。〈論人工智慧時代演算法爭議諮詢及調解〉，《月旦法學》，第 325 期，頁 109-111。

陳起行，2023 年 2 月。〈初探大數據與人工智慧對未來立法影響〉，《月旦法學》，第 333 期，頁 81-82。

陳陽升，2023 年 8 月。〈從法治原則探索人工智慧之應用界限〉，《臺灣法律人》，第 26 期，頁 97-106。

陳譽文，2017 年 4 月。〈人工智慧規範性議題綜觀〉，《科技法律透析》，第 29 卷第 4 期，頁 43-51。

彭睿仁，2022 年 7 月。〈從法制層面論德國因應疾病大流行之防治措施—兼論科技防疫工具運用之比例原則檢驗〉，《臺灣科技法學叢刊》，第 4 期，頁 43-90。

程法彰，2022 年 12 月。〈區塊鍊技術運用與歐盟個人資料保護的規範政策折衝〉，《萬國法律》，第 246 期，頁 79-85。

黃崇祐，2023 年 5 月。〈穿透認知作戰指揮調度〉，《清流雙月

刊》，第 45 期，頁 10-15。
楊智傑，2019 年 7 月。〈歐盟與德法網路平臺假訊息責任立法比較與表意自由之保護〉，《憲政時代》，第 45 卷第 1 期，頁 43-106。
楊智傑，2023 年 1 月。〈網路選舉宣傳揭露資助者、外國勢力與言論自由〉，《憲政時代》，第 46 卷第 4 期，頁 495-557。
楊智傑、鄭富源，2024 年 3 月。〈歐盟人工智慧法與生成式 AI 規範〉，《國會季刊》，第 52 卷第 1 期，頁 1-30。
溫祖德，2020 年 4 月。〈法人犯罪量刑與法令遵循〉，《刑事政策與犯罪防治研究》，第 24 期，頁 115-154。
趙萃文，2023 年 11 月。〈在 AI 取代人力的那一天—談 AI 技術革新下的法律挑戰〉，《清流》，第 48 期，頁 33-38。
劉汗曦，2022 年 12 月。〈從數位憲政與數位信任看我國健保資料庫的爭議與使用〉，《月旦法學》，第 331 期，頁 49－51。
劉奕成、葉柏廷，2023 年 6 月。〈ChatGPT 問世五年內金融服務業的 AI 戰略〉，《當代法律》，第 18 期，頁 88-99。
蔡宜臻，2023 年 10 月。〈人工智慧侵權責任新局？從歐盟人工智慧責任指令草案立法方向觀察〉，《科技法律透析》，第 35 卷第 10 期，頁 47-53。
鄭旭高，2023 年 11 月。〈人類與 AI ：共榮或毀滅？〉，《清流雙月刊》，第 48 期，頁 22-26。
簡立峰，2023 年 10 月 18 日。〈會創作也會瞎掰 AI 背後是機率〉，《天下雜誌》，第 784 期， AI 與人才專欄。

魏伶娟，2022 年 7 月。〈人工智慧浪潮對民事責任建構的挑戰－以智慧醫療器材之應用為例〉，《中正財經法學》，第 25 期，頁 1-60。

顧振豪，2023 年 6 月。〈生成式人工智慧與法律的和諧共舞〉，《當代法律》，第 18 期，頁 47-52。

（三）其他媒體

〈人工智慧的美麗與憂愁〉，《工商時報》，2023 年 7 月 5 日，社論。

〈國家人權委員會新聞稿〉，112 年 3 月 29 日，最後瀏覽日：2023/12/06。

臺灣人工智慧行動網，〈歐盟人工智慧規則草案之初探—以市場、風險、價值及信賴為核心的管制架構〉，2022 年 3 月 4 日。

吳武典，〈臺灣土壤能孕育 AI 教父？〉，《聯合報》，2023 年 6 月 10 日，民意論壇。

李念祖，〈ChatGPT 提供的資訊是受憲法保障的言論嗎？〉，《中國時報》，2023 年 3 月 23 日，時論廣場。

林臺森，〈　化識讀能力，嚴防中共深偽入侵〉，《青年日報》，113 年 4 月 16 日。

茅毅編譯，〈賀錦麗找科技巨擘研商 AI 風險〉，《聯合報》，2023 年 5 月 6 日。

孫宇青報導，〈美應把握 AI 技術優勢抗中數位專制〉，《自由時報》，2023 年 2 月 22 日。

徐子苓，〈資通法把關且非軟體我公部門暫不禁 ChatGPT〉，《自

由時報》，2023 年 4 月 7 日。

徐作聖，〈AI 內閣蹭風潮產業怎信賴〉，《聯合報》，2024 年 4 月 17 日。

高詣軒、陳曉慈報導，〈AI 峰會宣言共同對抗潛在風險〉，《聯合報》，2023 年 11 月 2 日。

張瑞雄，〈當企業比政府更有權〉，《中國時報》，2023 年 9 月 23。

游梓翔，〈AI 衝擊傳播業，你是 80 還是 20 ？〉，《聯合報》，2023 年 6 月 26 日。

楊堯茹，〈孫曉雅：假訊息介選臺美都深受其害〉，《自由時報》，2023 年 10 月 18 日。

漢坤律師事務所，〈歐盟 AI 法案立法觀察〉，2023 年 5 月 14 日。

趙萃文，〈GPT 假新聞 資訊戰〉，《自由時報》，2023 年 4 月 2 日。

趙萃文，〈我國 AI 基本法立法宜緩不宜急〉，《自由時報》，2023 年 8 月 3 日。

盧天麟、呂宜諠，〈發展自主衛星產業，強化數位韌性與通訊安全（上、下）〉，《工商時報》，2023 年 2 月 21、22 日。

鍾麗華，〈社、媒遭中國滲透臺灣排名世界第一〉，《自由時報》，2022 年 3 月 25 日。

蘇紫雲，〈中共操弄認知作戰，「垃圾偽軍」為禍全球〉，《青年日報》，113 年 4 月 15 日。

二、英文部分

（一）專書

Macnish, Kevin, 2020. *Big Data and Democracy*. Edinburgh, UK: Edinburgh University Press.

Moran, Michael, 2003. *The British Regulatory State: High Modernism and Hyper-Innovation*. New York: Oxford University Press.

（二）網際網路

European Commission, 2019/04/08. "Ethics Guidelines for Trustworthy AI," <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477>.

European Council, 2023/12/09. "Artificial Intelligence Act: Council and Parliament strike a deal on the first rules for AI in the world," <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificialintelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>.

iThome, 2023/06. "Open letter to the representatives of the European Commission, the European Council and the European Parliament," <https://www.igizmo.it/wp-content/uploads/2023/06/Open-Letter-EU-AI-Act-and-Signatories.pdf>.

Reuters, 2024/01/12. "EU wants AI Act to be global benchmark, but Asian countries are not convinced," <https://www.scmp.com/tech/tech-trends/article/3228050/eu-wants-ai-act-be-global-benchmark-asiancountries-are-not-convinced?module=perpetual_scroll_0&pgtype=article&campaign=3228050>.

Stewart, Phil, 2017/11/27. "China Racing for AI Military Edge over US: Report," <https://www.reuters.com/article/idUSKBN1DS0GN/>.

The Economist, 2017/05/06. “The world’s most valuable resource is no longer oil, but data,” <https://myppt.cc/Z4piA0/>.

Toh, Michelle, 2024/01/12. “Serious concerns: Top companies raise alarm over Europe’s proposed AI law,” <https://edition.cnn.com/2023/06/30/tech/eu-companies-risks-ai-law-intl-hnk/index.html>.