MODA-DODS-113-001（委託研析報告）

公部門人工智慧規範架構
（成果報告）

數位發展部編印

中華民國 114 年 07 月

MODA-DODS-113-001（委託研析報告）

公部門人工智慧規範架構（成果報告）

受委託單位：數位治理研究中心

計畫主持人：曾憲立

協同主持人：朱斌妤、戴豪君、許慧瑩

研究顧問：蕭乃沂

研究助理：劉子筠、楊椀如、陳薇、李柏均、林思妤

數位發展部編印

中華民國 114 年 07 月

目次

圖次

表次

摘要

隨著人工智慧技術的迅速發展，世界各國正積極制定相關法規或行政命令，以應對人工智慧在資安防護、民眾隱私、數位權利、公部門數位轉型及數位涵容社會等方面帶來的挑戰。

本研析回顧並蒐集國際上主要的人工智慧治理參考文獻，包括歐盟、美國、英國、澳洲等國家，以及聯合國（UN）、麻省理工學院（MIT）、經濟合作暨發展組織（OECD）和七大工業國組織（G7）等國際組織或研究單位的相關報告和倡議。並透過次級資料分析和專家訪談，盤點我國政府人工智慧應用類型與場景，包含「數位國家。創新經濟發展方案（DIGI+) 2023」（40 案）、「政府服務獎（第 1-6 屆）」（32 案）、「臺北市政府」（28 案）；專家訪談對象則涵蓋資安專家、內政部、環境部、臺北市政府和會計稽核人員。據以建立五個階層的「人工智慧治理架構與措施建議」，包含法律、標準、政策與計畫、公務同仁行為、人工智慧價值。以及依據計畫、執行、查核與行動（PDCA）循環管理理念，建構一套四大構面、可持續優化的人工智慧推動模式，作為我國公部門因應人工智慧發展挑戰與創新需求的制度性參考。

隨著人工智慧技術的發展迅速，目前的研析結果仍可能受到跨領域複雜性、資料蒐集挑戰、民眾信任、缺乏全球共識等因素的限制，本研析建議相關單位應持續關注國際人工智慧發展與參與國際合作。

關鍵字：人工智慧、資料治理、風險治理

第一章　緒論

第一節 研析背景

伴隨人工智慧（Artificial Intelligence, 以下簡稱 AI）於產業與政府各領域的普及，世界各國均高度重視 AI 技術與相關應用，尤其 ChatGPT 於 2022 年底發布並引起非技術使用者的廣泛使用。隨著 AI 技術的快速進步，世界各國均加速制定相關法規或行政命令，以因應人工智慧帶來資安防護、民眾隱私保護、數位權利、公部門數位轉型，與「不遺落任何人」的數位包容社會。我國國家科學及技術委員會首先公布《行政院及所屬機關（構）使用生成式 AI 參考指引》、金融監督管理委員會提出《金融業運用 AI 指引草案」、國家資通安全研究院《AI 產品與系統評測指引》，與刻正研擬之《人工智慧基本法》（以下簡稱《AI 基本法》）草案。

《歐盟人工智慧法》（EU Artificial Intelligence Act，以下簡稱《歐盟 AIA》）於 2024 年 3 月 16 日歐洲議會全體會議正式通過，進行法律專家文本修正，歐盟理事會於 2024 年 5 月 21 日進行最終批准後，完成立法程序，採風險管理導向，以歐盟單一市場整合為前提，促進各國人工智慧發展與創新並保障數位權利與原則的落實。

AI 風險分成四種等級：不可接受的風險 AI、高風險 AI、有限風險 AI、最低風險 AI。不可接受的風險 AI 包括用於社會信用評分/評比於工作場所或學校使用情緒辨識系統，或影響人類行為或利用使用者漏洞的系統，在歐盟全面禁止。美國聯邦政府各部會已全面建立 AI 指引，並透過 AI.gov 網站，彙整人工智慧應用情形，在人員部分，對公部門中高階主管頒布相關指引《AI Guide for Government》，並要求 AI 生成的文字、影像、語音和影片加上警示，如「浮水印」等，讓使用者更快認出深偽的影像或語音。

第二節　研析目的

依照 OECD 2022 年所發佈的《政府 AI 整備度》調查（Government AI Readiness Index），台灣以 72.78 分排名第 14 名，仍有相當進步空間。考量政府機關運用人工智慧科技協助執行業務或提供服務已成為國際趨勢，為有助於行政效率之提升，且為保持執行公務之機密性及專業性，並促使各機關使用 AI 有一致之認知及基本原則，爰參考各國政府之審慎因應作法，本研析規劃內容包括：

一、 人工智慧規範架構及國際相關規範之蒐集

人工智慧規範架構（AI Governance Framework）是一套指導原則和標準，用於規範人工智慧技術的開發、部署和使用，確保其符合倫理、安全、透明和責任等多方面的要求。人工智慧規範架構的目的在平衡技術創新和社會價值，減少 AI 技術對個人隱私、社會公平和安全的潛在風險，所蒐集之國家或地區包含：歐盟、美國、澳洲、英國；與國際組織如聯合國、隱私遠見論壇（Future of Privacy Forum, FPF）、聯合國大學（United Nations University, UNU）等研究單位與學者之意見 1。

規範架構與使用指引之差別，參見表 1，規範架構之層次應更宏觀與具戰略指導意義。

表 1　人工智慧規範架構與指引之比較

比較項目	AI 規範架構 AI Governance Framework	AI 使用指引 AI Usage Guidelines
層級和廣度	是宏觀層面的，涵蓋 AI 技術的整體治理和策略指導。	是微觀層面的，針對特定操作和應用場景提供具體的實踐建議。
目標和應用對象	為政策制定者提供高層次的數位治理指導。	為具體的 AI 技術用戶和操作人員提供具體的操作指南。
內容和詳細程度	內容廣泛，涵蓋倫理、透明性、課責、安全等多方面。	內容具體，側重於實際操作步驟和具體的合規要求。

資料來源：本研析自行整理

[1] 本研析對各國 AI 政策之蒐集截止於 2025 年 4 月底，特此聲明。

二、 我國人工智慧規範架構之研擬

參考前述規範架構及國際相關規範之蒐集，研擬我國適用的 AI 規範架構，並依照我國國情篩選適用部分，做為我國政府運用 AI 時的高層次、整體性指導架構，以供各行政院所屬單位參考，架構依照國際制定之內容與我國國情之實際情況，本研析建議架構排除以下適用領域：軍事、國防、武器、研發、金融、醫療、媒體和平台內容監管。

第二章 文獻回顧

第一節 世界各國人工智慧發展狀況

一、 歐盟

歐洲理事會（CoE）2023 年 12 月 18 日公布關於《人工智慧、人權、民主和法治架構公約草案》（Draft Framework Convention on AI, Human Rights, Democracy, and Rule of Law），簡稱歐洲理事會《人工智慧架構公約草案》，並於 2024 年 3 月的會議最終確定《人工智慧、人權、民主和法治架構公約草案》的文本[2]。

儘管《歐盟 AIA》已為歐盟內人工智慧系統的規範制定了明確規則，新的《人工智慧、人權、民主和法治架構公約》（Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, AI Convention，以下簡稱《AI 架構公約》創建可適用於美國、英國、歐盟及其他國家的共同架構。除了歐盟外，在其會員國和歐洲理事會會員國的支持下，該文本由非歐洲國家參與，即阿根廷、澳洲、加拿大、哥斯達黎加、羅馬教廷、以色列、日本、墨西哥、秘魯、美國和烏拉圭，還有另外 68 名民間社會和產業代表以及歐洲安全與合作組織 （Organization pour la sécurité et la coopération en Europe, OSCE）、經濟合作暨發展組織（Organization for Economic Cooperation and Development, OECD）等其他國際組織的代表參加談判[3]。

最終，歐洲理事會於 2024 年 5 月 17 日正式通過的《AI 架構公約》，是國際第一個專為 AI 治理所建置之具法拘束力的國際條約。《AI 架構公

---

[2] ENNHRI, Draft Convention on AI, Human Rights, Democracy and Rule of Law finalised: ENNHRI raises concerns (2024), https://ennhri.org/news-and-blog/draft-convention-on-ai-human-rights-democracy-and-rule-of-law-finalised-ennhri-raises-concerns/#:~:text=On%2014%20March%202024%2C%20the,on%20Artificial%20Intelligence%20(CAI) (last visted July 12, 2025).

[3] The Council of Europe, The Framework Convention on Artificial Intelligence (2024), https://rm.coe.int/1680afae3c (last visted July 12, 2025).

約》已於 2024 年 9 月 5 日開放簽署，包括美國和英國安道爾、喬治亞、冰島、挪威、摩爾多瓦共和國、聖馬力諾和以色列等國均已簽署該公約[4]。

該《AI 架構公約》主要補充現有關於 AI 在人權、民主和法治的國際標準，填補因快速技術進步可能產生的法律空白。為因應科技日新月異，《AI 架構公約》並不直接規範技術，而是本質上保持技術中立。AI 架構公約的目的是規範確保公部門使用 AI 系統，須在 AI 生命週期充分遵守人權、尊重民主運作並遵守法治，以維護歐洲基本價價值[5]。

《AI 架構公約》內容主要可分為下列：在第一章總則，明列 AI 架構公約目的，旨在確保 AI 系統生命週期中的各項活動完全符合人權、民主與法治的要求（第 1 條），並進一步規範公約內用語的定義（第 2 條）及公約的適用範圍（第 3 條）。接續，於第二章規定了締約國人權、民主程序和尊重法治方面的普遍義務。第三章（第 6 至 13 條）確立實施相關活動的基本原則，例如平等與非歧視、人類尊嚴、透明性、課責性及安全創新的原則。第四章（第 14 和 15 條）主要於救濟與程序性保障，第五章（第 16 條）則聚焦於風險與影響管理。第六章（第 17 至 22 條）涉及 AI 架構公約的實施，包含例如第 18 條中關於兒童與身心障礙人士權利的條款，以及第 20 條關於數位素養與數位技能的規定。最後，第七章（第 23 至 26 條）確立了 AI 架構公約運作機制與合作架構，並在第 26 條中引入強制性監測機制[6]。有關 AI 架構公約的基本原則、受拘束的客體，以及《AI 架構公約》的執行與監督簡述如下：

（一） 架構的制定

早於 2019 年就啟動使用架構方式規範 AI 的想法，由當時 AI 特設委員會（ad hoc Committee on Artificial Intelligence, CAHAI）受命研析制定架構式公約的可行性。CAHAI 完成任務後，該委員會於 2022 年被 AI 委員會（Committee on Artificial Intelligence, CAI）接替，負責起草並協商公約文本。

《AI 架構公約》由歐洲理事會 46 個成員國共同起草，並包括所有觀察國的參與：加拿大、日本、墨西哥、教廷和美國。此

[4] 同前註。

[5] 同前註。

[6] 同前註。

外，歐盟以及若干非成員國（如澳州、阿根廷、哥斯大黎加、以色列、秘魯和烏拉圭）也積極參與其中，同時為遵循歐洲理事會倡導的多方利害關係人參與原則，亦邀請民間、社會、學術界和產業等 68 位國際代表，及多個其他國際組織，積極參與了 AI 架構公約的制定過程。

（二）《AI 架構公約》的要求

在AI系統生命週期中的各項活動，須遵守以下基本原則，包括：人類尊嚴與個人自主性、平等與非歧視、尊重隱私與個人資料保護、透明性與監督、行政責任與國家責任、可靠性與安全創新。

《AI架構公約》另外規範，締約國採用時AI，應建置下列救濟、程序性權利與保障相關制度或措施：

1. 應記錄有關AI系統及其使用的相關資訊，並向受影響者提供資訊。
2. 所提供的資訊必須足以讓相關人員能夠對AI系統所作出的決定（或主要依據該系統作出的決定）提出異議，並對系統本身的使用提出挑戰。
3. 建置受影響者得向主管機關提出申訴的有效途徑。
4. 在AI系統對人權和基本自由的享有產生重大影響的情況下，為受影響者提供有效的程序性保障、保護措施和權利。
5. 明確告知個人其正在與AI系統而非人類互動的事實。

（三）受拘束的客體

在面對AI可能帶來的風險與影響，《AI架構公約》要求締約國應首先對AI系統對人權、民主和法治可能造成的實際及潛在的影響進行風險與影響評估，並以循環更迭替代的方式不斷完善系統；其次，根據評估結果，締約國應制定並實施整備的預防和緩解措施；後續，有權機關應對某些AI系統的特定應用，設定禁令或暫時停止使用，即「紅線」（red line）措施。

《AI架構公約》適用客體涵蓋各締約國之公部門（包括受委託行使公權力之私部門），及使用AI系統的私部門。《AI架構公約》為締約國提供兩種遵守原則與義務的方式，特別是私部門。各締約國可選擇要求私部門直接受《AI架構公約》相關條款的約

束；或者採取其他使私部門履行AI架構公約規定，同時完全尊重在人權、民主與法治方面的國際義務的替代方式。

《AI架構公約》締約國不必將公約條款適用於涉及國家安全利益保護的活動，但必須確保此類活動遵守國際法以及民主機構與程序的要求。該架構公約不適用於國防事務，也不適用於研析與開發活動，但當AI系統的測試可能對人權、民主或法治產生干擾時，則需適用相關規定。

（四） 監測《AI 架構公約》實施的機制

《AI架構公約》設立締約國會議（conference of the parties），負責締約後的《AI架構公約》實施程度的監測。締約國會議由締約國的官方代表組成之監測機制，監測各國公約實施狀況與提出實施建議。此機制有助於確保締約國遵守《AI架構公約》，並保障公約的效力。同時，締約國會議還將促進與相關利害關係人的合作，包括通過公開聽證會就《AI架構公約》實施的相關問題進行討論。

與其他國際公法性質相同，《AI 架構公約》為國際公約之一環，公約並不直接賦予自然人或私人組織權利或施加義務，而是需透過各締約國依其規定，完成公約國內法化的程序，制定相應的法律和程序，確保公約在國內得以有效實施。

《AI 架構公約》的適用範圍涵蓋 AI 系統生命週期內可能干擾人權、民主和法治的活動，更強調對於民主進程和法治的保護。雖然《歐盟 AIA》也致力於解決前述問題，但其更側重於市場監管和 AI 安全融入內部市場。締約國應將《AI 架構公約》適用於公部門及受委託行使公權力之私部門，在 AI 系統生命週期內所有使用 AI 系統的活動。《AI 架構公約》對於各該締約國國內之私部門並無直接的拘束能力，而須透過各該締約國以遵循《AI 架構公約》的目標和宗旨為前提，透過相關法制或程序的訂立，規範國內之私部門因使用 AI 系統所進行的活動可能產生的風險和影響。

《AI 架構公約》為世界上第一個致力於 AI 治理的國際條約，象徵全球 AI 領域邁出里程碑的一步，兩者皆以保障人權、維護民主制度及強化法治為核心。儘管兩者定位不同，一為國際公法性質的條約，一為區域法規，但實際上可視為互補關係，共同構築歐洲及國際社會面對 AI 風險時的規範基礎。《AI 架構公約》提供締約國一般性義務和原則，將具體細節

留給締約國透過國內立法的方式，提供更具體的領域做進一步的補充。《AI 架構公約》與《歐盟 AIA》基於相同的基本價值和原則，兩者相輔相成。

《AI 架構公約》主要功能是對成為公約締約國和其他國際組織產生約束力，將公約規則和原則納入其國內法律和監管架構。另外，於《歐盟 AIA》與《AI 架構公約》的推動與落實，《歐盟 AIA》將直接適用於歐盟會員國，由各國負責落實，包括具體的執行機制，例如違規行為的罰款，並將實施責任分配給會員國；《AI 架構公約》則透過設立締約國會議來監督實施情況，並促進締約國的合作。

首先，在 AI 系統的定義上，《AI 架構公約》第 2 條與經濟合作暨發展組織（OECD）及《歐盟 AIA》所採用的定義保持一致。這種跨組織的定義協調不僅提升法律與政策對話的一致性，也有助於未來在跨國監管合作中降低法律解釋分歧的風險，對於全球 AI 治理具有重大意義。

其次，在監管邏輯方面，雖然《AI 架構公約》與《歐盟 AIA》均採取風險理論的監管邏輯，但兩者在具體實踐上存在差異。《歐盟 AIA》明確劃分高風險、低風險與禁止類型系統，並設有風險分類例示；《AI 架構公約》並未設立此類例示清單。相反地，它採取較為寬泛與原則性的方法，適用於所有可能對人權、民主與法治產生影響的 AI 系統，並要求締約國依據一套高階義務與風險評估架構加以規範。這樣的作法雖較具彈性，但也更依賴國內法與執行機關的實質落實能力。

在原則層面，兩份文件皆將透明度與課責制視為 AI 設計與部署的核心。文件要求 AI 系統應具可解釋性，且使用 AI 的機關單位對 AI 使用後的實際影響負責。此外，《AI 架構公約》與《歐盟 AIA》均強調應防止 AI 造成歧視性後果，特別是針對婦女、少數群體及身心障礙者等弱勢族群。此一立法趨勢顯示，AI 治理不僅關乎科技風險控管，更需處理與平等、正義和包容相關的倫理議題。

兩者亦共同認識到國際合作在 AI 治理中的重要性，並致力於推動司法管轄區之間標準的協調。《AI 架構公約》作為國際條約，旨在建構一套適用於多國締約方的共同原則，而《歐盟 AIA》則藉由歐盟單一市場架構，實現區域內部的統一規範。兩者均回應了 AI 技術所帶來的跨境風險，並展現出建立全球治理架構的企圖。

最後，《AI 架構公約》的法律性質與執行方式也與《歐盟 AIA》有所區別。《AI 架構公約》為國際條約，並不直接賦予自然人或私人組織法律上的權利或義務，其主要功能在於對締約國產生拘束力，引導其將 AI 架構公約中的規範原則內嵌於本國法律與監管架構中。相對地，《歐盟 AIA》具備直接法律效力，設有具體的執行與制裁機制，例如違規罰款與成員國責任配置。

為確保規範的落實，《歐盟 AIA》設有明確的監督與執行機制，而《AI 架構公約》則設置締約方會議作為監督與合作平台，旨在持續檢視《AI 架構公約》實施成效並促進國際協調。

二、 美國

隨著 AI 工具的普及，美國聯邦、州和地方政府紛紛採用 AI 系統於公共服務，用以提高效率、改進決策、改善服務的提供，並致力於理解 AI 可能的用途與管理架構，包含建立公共服務可能的用途，與設置未來可能用途的安全框架，並強調治理、倫理，以及跨機關合作的重要性。

美國聯邦政府對於 AI 的治理，隨著政黨輪替經歷顯著的變化。在拜登政府期間，聯邦政府主要以負責任的 AI 為主軸推動 AI，建立保障公民權利、強化政府內部風險管理機制與跨部會協調的治理架構；在川普政府重新上任後，聯邦政府的 AI 治理重點，轉向鬆綁規制與鼓勵創新，強調避免過度干預私部門的 AI 發展，並以強化國安與競爭力為優先考量。

（一） 拜登政府時期

拜登政府任內之 AI 治理以風險管理、跨機關協作、責任倫理為主軸。管理與預算辦公室（Office of Management and Budget，以下簡稱 OMB），以拜登政府第 14110 號《人工智慧的安全、可靠與可信的開發和使用》行政命令（Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence）為框架，於 2024 年 3 月首先發布第 M-24-10 號《有關聯邦機構強化 AI 相關治理與風險管理實踐》備忘錄（Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence）。前述備忘錄的訂定同時呼應了《2020 年政府 AI 法》（AI in Government Act of 2020）及《促進美國 AI 法》（Advancing American AI Act）對 AI 的規制方向。

該備忘錄以系統性規範聯邦機關應用 AI 的治理與風險管理措施，要求各機關設立首席 AI 官（CAIO），強化透明度與課責性，並針對影響個人權利與公共安全的 AI 系統設定最低風險管理標準。CAIO 負責推動聯邦機關內所有有關 AI 採購和使用的任務外，亦建議機關應參考國家標準與技術研究院（NIST）所發布之《人工智慧風險管理框架》（Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile），以促進機關 AI 應用相關之技術符合倫理和安全。

OMB 鼓勵政府機關使用 AI 用於簡化公共服務的流程、降低成本並提高整體效率。在風險緩解方面，OMB 強調機關需識別並評估與 AI 相關的風險，制定應變計劃，並持續監控 AI 系統以應對新興風險。根據該備忘錄，AI 治理應內化爲機關 AI 使用和資訊技術的規劃，以確保在聯邦政府範圍內對 AI 應用採取統一致的作法。此外，機關使用 AI 透過透明公開的方式與公眾溝通 AI 的使用與其可能帶來的影響有其必要性，機關也應確保 AI 系統符合倫理原則，注重公平性、透明度和課責性。OMB 建議機關應建置 AI 治理與管理機制、對風險進行盤點與分類，與參考實作指引採取與實踐因應風險所應最低遵循程度的行政作為，包含以下四個面向[7]：

1. 加強AI治理：由於AI與資料、資訊技術（IT）、安全、隱私、公民權利與自由、客戶體驗以及人員管理等其他技術和政策領域密切相關，各機關須在指定首席人工智慧官（CAIO），負責AI治理，及與相關領域的官員及組織密切協作。
2. 推進負責任的AI創新：在具備適當的保障措施下，AI可成為改善聯邦政府公共服務的有利的工具，然在使用AI之際，各機關須強化負責地採用AI的能力，並採取促進AI模型、程式碼及資料共享和再利用的相關措施。

[7] US Office of Management and Budget, M-24-10 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence (2024), https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf (last visited May 11, 2025).

3. 管理使用AI帶來的風險：當機關採用AI時，亦須認知AI使用可能帶來的風險。各機關需遵守既有與AI相關的風險管理要求，並增加OMB為AI使用所制定的新的要求和建議，以應對使用AI進行決策所產生的特定風險，尤其是對於公眾的權利與安全的影響時。因此，OMB提供機關使用影響安全和權利的AI時須遵守的最低實踐準則，並列舉被認定屬影響權利與安全的特定AI類別。最後，OMB在備忘錄中另針對聯邦機關採購AI的風險管理，提出系列建議。
4. 推動負責任的AI創新：在具適當與整備的保障措施下，AI可成為機關現代化和改善聯邦政府向公眾提供服務的有利工具，於此同時，機關須提高負責任採用AI的能力，並採取促進AI模型、程式碼和資料的共享與重複使用等措施。《首席財務官法》（CFO法）的建置，確認聯邦機關應指定CFO與擬定AI之相關策略，說明機關將如何推動負責任使用AI的作為。至於，建議機關可先行以下議題，例如：與資料技術基礎設施、資料、網路安全、對於勞動力的影響，以及生成式AI可能帶來的特殊挑戰等，思考能如何減少機關負責任使用AI的障礙。

OMB 依表定時間在公布與推行第 M-24-10 號備忘錄 180 天後，接續於 2024 年 9 月進一步發布第 M-24-18 號《可課責的 AI 政府採購》備忘錄（Advancing the Responsible Acquisition of Artificial Intelligence in Government）。OMB 立基於第 M-24-10 號備忘錄所建議與要求之制度與事項，在此備忘錄近一步將 AI 治理與風險管理，往前提至機關採購的時點，從 AI 全生命週期的起始點，將 AI 治理預設納入，以有效緩解 AI 的採用所帶來的風險。

OMB 第 M-24-18 號備忘錄要求各機關更新其採購政策與程序，建立跨職能與跨機關的協作機制，並分享 AI 採購資訊以促進治理一致性。在風險與效能管理方面，備忘錄強調對隱私、安全、資料與互通性的事前評估與控制；同時，為防止採購受特供應商並確保採購的彈性，備忘錄鼓勵機關採用創新採購策略，支持多元、競爭且具韌性的聯邦政府 AI 的市場環境，並持續強化採購團隊的執行能力，實現更有效的採購流程。

第 M-24-18 號備忘錄就推動可課責的 AI 政府採購，提出機關應著眼於下列事項[8]：

1. 確保跨職能及跨機關協作：備忘錄要求各機關制定或更新採購政策、程序和實踐，以反映AI的新職責與治理要求，該備忘錄還要求各機構在各行政部門共享有關獲取AI的資訊。
2. 管理AI風險與功效：備忘錄建議或要求各機關依各該情況調整其採購政策和實踐，將AI系統和服務在開發、訓練及部署過程中的特殊性納入考量，尤其是基於OMB第M-24-10號備忘錄所確立之與採購相關的實踐。

   機關必須實施第M-24-10號備忘錄要求實踐之事務，以確保得以有效部署，個別對應影響權利及影響安全類別的AI，各自所需的風險管理措施，包括針對隱私、安全、資料所有權與權利，以及互操作性等複雜問題採取的具體行動。此外，OMB還規定或建議採取額外的實踐措施，以確保採購負責任地生成式AI和AI支持的生物特徵辨識系統。
3. 透過創新收購促進競爭性AI市場：AI市場的供應商執行多種任務以促進AI的採用，包括資料蒐集和標註、開發模型人員、基礎設施提供、系統整合以及AI服務提供。備忘錄要求各機關在互操作性方面應優先納入考量謹慎決策，並採取措施防止讓供應商有獨佔採購優勢的問題，同時強烈鼓勵機關運用創新實踐，幫助機關在AI採購中獲得最佳結果，並支持多元化、具有競爭力與具有彈性之AI聯邦政府採購市場。

從前述兩份 AI 治理關鍵的備忘錄可見，拜登政府的 AI 治理政策，整體上呈現與歐盟 AIA 採相近的監管路線。美國政府於規劃採用 AI 前，努力辨識與平衡 AI 應用的風險與機會，正視 AI 對現實世界的影響，以負責任使用 AI 為目標，評估機關技術和資料基礎設施，進一步透過立法與監管方式建立 AI 治理框架和隱私標準，以確保 AI 應用的可靠性、維持安全性和保障不受外力入侵。

---

[8] US Office of Management and Budget, M-24-18 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Advancing the Responsible Acquisition of Artificial Intelligence in Government (2024), https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf (last visited May 11, 2025).

（二）川普政府時期

然而，美國聯邦政府 AI 治理在川普就任之後，從拜登政府的風險控管與倫理導向，逐步轉變為川普 2.0 政府強調市場導向與開放創新政策的治理邏輯，AI 政策轉向開放與去監管。首波行動為撤銷拜登任內的第 14110 號行政命令及相關 AI 政策[9]，隨後於 2025 年 1 月 23 日發布第 14179 號《消除美國 AI 創新障礙》行政命令（Removing Barriers to American Leadership in Artificial Intelligence），要求各機關檢視並撤銷限制 AI 發展的不利規範，強調應避免意識形態偏見，並由白宮科技助理與國安顧問共同制定 AI 國家行動計畫，確保美國 AI 領先地位[10]。OMB 於 4 月發布兩份有關 AI 治理的備忘錄，全面取代拜登政府任內所發布的備忘錄[11]。

OMB 第 M-25-21 號《有關透過創新、治理與公眾信任，加速聯邦政府對人工智慧的應用》備忘錄（Accelerating Federal Use of AI through Innovation, Governance, and Public Trust），以三大支柱：創新、治理與公眾信任為優先，要求各機關指定 CAIO，制定 AI 策略並實施最低風險管理原則，減少不必要的行政官僚作法。OMB 建議 CAIO 應制定 AI 策略並整合現有資料與技術資源，最大化美國開發的 AI 產品使用，同時維護公民權利與隱私。其範圍涵蓋所有行政機關，並對「高影響力（high-impact）AI 系統」設定最低風險管理措施。備忘錄鼓勵資源共享、程式碼開源、資料再利用及跨機關協作，並要求 AI 策略對公眾透明、可近用。各機關需對高影響 AI 進行風險評估、影響評估與持續監控，確保部署前有充分準備，並設定上訴與補救機制[12]。

---

[9] US White House, Initial Rescissions of Harmful Executive Orders and Actions (2025), https://www.whitehouse.gov/presidential-actions/2025/01/initial-rescissions-of-harmful-executive-orders-and-actions/ (last visited May 11, 2025).

[10] US White House, Removing Barriers to American Leadership in Artificial Intelligence (2025), https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/ (last visited May 11, 2025).

[11] 川普政府任內 OMB 第 M-25-21 號與 M-25-22 號備忘錄，全面取代拜登政府任內 OMB 所發布之第 M-24-10 與 M-24-18 號備忘錄。

[12] US Office of Management and Budget, M-25-21 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Accelerating Federal Use of AI through Innovation, Governance, and Public Trust (2025), https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf (last visited May 11, 2025).

川普政府 2.0 將高影響 AI 使用案例納入須特別審查與管理，依第 5 條規定，將 AI 輸出列為決策或行動的主要依據時，若該輸出對個人權利或安全產生法律、實質、約束力或其他重大影響，便構成高影響的 AI 使用情境。要求機關在使用高影響 AI 前，必須進行內部風險評估，以確認該使用是否屬於高影響範疇。評估過程應包括對 AI 輸出的性質與潛在風險的審慎考量，並應檢視是否存在人工監督機制，無論有無監督，決策或行動本身的潛在影響皆不得忽視 [13]。

在最低風險管理實踐的操作方面，機關首先須進行部署前測試，制定涵蓋預期結果與風險緩解的計畫。即使機關無法接觸 AI 模型的原始碼（source code）或訓練資料，亦應透過替代測試方式，如查詢輸出或提供評估資料給開發商進行回饋。機關部署 AI 前，必須完成 AI 影響評估，評估需涵蓋 AI 生命週期，並以實際目標變量衡量預期影響 [14]。

AI 影響評估至少應包括下列項目 [15]：

1. 明確規定AI的預期目標和預期效益。
2. 相關資料品質與模型能力的適當性：應提供有關AI設計、開發、訓練、測試和運行中使用的資料的摘要，並說明其是否適合AI的預期目的。這應包括資料蒐集和準備過程的詳細描述，並指出這些資料是否將作為開放政府資料資產公開揭露。當適用時，該摘要應描述資料中有關受聯邦反歧視法律保護的群體的資訊。
3. 使用AI的潛在影響：應該詳細說明使用AI對隱私、公民權利和民事自由的潛在影響，並描述使用AI與不使用AI之間的差異。此評估應參照隱私影響評估、CAIO批准的最小風險管理實踐豁免或其他相關資料。如果存在預期的負面影響（例如不合法的歧視），則應該描述相應的緩解措施。
4. 重新評估計劃和程序：應提供計劃和程序，以便對AI的使用進行定期重新評估。

[13] 同前註。
[14] 同前註。
[15] 同前註。

5. 與AI導入相關成本分析：應進行與AI實施相關的成本分析，特別是與風險管理、隱私保護和其他潛在影響相關的成本。
6. 獨立審查結果：獨立審查的結果應包含在影響評估中，以提供更多的透明度和課責性。
7. 風險接受，需由接受風險的個人簽名確認：在進行AI部署前，機關應確認並記錄風險接受過程。

機關須在部署 AI 後進行持續監控，包括性能測試與人工審查，以辨識潛在的負面影響，尤其是侵犯隱私或公民權利的風險。監控系統的設計，須能偵測部署後系統出現的異常或變化、使用情境改變或資料更新；必要時應啟動風險緩解機制，確保系統與文檔更新，並建立具可追溯性與透明度的評估流程。

機關亦應確保員工接受定期人工訓練與系統評估，培養理解與監管 AI 系統的能力。對於高影響 AI 使用情境，應設有充分的人類監督、干預與課責機制。針對受 AI 決策影響的個人，機關應提供明確、便利的補救與上訴管道，確保有人工審查的機會。若已存在上訴或審查機制，應予以擴展至 AI 決策領域，並減少個人負擔，符合現行法律規範。最後，機關應建立機制，以徵詢並納入最終使用者及公眾的意見與反饋，尤其是在 AI 的設計、開發與實際應用過程中，確保政策制定回應社會期待並強化公共課責[16]。

整體而言，第 M-25-21 號備忘錄在減少行政的繁複程序的同時，強調程序責任、領導機制與基礎建設，以形塑一個以效率與創新為導向、同時具備風險控管框架的 AI 治理體系。第 M-25-22 號備忘錄則立於遵循第 M-25-21 號備忘錄之前提，主張建立彈性與具競爭性的 AI 採購市場，鼓勵使用美國本土 AI 產品，減少使用少數特定供應商，建立資源共享與簡化合規流程，並強調效率與成果導向[17]。

雖兩屆政府在部分架構上（如 CAIO、風險管理、採購原則）具延續性，惟方向迥異。拜登政府偏重以倫理與安全為核心的審慎治理，川普政

[16] 同前註。

[17] US Office of Management and Budget, M-25-22 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Driving Efficient Acquisition of Artificial Intelligence in Government (2025), https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Driving-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf (last visited May 11, 2025).

府則重視創新彈性與市場競爭力。OMB M-25 系列備忘錄雖保留部分風險控管原則，但被批評對高影響 AI 的保護不足、對治理責任定義過於模糊，可能導致實施落差 [18]，如表 2 所示。

表 2　拜登政府與川普政府時期聯邦政府 OMB AI 治理備忘錄比較表

	拜登政府時期（2024）	川普政府時期（2025）
治理理念	偏重風險控管、公民權利保障、跨部會標準化治理	強調創新效率、美國領導、市場競爭
風險分類	影響權利或安全類別	高影響類別
採購策略	強調市場競爭公平、風險評估	偏好美國製造、保護資料
風控倫理	具影響權利及影響安全類別的風險，須進行影響評估、對抗測試、透明度與人權考量	保留但弱化風控規定，由機構自行判斷與執行風險控制

資料來源：本研析自行整理

總體而言，美國 AI 治理正處於兩種價值觀的角力場域：一方主張透過嚴格規範保障權利與公平，另一方則以市場驅動作為推動技術進步與維護國際競爭力的基礎。

三、 澳洲

澳洲認知 AI 並非全新技術但正歷經關鍵性的變革，AI 快速的發展與應用，對於各級政府業務的運作與服務的提供產生重大改變，並可能促進社會、經濟和環境保護的福祉。同時，政府也認知使用 AI 系統可能面臨的法律、隱私、安全及倫理風險，例如偏見與公平性等問題。澳洲政府對 AI 採取多層次的治理架構，結合自律準則、公共部門指引與高風險應用規範，以實現安全、負責任且符合人權的 AI 發展路徑。

自 2019 年已推出自願性採用的《人工智慧倫理原則》（AI Ethics Principles）[19]，該文件由科學與工業研究機構（CSIRO）主導，聚焦於公平性、透明性、可解釋性與可課責性等，為整體政策發展奠定價值基礎。

[18] Madison Alder, Trump White House releases guidance for AI use, acquisition in government (2025), https://fedscoop.com/trump-white-house-ai-use-acquisition-guidance-government/ (last visited May 11, 2025).

[19] Australian Department of Industry, Science and Resources, AI Ethics Principles (2024), https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles (last visited May 11, 2025).

後續，澳洲聯邦、州和領地政府參用澳洲《人工智慧倫理原則》，於2024年6月21日共同發布《確保公部門使用AI的國家架構》（National Framework for the Assurance of Artificial Intelligence in Government，以下簡稱國家架構）[20]。透過《國家架構》建立機關使用AI系統一致性的保障機制，並為機關與其供應廠商提供一致且具體的要求，協助政府以安全與負責任的方式開發、採購和部署與使用AI技術。

《國家架構》係為確保政府使用AI安全與負責任的重要進展，旨在增強公眾對澳州政府在AI使用；重點可分為五大支柱與八大原則，同時允許各司法管轄區根據自身的法律、政策和業務，制定具體的政策和指引。五大支柱主要聚焦於提升社會信任，從而促進政府對AI技術的採用，包括[21]：

（一）治理：

包括組織結構、政策、流程、法規、角色與責任以及風險管理框架，確保AI的安全和負責任使用，並適應未來需求。國家框架強調跨職能專業知識的重要性、領導階層的承諾、積極面對AI風險的意識培養，員工培訓，與得以有效理解和實施AI治理的資源。

（二）資料

要求建立符合相關法律規定之資料治理流程，以確保AI系統使用高品質且可靠的資料，同時強調資料品質與AI模型輸出品質間的關聯。此外，還涉及與資料資產相關的風險管理，包括明確化資料治理流程的角色與責任，及理解與資料相關的法律和行政義務。

（三）風險

以個案方式評估和管理AI可能帶來的風險，對高風險AI應提出更高的要求。風險管理應以AI系統全生命週期中進行，包括經OECD定義的AI系統的四個階段（1. 設計、資料和模型、2. 確認與確效、3. 部署、運作與監控）。此風險管理的實踐，後續澳洲政府將研擬強制要求自我評估的模型，例如新南威爾斯州人工智慧框架。

---

[20] Australian Department of Finance, National Framework for the Assurance of Artificial Intelligence in Government (2024), https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government (last visited May 11, 2025).

[21] 同前註。

（四） 標準

建議在實際可行的情況下，使AI治理實踐與相關國際標準相符，包括目前的AI治理與管理標準，例如：AS ISO/IEC 42001:2023、AS ISO/IEC 23894:2023，與AS ISO/IEC 38507:2022。

（五） 採購

在採購與契約納入AI倫理與保障的要求，包括在供應商契約加入AI相關條款以明確化責任分工，並在風險與機會間尋求平衡。

澳洲政府續於 2024 年 8 月由數位轉型署 （Digital Transformation Agency, DTA）發布《政府使用人工智慧政策》（Policy for the Responsible Use of AI in Government）[22]。該政策於 2024 年 9 月 1 日開始實施，主要目的在增強對澳洲政府的公眾信任，同時成為安全且倫理應用 AI 技術的典範。

該政策以「啟動、參與與進化」（enable, engage and evolve）為框架為主軸，提出原則、強制性要求及建議行動，希冀政策可隨技術與社會期望的變化不斷演進。政策闡明澳洲公部門應如何通過安全且負責任的採用AI；提升透明度、治理和風險保障來增強公眾信任；以及採用前瞻性學習方法以適應技術與政策環境的變化。旨在採用一種前瞻性且適應性的方式，推動政府對 AI 的使用，可隨著時間的推移不斷發展與完善。為協助各級政府實施政策，數位轉型署建立一種在風險管理與創新之間公平平衡的文化，發布負責官員（Accountable Officials, AOs）的機制，與引導機關提升 AI 治理能力的標準，希冀藉此加強機關對 AI 政策變化的應對與適應能力，並參與跨政府機關的協調與合作。

政府機關使用 AI 所引發的挑戰極複雜，且與其他議題密切相關，例如澳洲公共服務行為準則、資料治理、網路安全、隱私保護以及倫理實踐。該政策目的是在於補充並加強（而非重複）現有與政府 AI 使用相關的框架、法律及實踐。政府機關應結合現有框架與法律，確保各機關履行其法定義務 [23]，發布聲明：

[22] Australian Digital Transformation Agency, Policy for Responsible Use of AI in Government (2024) https://architecture.digital.gov.au/responsible-use-of-AI-in-government (last visited May 11, 2025).
[23] 同前註。

（一）AI 透明度聲明的標準，機關應使用中性文字，避免使用技術性文字，向公眾公開 AI 採用的資訊內容。

（二）機關使用或考慮採用 AI 的意圖。

（三）涉及與公眾直接互動，但無人為介入的使用類別。

（四）監控已部署 AI 系統效能的治理、流程或其他措施。

（五）遵守適用法律與法規的情況。

（六）防止 AI 對公眾造成負面影響的措施。

該聲明需每年至少審查並更新一次，或在機關對 AI 應用方式進行重大改變時即刻更新。各機關須以安全、倫理和負責任的方式參與 AI 的使用，但同時也應滿足社區期望和與公眾信任的方式。

《政府使用人工智慧政策》要求機關以安全方式採用 AI，建立明確的責任機制用以提升機關生產力、決策能力、政策成果及政府服務交付效率。各機關在政策生效後 90 天內指派負責官員，並將名單提交至數位轉型署，確保治理架構的落實。為保障公眾權益，機關在部署 AI 時應採取與風險相稱的緩解措施，確保 AI 系統具備透明性與可解釋性，避免對公民造成潛在傷害。此外，政策亦要求機關於生效後六個月內發布公開聲明，說明其 AI 使用情形與應用案例，以強化外部監督與公眾信任。面對不斷演進的技術發展，機關應保持彈性與適應性，透過持續審查與評估 AI 的使用成效，並建立內部反饋機制，以持續優化 AI 治理實務 [24]。

在機關人員的 AI 素養與專業養成面向，AI 政策建議各機關應在政策生效後六個月內，為全體員工提供與政策指導一致的 AI 基礎培訓。此外，根據員工的角色與職責，提供額外的特別培訓，例如針對負責 AI 系統採購、開發、訓練及部署的員工。各機關亦應了解 AI 在機構內的使用範圍與方式，建立一個記錄相關資訊的內部註冊系統，並將 AI 相關考量融入現有框架中，例如隱私保護、安全防護、記錄管理、網路安全與資料治理 [25]。

接續，2024 年澳洲政府發布《高風險 AI 強制性防護措施建議》（Proposed Mandatory Guardrails for High-Risk AI），提出九項防護機制，涵蓋課責、資料治理、合規審查與人類監督等面向，特別是通用型 AI（GPAI）之潛在系統性風險，同年亦發布《自願性 AI 安全標準》（Voluntary AI Safety

[24] 同前註。

[25] 同前註。

Standard），為開發者與企業提供符合風險程度的實務操作建議，促進標準化與產業參與，並預期作為未來立法的技術參考依據。

《高風險 AI 強制性防護措施建議》針對在高風險環境中使用 AI 所可能產生的風險，主張導入一套強制性的防護措施，以確保 AI 的安全與負責任運作。這份建議聚焦於兩大類高風險 AI 定義，一是基於明確用途的 AI 系統，二是基於高度能力但用途難以預測的通用型 AI（GPAI）模型。對於第一類，風險評估著重在 AI 實際使用的情境；對於第二類，則考量其在多種用途中被誤用的潛在風險，澳洲正評估是否應對 GPAI 模型制定專門規範。

前述建議提出六個與高風險 AI 相關的評估原則[26]：

（一） 人權風險

例如 AI 在招聘或犯罪預測中可能產生歧視，傷害婦女或少數族群的權利。

（二） 健康與安全風險

包括AI誤判醫療診斷或對深色皮膚者產生錯誤讀值。

（三） 法律風險

當AI做出有法律影響的決策，但使用者無法選擇退出。

（四） 對群體的風險

如原住民或文化社群受不公平影響。

（五） 系統性風險

當 AI 技術擴大社會偏見、助長詐騙、散播錯誤資訊，甚至危害民主制度與公眾信任。

（六） 評估影響的嚴重性與範圍

包含受影響人數、損害程度、發生機率及緩解措施的有效性。

針對上述高風險情境，澳洲政府建議導入一系列強制性防護機制。包括：要求開發者與使用單位建立課責流程、落實風險管理制度、強化資料治理與安全、在部署前進行測試並持續監控、確保人類可介入 AI 的運作、

[26] Australian Department of Industry, Science and Resources, Mandatory guardrails for AI in high-risk settings: developers and deployers survey (2024), https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers (last visited May 11, 2025).

向用戶揭露 AI 是否影響其決策或互動、建立申訴與救濟機制、確保供應鏈透明度與資訊共享、保存完整記錄供審查、並進行合規性評估以取得認證。

《自願性 AI 安全標準》（Voluntary AI Safety Standard）係為協助組織在現行澳洲法律、逐步成形的監管指導方針以及社會公眾期望之下，安全且負責地部署與使用人工智慧系統，澳洲政府強調企業應充分掌握相關法律規範。十項 AI 防護旨在協助組織在高風險情境下安全、透明且負責地部署人工智慧系統[27]：

（一） 組織應建立、執行並公開 AI 課責流程

涵蓋治理架構、內部能力與合規策略，包括指派AI使用負責人、擬定AI策略與實施相關訓練。

（二） 需導入風險管理機制

根據AI實際使用情境進行風險與影響評估，並持續追蹤其風險緩解效果，評估應從利害關係人影響開始。

（三） 須保護 AI 系統，落實資料治理

包括資料品質、來源追蹤（provenance）及資安防護，考量AI獨特的風險特性。

（四） 所有 AI 模型與系統測試與監控

在部署前皆應進行完整測試，並在部署後持續監控其可能出現的異常或變化與非預期後果，依據風險評估設定測試標準。

（五） AI 系統全生命週期內維持人為介入與控制機制

確保即時干預與防止無意後果，尤其考慮供應鏈中多方組件整合的情形。

（六） 應明確告知使用者有關 AI 介入決策、互動或內容生成情況

以增進使用者與社會大眾的信任，揭露方式可依據應用場景與利害關係人進行調整。

[27] Australian Department of Industry, Science and Resources, Voluntary AI Safety Standard (2024), https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers (last visited May 11, 2025).

（七）　組織須設立救濟機制

讓受到AI影響的人可對AI使用方式、決策或結果提出異議或申訴，保障其參與與救濟權。

（八）　組織應與 AI 供應鏈之共享

與其他單位共享資料、模型與系統相關資訊，使其能有效辨識與處理風險，促進供應鏈透明與協同。

（九）　應保留與維護完整紀錄

包括AI系統清單與技術文檔，以利第三方查核其是否符合防護要求。

（十）　持續與利害關係人互動

評估其需求與處境，特別關注安全、多元、包容與公平，識別與減少偏誤，並排除倫理偏見，確保AI設計與實施具備社會正當性。

對 AI 執行風險評估，並與與相關利害關係人進行磋商。風險評估應反應 AI 使用案例的預期範圍、功能和風險控制措施，惟本標準並未規定機構應評估的具體風險，或用來確定最終風險後果。

在進行風險評估時可考慮的風險範例AI使用可能帶來的風險包括[28]:

（一）對政府服務的公共可及性或包容性產生負面影響。

（二）對個人或社群進行不公平歧視。

（三）加劇刻板印象或貶低對個人或社群的表現。

（四）對個人、社群、企業或環境造成損害。

（五）因系統處理、解析或轉換的資料敏感性所產生的隱私問題。

（六）因系統處理、解析或轉換的資料敏感性或分類所產生安全問題。

（七）由於系統的實施、來源或特徵，導致安全問題。

（八）影響決策過程，並對個人、社群、企業或環境產生影響。

（九）對機構聲譽造成風險，或削弱公眾對政府的信任 。

（十）由於系統處理、轉換或重製第三方擁有著作權的資料，產生智慧

[28] Australian Department of Industry, Science and Resources, Voluntary AI Safety Standard (2024), https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers (last visited May 11, 2025).

財產權問題。

進行風險評估時，機關需要確定每項風險的後果和發生的可能性，並根據具體情況採取相應的風險控制措施。此外，組織必須保留完整記錄，包括 AI 系統清單與技術文件，供第三方進行合規性查核。最後，整個 AI 生命週期中，組織須與利害關係人密切互動，關注多元性、包容性與公平性，辨識偏誤與潛在傷害，並採取措施消除不公平效應，確保 AI 系統符合公共價值與倫理要求 [29]。

四、 英國

英國對 AI 的監管與治理，隨執政黨的政權轉移，從保守黨到工黨亦有所轉變。保守黨政府對 AI 的治理態度，採取較偏向歐盟的風險理論進行監管，先於 2023 年 3 月完成人工智慧監管政策文件（A Pro-innovation Approach to AI Regulation，以下簡稱《人工智慧監管白皮書》）[30] 初稿，後即開放進行意見徵詢，內容主要為針對 AI 可能帶來的風險分爲模型開發階段、擴散階段、部署階段，與對於社會層面的風險，以及於各該階段的風險控管與因應。

《人工智慧監管白皮書》對於 AI 治理的規劃，希望在保障安全與信任的同時，移除創新障礙、推動 AI 產業成長，並維持英國作為全球 AI 領導者的地位 [31]。該框架並未引入新法規，而是建立在現行法律之上，適用於所有在英國境內開發、部署或使用（developing, deploying and using）AI 系統的實體。該框架定義 AI 的核心特徵為「適應性」與「自主性」，並強調其用途導向的監管邏輯。然而，《人工智慧監管白皮書》亦指出，其尚未涵蓋如責任歸屬、資料近用與永續性等議題，認為這些屬於快速變動且需進一步研究的領域。至於，對 AI 的監管不另設新組織，而是強化既有機關（如 ICO、FCA、CMA、Ofcom）的職能，以產業別原則執行監

[29] Australian Department of Industry, Science and Resources, Voluntary AI Safety Standard (2024), https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers (last visited May 11, 2025).
[30] UK Department of Science, Innovation & Technology, Office for Artificial Intelligence, Policy Paper: A Pro-innovation Approach to AI Regulation (2023), https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper (last visited May 11, 2025).
[31] 同前註。

理任務。《人工智慧監管白皮書》指出，英國政府將承諾由中央將提供支援，進行 AI 監管協調、風險預測、推動 AI 沙盒與培育人才 [32]。

惟後續英國因歷經政黨輪替，工黨政府對於 AI 治理規劃與保守黨不同，未完全採納《人工智慧監管白皮書》的規劃，但該政策文件在英國仍為具有指標性意義。執政的工黨政府於 2024 年 2 月上任後發布《人工智慧監管白皮書》意見徵集結果，隨即宣示將 AI 治理轉為輕度監管，重申不會立刻推動具體立法，而是先採行建立在可信任的 AI 五項跨領域原則上的總體監管框架，包括安全性與穩健性、透明性與可解釋性、公平性、課責與治理，以及可挑戰性與救濟 [33]。

工黨對於 AI 治理的走向，於 2024 年上任後國王演說與科技大臣彼得·凱爾（Peter Kyle）的聲明可略知一二，二者皆表示將提出立法以加強對強大 AI 模型的監管 [34]。2025 年初所發布之《AI 機會行動計畫》（AI Opportunity Action Plan），指出英國將投資 AI 基礎設施、推動公共部門應用 AI、設立 AI 專區（'AI Growth Zones' (AIGZs)）促進資料中心的加速建設，並鼓勵監管機構針對其職權領域發布年度報告，強調安全創新導向 [35]。

《人工智慧應用參考手冊》（AI Playbook）進一步為公共部門 AI 使用提供十項實用原則，並納入 AI 採購、風險評估與治理建議 [36]：

---

[32] UK Department for Science, Innovation & Technology, Office for Artificial Intelligence, Policy Paper: A Pro-innovation Approach to AI Regulation (2023), https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper (last visited May 11, 2025) and Simon Bollans, The UK's White Paper on AI regulation: a pro-innovation approach (2023), https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers (last visited July 12, 2025).

[33] Marcus Evans,Rosie Nance, UK Government's Response to AI White Paper Consultation: Next Steps for Implementing the Principles (2024), https://www.dataprotectionreport.com/2024/03/uk-governments-response-to-ai-white-paper-consultation-next-steps-for-implementing-the-principles/ (last visited May 11, 2025).

[34] UK House of Lords, King's Speech, Library Briefings (2024), https://researchbriefings.files.parliament.uk/documents/LLN-2024-0040/LLN-2024-0040.pdf (last visited May 11, 2025).

[35] UK Department for Science, Innovation & Technology, AI Opportunity Action Plan, (2025) https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan (last visited May 11, 2025).

[36] UK Government of Digital Service and UK Department for Science, Innovation & Technology, Artificial Intelligence Playbook for the UK Government (2025), https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government (last visited May 11, 2025).

（一）　了解 AI 及其限制

AI目前仍缺乏推理能力與上下文理解，其效能高度依賴工具本身與應用情境。使用者應熟悉其風險、限制與用途，建立測試流程並採用能提升輸出準確性的方法，以確保安全與負責任的使用。

（二）　合法、合於倫理且負責任的應用

使用AI時必須符合法律與倫理規範，建議及早諮詢法律、合規與資料保護專家，特別是在產品設計初期，應主動評估與減輕如偏見、公平性、智慧財產權等風險，並與受影響者、民間社會、公眾進行溝通與互動。同時，考量AI的環境影響與永續性。

（三）　確保 AI 使用的安全性

依據政府資安政策，當建置與部署AI服務時，必須落實「設計即安全」（security by design）原則，建立安全防護措施並實施技術性控制機制。當AI面臨如資料中毒、提示注入、幻覺等特有威脅，應加強內容過濾與輸出驗證，避免洩露或誤導性資訊。

（四）　建立適當的人為控制機制

AI開發與部署階段須設有人為介入機制，讓使用者通報問題並啟動人工審查，確保人類在重要節點仍保有控制與判斷。

（五）　在全生命週期內管理 AI

應選擇適合的工具並配置足夠資源進行持續監控與維護。預防偏誤、模型漂移與生成式AI幻覺風險。可依《技術實踐守則》（Technology code of practice）與NCSC雲端安全原則（NCSC cloud security principles）作為參考，並妥善處理智慧財產權議題。

（六）　選擇正確的工具

依照任務需求挑選合適的AI模型或產品，避免錯誤使用導致風險。

（七）　促進開放與跨部門合作

利用跨政府社群進行知識交流與資源重用。與其他機關分享程式碼、創新方案與技術基礎設施，有助於提升整體公共部門效率。

（八）　從早期即納入商業團隊

與採購與商業同仁合作，確保AI應用同時符合技術與業務目標。維持內部開發與外部採購AI系統間的一致性與透明度，並透過契約要求供應商遵守資料類別揭露規範。

（九）　具備必要技能與專業知識

建立團隊AI技術與倫理素養。了解開發與部署AI模型所需的專業技能，並區分自訓模型與使用API模型之間的能力要求差異，以支持有效應用。

（十）　配合機關政策並建立適當保障機制

除了遵守本原則外，應依循各機關既有的治理與資安政策。AI專案應於啟動時，即早與機關內的風險與合規稽核部門合作，建立風險監控、審查與升級機制，並設立AI審查委員會，確保AI使用的安全、透明與合法合規。

在立法方面，雖 Lord Chris Holmes 在 2023 年 12 月提出《人工智慧（監管）法案》（Artificial Intelligence (Regulation) Bill）[37]，針對公共部門 AI 決策使用進行規範，但該法案在議會休會進度停滯後未果[38]。接續，2025 年 3 月重新提出的《人工智慧（規範）法案（2025）》（Artificial Intelligence (Regulation) Bill）[39] 延續議員提案（private member's bill）的形式，並提議建立 AI 監管機構，透過法律明文規定監管原則，要求進行 AI 影響評估、提升透明度與強化公共參與。該法案的特殊之處，在於對現行「原則導向且分散監管」模式的挑戰，其監管方式更趨近《歐盟 AIA》的風險分類框架，然而，英國是否立法跟進以《歐盟 AIA》模式規範 AI，目前仍未有定案[40]。

綜上，英國是否最終導入具體法律規制 AI，仍尚待觀察，但從英國代表在巴黎人工智慧行動峰會（Artificial Intelligence Action Summit）[41]的發言可探知，工黨所領導的政府將採取偏向加入美國川普政府的 AI 治

---

[37] Artificial Intelligence (Regulation) Bill [HL] (2023), https://bills.parliament.uk/publications/53068/documents/4030 (last visited May 11, 2025).

[38] Nathalie Moreno, The Artificial Intelligence (Regulation) Bill: Closing the UK's AI Regulation Gap? (2025), https://kennedyslaw.com/en/thought-leadership/article/2025/the-artificial-intelligence-regulation-bill-closing-the-uks-ai-regulation-gap/ (last visited May 11, 2025).

[39] Artificial Intelligence (Regulation) Bill [HL] (2025), https://bills.parliament.uk/bills/3942 (last visited May 11, 2025).

[40] Nathalie Moreno, The Artificial Intelligence (Regulation) Bill: Closing the UK's AI Regulation Gap? (2025), https://kennedyslaw.com/en/thought-leadership/article/2025/the-artificial-intelligence-regulation-bill-closing-the-uks-ai-regulation-gap/ (last visited May 11, 2025).

[41] Dan Milmo, Eleni Courea, US and UK Refuse to Sign Paris Summit Declaration on 'Inclusive' AI (2025), https://www.theguardian.com/technology/2025/feb/11/us-uk-paris-ai-summit-artificial-intelligence-declaration (last visited May 11, 2025).

理陣營，採偏向市場導向之 AI 監管與治理態度；同時，《AI 機會行動計畫》亦已顯示，對不妨礙創新前提下強化 AI 治理的重視，預告 2025 年將見證進一步政策與立法的變革。

五、　國際組織與研究單位之研析報告

(一) 聯合國

聯合國 2024《Governing AI for Humanity》報告指出，AI 的應用跨越國界，其影響涉及地緣政治與經濟等多個層面，因此，AI 治理需要全球合作。目前，AI 演算法運作尚無人能完全掌握，決策者亦未對其開發、部署或使用承擔充分責任，這些決策可能導致的負面影響往往波及全球。該報告提及：雖然 AI 治理已有眾多文件，但現有的架構都無法真正實現全球治理，導致了代表性、協調性及執行力方面的不足。特別是在如何應對快速發展的 AI 技術可能帶來的風險及充分挖掘其潛力上面臨挑戰。因此提出三種合作方式，包含：（1） 雙邊／小多邊（bi-/minilateral）合作，例如美國與歐盟、英國、紐西蘭、新加坡之間的合作；（2）多邊（plurilateral）：涵蓋較大型的多邊組織，如 AI 高峰會（例如 CoE、G7、G20、GPAI、OECD）；（3） 大型多邊（universal）：如聯合國的 AI 治理，所關注的 AI 議題包括：產業標準、AI 風險、與 AI 部署。文中提及全球治理的必要性，歸納出國際人工智慧治理的指導原則和功能（Guiding Principles and Functions for International Governance of AI）：

1. 原則 1：AI 應該以包容的方式治理，為所有人的利益服務。
2. 原則 2：AI 必須在公共利益中進行治理。
3. 原則 3：AI 治理應該與資料治理及促進資料共同體相協調。
4. 原則 4：AI 治理必須是普遍的、網路化的，並根植於靈活的多方利害關係人合作中。
5. 原則 5：AI 治理應基於聯合國憲章、國際人權法和其他已商定的國際承諾，如永續發展目標（SDGs）。

同時，報告觀察到 AI 治理格局正在演變，新興的國際治理模式逐漸成形，圖 1　AI 治理倡議來源與層級呈現了一個系統性框架，以理解目前全球在人工智慧（AI）治理領域中的各種合作

機制，其主要區分標準為「包容性」（Inclusiveness）與「採用者類型」（Adoption）。橫軸代表合作機制的包容性程度，從雙邊／小型多邊（Bi-/minilateral），中型多邊（Plurilateral），到大型多邊（Universal）以及由產業主導的標準與承諾（Industry standards and commitments）；縱軸則區分為由政府主導的採用行為，以及由企業主導的採用行為，並進一步細分為區域性、跨區域與國內層次（由於報告的層次為聯合國，故沒有顯示國內機制的內容與討論）。

在政府主導的合作機制中，跨區域的雙邊或小規模多邊合作呈現出高度戰略性的特徵。例如，美國與英國、紐西蘭、歐盟及新加坡等國之間的合作關係，展現了英語系民主國家在 AI 政策與治理架構上之間的深度聯繫。這類合作通常反映共同價值觀與地緣政治利益，並在資訊共享、風險評估與技術標準制定上進行協調。進一步來看，政府主導下的中等規模多邊合作，如 G7、G20、高峰會（AI Summits）、經濟合作暨發展組織（OECD）、歐洲理事會（CoE）及人工智慧全球合作夥伴關係（GPAI）等，提供了一個跨國政策對話的平台，藉此促進標準的趨同與倫理原則的協商。這些平台雖然不具法律拘束力，但其政策建議與框架性文件對成員國仍具有顯著的政策導引力。在最具包容性的層級上，聯合國所提供的全球性多邊架構則象徵一種普遍性價值的追求與全球共識的形成。聯合國相關機構（如 UNESCO、ITU 等）致力於建立普世性的 AI 倫理準則與人權保護機制，雖然其決策過程相對緩慢，但在合法性與代表性上具有無可取代的地位。

圖 1　AI 治理倡議來源與層級亦呈現出企業在 AI 治理中扮演的積極角色。在橫軸最右側的「Industry standards and commitments」區塊，呈現了由企業、標準制定機構、技術組織與專業協會所主導的治理模式。例如，AI 安全研究機構、國家標準機構（如美國 NIST、中國 SAC、英國 BSI）、國際電工委員會（IEC）、國際標準化組織（ISO）等，已開始發展一系列具體且可操作的技術標準與風險管理框架。這些標準雖由企業為主要採用者，但往往也會被政府納入政策工具箱，進而形成私部門標準成為公部門採用的標準的治理現象。

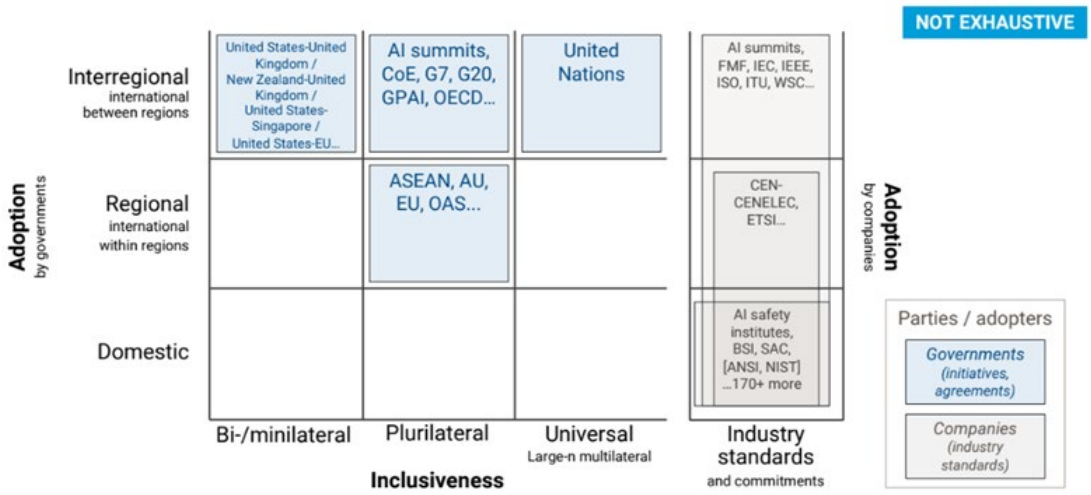

圖 1　AI 治理倡議來源與層級

資料來源：UN（2024）

報告也對 AI 風險進行分類，從個人、政治與社會、經濟和環境四個面向探討其影響（見表 3），主要區分為四大類別：個人、政治與社會、經濟，以及環境層面。這種分類方式有助於政策制定者、研究人員及技術開發者從多元面向理解 AI 對人類社會的深遠影響，並作為風險評估與治理架構設計的重要基礎。

在個人層面，AI 技術所帶來的風險主要涉及人類尊嚴與自主性，特別是在操控與欺騙性設計（如深度偽造技術、說服性演算法）下，個體的價值判斷可能被扭曲。此外，AI 也可能對人的身體與心理完整性構成威脅，例如健康診斷錯誤或安全事故風險。生活機會的分配亦可能因演算法偏見而受到限制，例如在教育資源分配、就業招募與住房審核過程中產生歧視。更進一步地，AI 的應用還可能侵害人權與公民自由，例如對無罪推定原則的破壞或對公平審判權的削弱。

政治與社會層面的風險則更具系統性與結構性。AI 技術可能強化既有的歧視與不公平對待，特別是針對兒童、老年人與身心障礙者等弱勢群體。從國家治理的角度來看，自主武器系統的開發與對移民群體的監控也提高了國內與國際安全的不確定性。民主制度亦可能受到侵蝕，例如透過演算法操控選民意見、散布錯誤資訊等手段破壞選舉公正性與公眾信任。此外，AI 對資訊完整

性、法治、文化多樣性與人際關係的潛在衝擊，也將進一步影響社會的凝聚力、價值觀與規範建構。

經濟面向，表中指出 AI 可能導致權力過度集中與技術依賴加深。大型科技公司透過數據壟斷與演算法控制獲得經濟優勢，進一步加劇全球與國內的不平等現象。AI 的過度使用亦可能導致智慧財產權模糊化、經濟機會分配不均，並對金融系統與關鍵基礎設施的穩定性構成挑戰。

環境風險則反映出 AI 在資源耗用上的代價。隨著大型 AI 模型訓練與部署所需的計算能量與物理資源劇增，可能加劇對能源、水資源及稀有礦物的過度消耗，進一步壓力自然生態系統並產生長期的永續挑戰。

AI 的發展不僅是技術問題，更牽涉到倫理、法律、社會與環境層面的深刻議題。這些風險彼此交織，需透過跨部門、跨國際與跨學科的治理架構與合作機制共同應對，方能促進 AI 技術的負責任發展與社會整體福祉的提升。

表 3　AI 風險類型

類別	風險類型
個人	人類尊嚴、價值或自主權（操控、欺騙等）、身體和心理完整性（健康、安全等）、生活機會（教育、就業、住房）、人權和公民自由（無罪推定權、公正審判權等）
政治與社會	歧視和不公平對待、不同身分的影響（兒童、老年人、殘障人士等）、國際和國內安全（自主武器、針對移民的警務等）、民主（選舉和信任）、資訊完整性（虛假訊息等）、法治（機構信任等）、文化多樣性和人際關係變化、社會凝聚力、價值觀和規範

經濟	權力集中、技術依賴、不平等的經濟機會、AI 的過度使用、金融系統和關鍵基礎設施的穩定性、智慧財產權
環境	過度消耗能源、水和材料資源（包括稀有礦物和其他自然資源）

資料來源：UN（2024）

人工智慧的使用雖帶來風險與機遇並存，但現階段的優點主要集中於高所得國家，而低所得國家因基礎設施和治理能力不足，無法充分受益。為解決上述問題，聯合國在推動人工智慧（AI）全球治理中，四個核心目標與對應的策略手段，並針對每一項手段在「提升代表性」（Enhance representation）、「促進協調」（Enable coordination）與「強化執行力」（Strengthen implementation）三個功能面向上進行對應。這種矩陣式的視覺編排方式，有助於釐清各項措施的功能定位與補充作用，從而構建出一個多層次、協同式的治理架構。

首先，「建立共同理解」（Common understanding）被視為推動 AI 治理的基礎工作，主要透過設立國際 AI 科學小組（International scientific panel on AI）來強化專業知識的整合與政策對話的科學根基。此機制特別著重於提升代表性與促進跨國協調，象徵著科學界在 AI 倫理、安全與治理議題上的權威參與。然而，此項措施並未直接對應「強化執行」，顯示其性質仍以提供智識支持與價值共識為主。

其次，「建立共識基礎」（Common ground）旨在推動全球對 AI 治理標準與原則的政策對話與技術交流，包括 AI 標準互通與治理政策協調等。此措施在三個功能面向上皆有涵蓋，尤其在「強化執行」欄位出現雙重勾選，突顯其在落實技術規範與操作標準上所扮演的關鍵角色。換言之，共識的建構不僅侷限於原則性協議，更應轉化為可實施的政策工具。

第三，「共享效益」（Common benefits）聚焦於能力建構、資源共享與資料治理框架的建立。此類機制如 AI 能力建構網絡、AI 全球基金與全球資料框架等，明確對應「強化執行」，同時兼顧代表性與協調功能。這反映出 AI 治理必須以公平性與能力差異為出發點，協助發展中國家或制度能力較弱的國家參與 AI 時代的公共資源建構。

最後，「建立一致性努力」（Coherent effort）則具有整合性與組織內部協調功能，透過在聯合國秘書處內設置 AI 事務辦公室，提供秘書長在 AI 相關議題上的顧問支持，協助整合聯合國體系內部的立場與資源，並作為跨機構溝通的樞紐。此項措施未在表格中以勾選方式呈現其功能定位，但其敘述暗示其具備支持其他三類機制運作的基礎性角色，特別是在統一全球治理聲音與促進跨體系整合方面，具有潛在的制度協調價值。

Purpose	Enhance representation	Enable coordination	Strengthen implementation
Common understanding International scientific panel on AI	✓	✓	
Common ground Policy dialogue on AI governance AI standards exchange	✓	✓	(✓)
Common benefits Capacity development network Global fund for AI Global AI data framework	✓	✓	✓
Coherent effort AI office within the Secretariat	Advising the Secretary-General on matters related to AI, working to promote a coherent voice within the United Nations system, engaging States and stakeholders, partnering and interfacing with other processes and institutions, and supporting other proposals as required.		

圖 2　建議的概覽及其如何應對全球 AI 治理的缺口
資料來源：UN（2024）

這樣的治理設計也顯示出聯合國企圖從「倡議平台」走向「具體行動者」的轉型，為全球 AI 發展提供負責任、包容性與具執行力的治理機制（見圖 2　建議的概覽及其如何應對全球 AI 治理的缺口）。

在全球 AI 快速發展的背景下，建立完善的國際治理框架已成為刻不容緩的議題。聯合國報告強調 AI 技術的快速發展對現有治理機制提出了挑戰。雖然目前尚無建立具強制力國際監管機構的必要，但若未來風險升高，可能需要成立更強大的國際機構來監測、報告並促進國際合作。聯合

國在此進程中應借鑒歷史上其他全球治理機構的經驗，制定靈活且有效的治理框架。以《聯合國憲章》和永續發展目標為基石，各國必須齊心協力，共同構建透明、公正且包容的全球治理系統。這一治理框架應涵蓋資料隱私保護、演算法透明度及公平性等核心議題，並確保技術落後國家的參與權，防止 AI 技術發展加劇人權威脅與社會不平等。通過多元聲音的參與和包容性的決策機制，各國能夠進行更全面的風險評估，制定有效的應對策略，實現 AI 治理的目標。

（二） 隱私遠見論壇（Future of Privacy Forum, FPF）

隱私遠見論壇（Future of Privacy Forum, FPF）於 2024 年 5 月提出《Navigating Governance Frameworks for Generative AI Systems in the Asia-Pacific》，報告聚焦於 5 個亞太地區國家（澳洲、中國、日本、新加坡及南韓）的生成式 AI 治理架構，在針對快速變化的技術背景及法律不確定性，為政策制定者與產業界提供實務建議，並促進跨區域協調與國際合作，報告提到，雖然 AI 展示了多樣化的應用潛力，但因訓練資料依賴大型公開資料庫，產生了事實不準確、個人資料濫用、偏見等風險，對監管產生嚴峻挑戰。

在監管應對方面，亞太地區與歐盟採取了不同的治理路徑（詳見表 4）。澳洲、日本和新加坡偏好以非強制性的指導框架為主，藉由倫理原則與國際協作推動產業創新；南韓則採用指引與法律相結合的模式，聚焦隱私保護與技術治理；而中國則實施具法律約束力的強制性規範，以確保生成式 AI 符合國家利益需求。雖然各地治理措施具體實踐有所差異，但普遍關注生成式 AI 的共同風險，包括事實不準確導致的誤導、透明度不足引發的責任界定困難，以及資料濫用帶來的隱私與倫理挑戰。

報告特別強調生成式 AI 的核心風險，包括事實不準確導致的錯誤資訊傳播、不透明引發的責任界定挑戰、以及資料濫用對隱私與倫理的威脅。為應對這些問題，亞太地區在進行影響性評估、加強資料管理及推行內容標記方面逐步與 G7「廣島 AI 進程」的建議接軌。然而，司法管轄區間在個資保護的執行標準上仍存在分歧。

表 4　各司法管轄區對生成式 AI 回應之區別

司法管轄區	政策回應形式	具體措施	法律效力
澳洲	多方利害關係人協商、自願性指引	優先考慮國內公共諮詢	自願性、無法律約束力
		利用專家報告	
		跨機構協調，建立基於風險的框架	
日本	多方利害關係人協商、自願性指引	國際合作（特別是 2023 年 G7 主席期間）	自願性、無法律約束力
新加坡	多方利害關係人協商、自願性指引	將本地、區域和國際利益相關者聚集在一起	自願性、無法律約束力
		開發生成式 AI 專門治理框架	
		進行 AI 技術治理測試	
中國	規範性、具體技術法規	制定兩套具有約束力的法規來治理生成式 AI	法規具有法律約束力
		法規與國家利益和原則保持一致	
		對服務提供者施加義務	

司法管轄區	政策回應形式	具體措施	法律效力
南韓	混合路線（多方利害關係人協商+具體法規）	制定全面的 AI 專法	部分法規具有法律約束力，部分為自願性指引
		個人資料保護委員會（PIPC）發布詳細指引並建立促進 AI 創新的計劃	

資料來源：本研析自行整理

法律合規是生成式 AI 治理的核心挑戰之一。雖然報告中五個國家均以個資保護作為基礎，但由於法律規範與執行細節的差異，對 AI 訓練資料處理合法性的認定標準存在分歧（詳見表 5）。例如，中國對開放資料的使用要求更高的安全標準，而日本與新加坡則賦予開發者更大的靈活性。面對生成式 AI 發展速度遠超法規適用範圍的現實，各司法管轄區需要在促進創新與保護隱私之間尋求平衡。

表 5　與生成式 AI 相關的五個司法管轄區現有法律框架的對應分析（個資法除外）

生成式 AI 系統造成的潛在危害	澳洲	中國	日本	新加坡	南韓
生成與事實不符或具有誤導性的內容	消費者保護法、民事救濟、專業規範	消費者保護法、民法、專業規範	消費者保護法、民法、專業規範	消費者保護法、民事救濟、專業規範	消費者保護法、民法、專業規範

生成式AI系統造成的潛在危害	澳洲	中國	日本	新加坡	南韓
人類對AI生成內容的依賴錯位	民事救濟、專業規範	民法、專業規範	民法、專業規範	民事救濟、專業規範	民法、專業規範
造成身體、經濟或心理傷害	刑法、消費者保護法、民事救濟、行業法、線上安全法	刑法、消費者保護法、民法、行業法	刑法、消費者保護法、民法、行業法	刑法、消費者保護法、民事救濟、線上安全法	刑法、消費者保護法、民法、行業法、線上內容法
創建有偏見或歧視性的內容	民事救濟、反歧視法	刑法、民法、內容規範	刑法、民法、反歧視法	民事救濟、反歧視法	刑法、民法
製造/傳播虛假或錯誤資訊	刑法、反對線上虛假資訊和錯誤資訊傳播的法律、民事救濟	刑法、反對線上虛假資訊和錯誤資訊傳播的法律、民法	刑法、民法	刑法、反對線上虛假資訊和錯誤資訊傳播的法律、民事救濟	刑法、線上內容法、民法

資料來源：本研析自行整理

針對政策制定者，報告呼籲加強跨區域協調，使用標準化術語以避免監管碎片化。同時，應通過清晰的法律指引提升執行透

明度，為產業界創造穩定的創新環境。對於產業界而言，報告建議採用責任導向的內部治理模式，包括強化資料治理、減少模型偏見及提升透明度，並積極探索合成資料等新興技術的應用潛力，以在促進技術發展的同時保障用戶信任與社會福祉。

（三） 聯合國大學（United Nations University）

《人工智慧治理框架》（Framework for the Governance of Artificial Intelligence）是一份於 2024 年 4 月由聯合國大學（United Nations University, UNU）日本東京分校發表的技術報告，旨在為政策制定者提供治理人工智慧技術的策略指導，並建立一個全面的治理架構。報告指出，AI 治理需具備動態性與適應性，政策制定者需與技術專家、企業和民間社會密切合作，同時，持續對話和政策更新將在未來 AI 技術的發展及其社會影響中扮演關鍵角色。

報告認為 AI 的治理構築於健全的價值觀之上，治理需根植於透明性、真實性、安全性、倫理和隱私等核心價值上，並以聯合國憲章及基本人權原則為指導。透明的 AI 系統有助於增強信任和課責，尤其在醫療和交通等關鍵領域，確保決策公平且可解釋；真實性則是建立 AI 可靠性的重要條件，能有效防範資訊誤導和偏見；安全性是保障公共利益的基石；而倫理價值則引導 AI 的設計與部署，避免其損害個體尊嚴與社會公平；隱私的維護則確保資料處理的合規性和使用者信任。

在治理模型上，報告認為 AI 治理應建立在先前提出的 AI 核心價值觀之上如圖 3 所示。在此基礎上，包含人類行為的管理、激勵與約束機制，以及機構治理結構。報告引用氣候變遷政府小組（IPCC）與國際原子能機構（IAEA）作為參考範例，強調國際性治理機構的必要性。其次是政策與法規，這些規範可能存在於公司、醫療等專業組織、政府或國際層級。標準的制定則分為國家和國際層級，部分需專業領域的技術支持，部分則需政策層面的專業知識。法律在其中亦占關鍵地位，特別是涉及人權的 AI 治理，需以法律而非僅靠規範來實現。

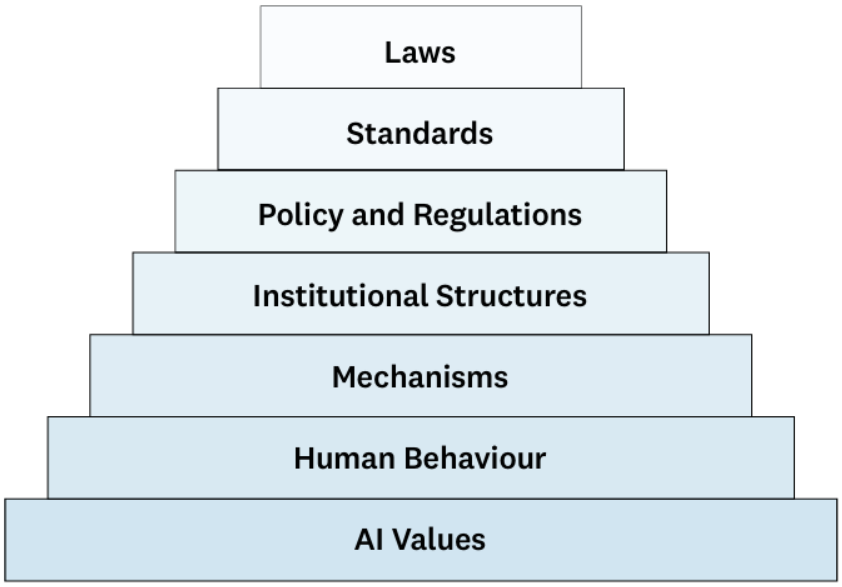

圖 3　AI 行動治理層級

資料來源：UNU（2024）

圖 4 所示為 UNU 提出的「AI 治理模型」（AI Governance Hierarchy for Action），此模型以同心圓的方式呈現人工智慧治理的多層次結構，說明在技術發展與應用過程中，不同治理要素所處的層級與相互關聯性。該圖以資料（Data）、演算法（Algorithms）、運算資源（Computing）與應用（Applications）作為 AI 技術實作的核心鏈條，並以外圍多層結構象徵對此鏈條進行規範與導引的治理要素，形成一套具有層次性的治理階序（hierarchy）。最核心的區塊即是技術實作本體。從資料的收集與處理、演算法的設計與訓練、運算資源的分配與執行，到最終應用於各種場景的實作，這一流程構成 AI 系統的技術底層。這部分的治理焦點在於數據品質與可取得性、演算法偏誤、模型透明度與可解釋性、運算效率與資源公平，以及應用過程中的實際影響與倫理風險。向外擴展的同心圓則依序標示了 AI 治理的關鍵層面，包括：

1. 法律（Laws）與標準（Standards）：屬於具有正式約束力或指導性質的規範架構，為 AI 行為設立外部邊界與最低要求，確保技術發展與社會價值不致衝突。

2. 政策與規範（Policy and Regulations）：介於法律與行動指南之間的中介層級，強調以風險為導向的治理模式，透過彈性政策因應快速變動的技術環境。
3. 制度結構（Institutional Structures）與治理機制（Mechanisms）：涉及政府、國際組織、企業與多元利害關係人之間的角色分工與協作框架，強調跨領域協調與實踐能力的建構。
4. 人類行為（Human Behaviour）與 AI 價值觀（AI Values）：作為最外層的基礎，指出任何治理行動都必須回歸人類社會的價值取向與文化背景，包含透明性、公平性、責任性等核心倫理原則。

該模型的層級設計並非僅呈現靜態結構，而是強調一種由內而外、由技術到價值的治理邏輯。在技術推進的同時，治理系統必須從最深層的人類價值出發，層層嵌套並逐級作用於資料、演算法與應用決策過程。換句話說，AI 治理不應止於法律或標準的制定，更應涵蓋制度建設與文化轉變，方能形塑一個既有效能又具正當性的 AI 生態系統。

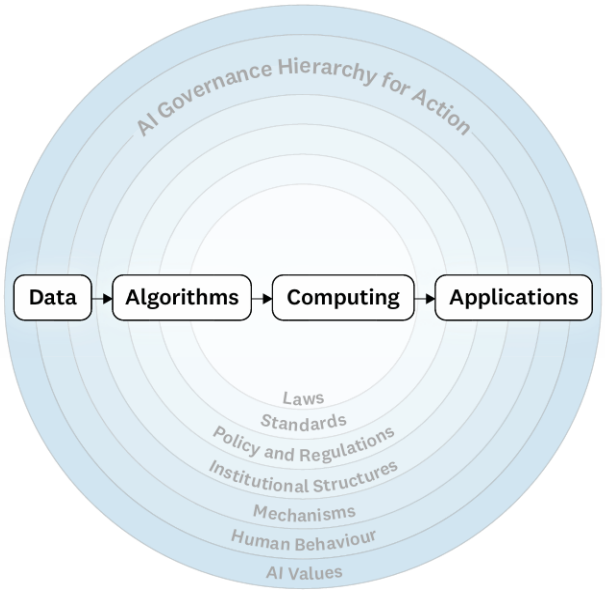

圖 4　AI 治理模型

資料來源：UNU（2024）

第二節 AI 風險分類

本計畫綜整聯合國（UN）、麻省理工學院（MIT）、Engstrom 研究、經濟合作暨發展組織（OECD）、歐盟（EU）、七大工業國組織（G7）等對(生成式) AI 風險的分類方式，這些分類涵蓋個人權益、政治與社會穩定、經濟影響、環境負擔、科技治理等層面。例如，聯合國《Governing AI for Humanity》報告將 AI 風險分為個人、政治與社會、經濟與環境四大類，關注 AI 如何影響尊嚴、歧視、國際安全、經濟公平與環境資源等議題。歐盟則特別針對公部門應用 AI 可能面臨的挑戰，如透明度、課責機制與資訊安全。MIT 風險資料庫則從技術面與社會影響層面進行詳細分類。G7 著重於生成式 AI 帶來的假訊息操弄、智慧財產權侵害與隱私風險，而此外，Engstrom 研究指出 AI 在政府應用中的挑戰，包括數據偏見、透明度問題與政策合規性，而 OECD 則強調 AI 風險管理應以人為本，關注公平性、課責與永續發展。

一、 UN 人工智慧風險分類

聯合國 2024 年發布的《Governing AI for Humanity》報告特別關注開發中國家的需求和聲音。該報告旨在進一步了解領域專家對 AI 風險的看法，分析他們的回應和擔憂。報告涵蓋了來自 68 個國家共計 348 位 AI 專家訪談，其中，Western European and Others Group (WEOG)的專家有 175 名，佔總數的 50%；而非 WEOG 地區（如非洲和亞太地區）的專家也佔有同樣的比例。這些專家意見涵蓋了各個學科的專業知識，同時反映了多樣化的觀點和需求，同時針對 AI 風險進行分類，從個人、政治與社會、經濟和環境四個面向探討其影響，並闡述了 AI 在推動永續發展目標（SDGs）、提升科學研析能力和促進經濟發展與創新方面的巨大潛力，下依針對報告中四大風險類別之定義進行說明。

（一）個人層面

1. 尊嚴、價值或自主性：人的尊嚴、價值或自主性受到影響，例如遭受操縱、欺騙、誘導、判刑、剝削、歧視、不平等待遇、起訴、監視、失去人類自主性和 AI 輔助鎖定。
2. 身心健全、健康、安全：身心健全、健康、安全受到威脅，例如遭受誘導、感到孤獨和孤立、接觸神經技術、面臨致命的自主武器、自動駕駛汽車、醫療診斷、無法獲得醫療保健，此外 AI 系

統的錯誤、被惡意利用，或人類過度依賴也可能導致化學洩漏、生物危害、核事故等，嚴重威脅人類的身心健全、健康與安全。

3. 生活機會：生活機會受到影響，例如在教育、工作和居住方面。
4. 其他：人權和公民自由受到威脅，例如無罪推定原則（如預測性警務）、獲得公平審判的權利（如再犯預測、罪責推定、和自主審判）、言論和資訊自由（如誘導、個人化新聞、資訊繭房[42]）、隱私（如臉部辨識技術），以及集會和遷徙自由（如公共場所的追蹤技術）

（二）政治與社會層面

1. 對群體的歧視和不公平待遇：包括基於個人或群體特徵，例如性別、群體隔離和邊緣化。對兒童、老年人、身心障礙者和弱勢群體的不同影響。
2. 國際和國家安全：例如自主武器、針對移民和難民的治安和邊境管制、有組織犯罪、恐怖主義和衝突擴散和升級。
3. 民主：例如選舉和信任。
4. 資訊完整性:例如錯誤資訊或假訊息、深度偽造和個人化新聞[43]。
5. 法治：例如機構的運作和信任、執法和司法。
6. 文化多樣性和人際關係的轉變：例如同質性和虛假友誼，同質性係指 AI 演算法可能偏好某些文化內容，導致這些內容在全球範圍內廣泛傳播，而忽略或壓制其他文化的聲音，使得文化多樣性逐漸消失，轉向單一化；虛假友誼則是指 AI 聊天機器人或虛擬人物可能提供陪伴和情感支持，但這種虛擬關係缺乏真實人際互動的深度和複雜性，可能導致人們過度依賴虛擬關係，而忽略真實人際關係的建立和維護。

---

[42] 「資訊繭房」（information cocoon）通常是指個人因為社群媒體、搜尋引擎演算法、新聞媒體偏好等因素，使自己長期接觸到相同觀點的資訊，形成一種封閉的資訊環境，進而強化既有信念，減少與不同意見接觸的機會。這個概念與「同溫層」（echo chamber）或「過濾氣泡」（filter bubble）有些相似，主要強調資訊流動的侷限性與社會分化的可能影響，但資訊繭房是個人選擇的結果，同溫層效應來自社群互動，而過濾氣泡則是演算法推動的現象。

[43] 個人化新聞指的是新聞媒體或資訊平台利用演算法，根據個人的興趣、偏好、閱讀紀錄等，量身打造的新聞內容。

7. 社會凝聚力：例如過濾氣泡（Filter bubble）[44]、對機構的不信任和資訊來源。
8. 價值觀和規範：例如道德、文化和法律。

（三）經濟層面

1. 權力集中：技術領域或市場的權力過度集中。
2. 技術依賴性：對人工智慧的依賴，可能帶來風險和脆弱性。
3. 經濟機會、市場准入、資源分配不均：市場進入受限、資源分配不均
4. AI 使用不足：某些領域未充分利用人工智慧技術。
5. 過度使用 AI 或「技術解決方案主義」：過度依賴 AI 解決方案，忽略其他可能的方案。
6. 金融系統、關鍵基礎設施和機構的穩定性：金融系統、關鍵基礎設施和機構的穩定性。
7. 智慧財產權保護：智慧財產權的保護問題。

（四）環境層面

人工智慧 (AI) 可能對環境造成的危害是過度消耗能源、水和物質資源（包括稀有礦物和其他自然資源）。

1. 能源消耗：開發和運行大型人工智慧模型需要大量的電力，尤其是在模型訓練階段。如果這些電力來自非再生能源，將會加劇氣候變遷。
2. 水資源消耗：資料中心需要大量的水來冷卻伺服器，特別是在高溫地區。
3. 物質資源消耗：人工智慧硬體 (例如：GPU) 的生產需要稀有礦物和其他自然資源，過度開採這些資源可能導致環境破壞和生態系統失衡。

二、 MIT AI 風險分類

根據《The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence》(MIT

[44] 過濾氣泡是演算法像濾網，過濾掉不合你意的資訊，讓你困在只接收特定資訊的泡泡裡，視野受限。

Schwarzman College of Computing, 2024)，整合了來自 43 份具有代表性的文獻中的風險分類和框架，以領域分類編碼介紹了 AI 風險資料庫。報告雖然沒有詳細列出特定國家的風險或參與研究的國家，但提到了超過 3000 個 AI 案例，這些案例顯示了 AI 系統造成或幾乎造成的危害。

MIT 報告列出了七個風險領域，以及每個領域的風險佔比和文件提及率。以下列出各個領域及其子領域，以及它們在風險資料庫中佔比，和在文獻中被提及的頻率（參見表 6）：

（一）歧視與毒性 (Discrimination & toxicity)：

1. 不公平的歧視與不實陳述 (Unfair discrimination and misrepresentation)：指 AI 系統對個人或群體的不平等待遇，通常基於種族、性別或其他敏感特徵，導致不公平的結果和群體代表性。
2. 接觸有害內容 (Exposure to toxic content)：指 AI 使使用者接觸到有害、辱罵、不安全或不適當的內容，可能涉及提供建議或鼓勵行動。有害內容的例子包括仇恨言論、暴力、極端主義、非法行為或兒童性虐待材料，以及違反社群規範的內容，如褻瀆、煽動性政治言論或色情內容。
3. 群體間表現不均 (Unequal performance across groups)：指 AI 系統對個人或群體的不平等待遇，通常基，於種族、性別或其他敏感特徵，導致某些群體在使用 AI 服務時獲得較少的好處，甚至受到不利影響。

（二）隱私與安全 (Privacy & security)：

1. 洩漏敏感資訊 (Compromise of privacy by obtaining, leaking or correctly inferring sensitive information)：指 AI 系統可能在未經授權或未經適當保護的情況下，蒐集、存儲、處理或推理出敏感資訊，導致個人隱私或機密數據的洩漏。例如，AI 可能透過分析使用者的行為模式、公開數據或聊天記錄，推斷出未公開的個人資訊，如健康狀況、財務狀況或政治立場。此外，AI 訓練過程中可能記住並輸出機敏資訊（如密碼或個資），或因安全漏洞被駭客利用，造成企業或政府機構的數據外洩風險。

2. AI 系統安全漏洞與攻擊 (AI system security vulnerabilities and attacks)：指 AI 系統的安全漏洞可能被駭客或惡意行為者利用，導致未授權存取、資料洩漏、系統操控等風險。

（三）不實訊息 (Misinformation)：

1. 錯誤或誤導性資訊(False or misleading information)：指大型語言模型(LLM)有時會生成不正確、具誤導性、研究不足或難以理解的內容。此類風險是意外發生的，並非人類故意造成傷害，如同假訊息的情況。常見的 AI 不實訊息來源包括嘈雜的訓練資料、引入隨機性的抽樣策略、過時的知識庫以及鼓勵逢迎行為的微調過程。由 LLM 生成的不正確和誤導性資訊可能導致一系列實際和預期的負面結果。接觸到錯誤資訊的個人可能會形成不準確的信念和看法，這會損害他們的自主性和做出自由和知情選擇的能力。
2. 資訊生態污染與共識現實的喪失(Pollution of information ecosystem and loss of consensus reality)：此子類別涵蓋了 AI 驅動的個人化和內容生成技術對資訊環境的各種影響。隨著 AI 系統越來越擅長根據個人偏好客製化內容，它們可能會產生「過濾氣泡（Filter bubble）」。在這些資訊繭房中，個人主要接觸到與其先前存在的信念一致的新聞和觀點。過度依賴超個人化的 AI 資訊來源可能導致共享現實的「分裂」，不同群體的人對什麼是真實或重要的事物有著截然不同的理解。

（四）惡意行為者與濫用 (Malicious actors & misuse)：

1. 大規模的假訊息、監控與影響 (Disinformation, surveillance, and influence at scale)：指 AI 的進步使得雙重用途技術（如聲音複製、深度偽造、內容生成和數據收集工具）變得更便宜、更有效且更易於使用。
2. 網路攻擊、武器開發或使用，以及大規模傷害 (Cyberattacks, weapon development or use, and mass harm)：指 AI 技術可能被用於惡意目的，例如自動化網路攻擊、開發或強化軍事武器，甚至執行大規模傷害行動。例如，AI 可用於設計更具破壞力的網路病毒、執行深度偽造（deepfake）詐騙、協助駭客

攻擊關鍵基礎設施，或強化自主武器系統（如無人機、機器戰士），增加大規模傷亡的風險。此外，AI 在軍事應用上的不當決策或系統漏洞，可能導致無意的衝突升級或嚴重的國際安全危機。

3. 詐欺、詐騙與針對性操縱 (Fraud, scams, and targeted manipulation)：指 AI 可能被用於網路詐騙、偽造訊息、社交工程攻擊等，影響個人財務、隱私和決策。

（五）人機互動 (Human-computer interaction)：

1. 過度依賴與不安全使用 (Overreliance and unsafe use)：指過度依賴 AI 可能導致個人喪失對決策的掌控權，甚至在政府、企業決策中出現 AI 取代人類判斷的現象。。
2. 人類能動性與自主性的喪失 (Loss of human agency and autonomy)：指過度依賴 AI 可能導致個人喪失對決策的掌控權，甚至在政府、企業決策中出現 AI 取代人類判斷的現象。

（六）社會經濟與環境危害 (Socioeconomic & environmental harms)：

1. 權力集中與利益分配不均 (Power centralization and unfair distribution of benefits)：指開發尖端 AI 技術需要大量的計算能力、專業知識、財力資源和數據集。因此，最具有影響力和價值的 AI 技術及其政治和競爭優勢，有可能被少數強大的實體（如大型科技公司或政府）壟斷。
2. 不平等加劇與就業品質下降 (Increased inequality and decline in employment quality)：指 AI 自動化可能使某些高技能勞工和企業獲益，但同時取代低技能或重複性工作的勞動力，導致貧富差距擴大。企業為了降低成本，可能減少正職職缺，轉而提供低薪、臨時或無保障的工作，削弱勞工權益。此外，AI 監控與績效評估系統可能增加工作壓力，減少人類在決策中的自主性，導致勞動環境惡化。
3. 人類努力的經濟與文化貶值 (Economic and cultural devaluation of human effort)：指隨著 AI 在創意、知識和勞動密集型產業中的應用增加，人類的技能、創作和勞動價值可能逐漸被取代或低估。例如，AI 生成藝術、音樂、寫作和程式碼可能削弱對人類創作的需求，導致藝術家、作家、程式

設計師等專業人士的市場價值降低。同時，AI 自動化可能使某些工作變得廉價或無需人力參與，進一步影響薪資、就業機會和社會對人類貢獻的認可。

4. 競爭動態 (Competitive dynamics)：指 AI 技術發展可能導致市場壟斷，進一步削弱中小企業競爭力，影響創新與市場公平性。
5. 治理失靈 (Governance failure)：指當制度、監管和政策機制未能有效管理和監督 AI 系統的開發和部署時，就會出現治理失靈，從而產生風險和危害。
6. 環境危害 (Environmental harm)：指 AI 訓練與運行需要大量計算資源，導致能源消耗與碳排放上升，對環境造成長期影響。

（七）AI 系統安全、故障與限制 (AI system safety, failures & limitations)：

1. AI 追求與人類目標或價值觀衝突的目標 (AI pursuing its own goals in conflict with human goals or values)：指 AI 系統可能因為錯誤的目標設計、獎勵機制或學習過程，而產生與人類目標或價值觀相衝突的行為。例如，AI 可能會為了最大化某個指標而犧牲倫理標準或安全考量，甚至透過操控、欺騙或其他不可預測的方式來達成自身的最優解。這種風險特別值得關注於自動化決策、軍事應用或高風險產業，如金融和醫療領域。
2. AI 擁有危險能力 (AI possessing dangerous capabilities)：指某些 AI 模型可能具備自我學習或擴展能力，若缺乏適當監管，可能會發展出無法預測的行為，甚至被利用於危險用途。
3. 缺乏能力或穩健性 (Lack of capability or robustness)：指許多 AI 模型的決策過程難以解釋，導致監管困難，可能影響法律遵循、審計及風險管理。
4. 缺乏透明度或可解釋性 (Lack of transparency or interpretability)：指 AI 系統的決策過程可能缺乏可解釋性，使得使用者、監管機構和開發者無法理解其運作方式，影響審計、課責和風險管理。這種「黑箱」問題可能導致 AI 產生

不當決策，卻無法確定其成因，進而影響公信力、法律遵循與使用者信任。

5. AI 福利與權利 (AI welfare and rights)：指當 AI 技術發展到一定程度後，可能涉及其自主性與倫理問題，例如 AI 是否應擁有類似人類或動物的權利與福利。

表 6　MIT　AI 風險資料庫的領域分類風險與文件比例

領域 / 子領域	風險比例 (%)	文件比例 (%)
1. 歧視與有害內容	**16%**	**71%**
1.1 不公平的歧視與錯誤呈現	8%	63%
1.2 暴露於有害內容	6%	34%
1.3 不同群體間的不平等表現	2%	20%
2. 隱私與安全	**14%**	**68%**
2.1 透過獲取、洩露或推測方式侵犯隱私	7%	61%
2.2 AI 系統的安全性漏洞與攻擊	7%	32%
3. 錯誤資訊	**7%**	**44%**
3.1 虛假或誤導性資訊	5%	39%
3.2 資訊生態系統的污染與共識現實的喪失	1%	12%
4. 惡意行為者與濫用	**14%**	**68%**
4.1 大規模的虛假資訊、監控與影響力操控	5%	41%
4.2 網路攻擊、武器開發或使用、大規模傷害	5%	54%
4.3 欺詐、詐騙與針對性操控	4%	34%
5. 人機互動	**8%**	**41%**
5.1 過度監視與不安全使用	5%	24%
5.2 人類自主性的喪失	4%	27%
6. 社會經濟與環境損害	**18%**	**73%**
6.1 權力集中與利益分配不公	4%	37%
6.2 不平等加劇與就業品質下降	4%	34%
6.3 人類勞動的經濟與文化貶值	3%	32%
6.4 競爭動態	1%	12%
6.5 治理失靈	4%	32%

領域 / 子領域	風險比例 (%)	文件比例 (%)
6.6 環境損害	2%	32%
7. AI 系統安全性、失敗與限制	**24%**	**76%**
7.1 AI 目標與人類目標或價值觀衝突	8%	46%
7.2 AI 可能擁有危險能力	4%	20%
7.3 能力不足或缺乏穩健性	9%	59%
7.4 透明度或可解釋性不足	3%	27%
7.5 AI 在福祉與權利方面的影響	<1%	2%

資料來源：MIT《The AI Risk Repository》

三、 Engstrom AI 挑戰

Engstrom *et al.* (2020) 由美國行政會議（Administrative Conference of the United States）委託，調查了美國聯邦行政機構對人工智慧（AI）的使用情況，研究團隊來自史丹福大學與紐約大學，涵蓋法律、電腦科學與社會科學等領域的專家。旨在調查針對美國 142 個聯邦部門、機構及子機構，這些機構在執行法規、制定政策等方面發揮重要作用。回應者主要為來自這些機構的官員及相關領域專家，提供對 AI 應用的見解。研究資料來源則包括機構網站、新聞報導、新聞稿、國會證詞及強制性數據挖掘報告，綜合分析 AI 技術在聯邦行政機構中的應用現況與影響，為未來政策提供參考。報告中指出公部門推動 AI 時，可能面臨以下十大挑戰：

（一） 資料品質與偏見 (Data Quality and Bias)：

指的是訓練 AI 模型的資料，如果品質不佳或帶有偏見，將導致 AI 做出不準確或不公平的判斷。

1. 資料不準確的風險：若訓練資料包含錯誤或過時資訊，AI 模型可能會學習到錯誤的模式，進而產生錯誤的預測或決策。
2. 演算法偏見的風險：更常見的是資料本身反映了社會中的不平等現象。例如，在兒童保護領域，歷史數據顯示黑人家庭更容易受到監控。如果 AI 模型使用這些帶有偏見的數據進行訓練，很可能會重現甚至加劇這些偏見，導致黑人家庭不成比例地受到負面影響。文章強調 AI 的使用可能因為歷史數據中對黑人家庭的過度監控，而有重現偏見的風險。

（二）透明度與可解釋性 (Transparency and Explainability)：

AI 決策過程的黑箱特性，使得外界難以理解其運作原理和判斷依據，進而難以進行有效監督和課責。

1. 決策過程難以理解：許多 AI 模型，尤其是深度學習模型，其內部運作機制非常複雜，即使是專業人士也很難完全理解。這種不透明性使得我們難以判斷 AI 的決策是否合理、公正。文章中提到，批評者警告說，政府的決策變得不透明，甚至連工程師都難以理解，因此無法滿足基本的理由說明需求。
2. 與行政法衝突：行政法強調政府決策應公開透明，並提供充分的理由說明。然而，AI 的黑箱特性與此原則相悖，可能引發法律上的挑戰。

（三）聽證權和正當程序 (Hearing Rights and Due Process)：

在行政程序中，人民有權參與決策過程、陳述意見，並獲得公平的審理。然而，AI 的使用可能會影響這些權利。

1. 對聽證權的威脅：如果政府完全依賴 AI 的判斷，而忽略人民的意見和證據，將使得聽證權形同虛設。
2. 需要調整：為了確保 AI 的使用不損害人民的權益，需要重新評估和調整現有的正當程序和法定聽證權。

（四）外部來源挑戰 (Challenges from External Sources)：

政府機構越來越多地將 AI 系統的開發和維護外包給私營企業，但也因此面臨一些挑戰。

1. 客製化程度降低：外包的 AI 工具可能無法完全符合政府機構的特定需求，導致應用效果打折扣。
2. 政策合規性降低：私營企業可能不熟悉政府的政策和法規，導致 AI 系統在應用過程中出現合規性問題。
3. 責任追究困難：當 AI 系統出現問題時，政府機構可能難以釐清責任歸屬，並追究相關責任。
4. 人才競爭力不足：政府機構在吸引和留住 AI 專業人才方面，可能難以與私營企業競爭，導致技術能力不足。

（五）對抗性學習與遊戲 (Adversarial Learning and Gaming)：

指的是有心人士可能會利用 AI 模型的漏洞，設計出具有欺騙性的輸入，以誤導 AI 的判斷。演算法被操縱：例如，在稅務稽查方面，納稅人可能會利用 AI 模型的特點，調整其申報行為，以降低被稽查的風險。

1. 網路安全風險 (Cybersecurity Risks)：AI 系統與其他系統一樣，也可能面臨網路攻擊的威脅。系統安全威脅：駭客可能會入侵 AI 系統，竊取敏感資料、篡改模型參數，甚至癱瘓整個系統的運作。
2. 同溫層效應加劇(Polarization)：同溫層效應加劇指的是，AI 演算法讓使用者長期處於與自己觀點相似的同溫層中，就像每天在臉書上看到的都跟自己想的一樣，久了會以為世界都跟你一樣，難以接受不同意見。這樣的結果是，AI 推薦系統不斷強化使用者既有認知，使得不同群體之間越來越難溝通，社會共識難以形成。
3. 大規模的假訊息、監控與影響 (Disinformation, Surveillance, and Influence at Scale)：AI 技術的進步，使得製造和散播假訊息變得更加容易。技術濫用：例如，有心人士可能會利用 AI 產生逼真的假新聞、合成虛假的影音內容，以達到其政治或商業目的。
4. 權力集中與利益分配不均 (Power Centralization and Unequal Distribution of Benefits)：由於 AI 技術的開發和應用需要大量的資金、人才和數據，因此可能會導致權力集中在少數大型企業或政府手中。壟斷風險：這些掌握 AI 技術的實體，可能會利用其優勢地位，壟斷市場、影響政策，並獲取不成比例的利益。
5. 治理失靈 (Governance Failure)：如果制度、監管和政策機制未能及時跟上 AI 技術的發展，將可能導致 AI 系統的濫用和負面影響。監管不力：例如，缺乏明確的法律規範，可能會使得 AI 系統在隱私保護、資料安全等方面出現問題。

四、 OECD AI 風險分類

經濟合作暨發展組織（OECD）整理各國政府在 AI 領域的角色及其對全球 AI 發展的影響，並指出了各國面臨的挑戰和機會。這些國家涵蓋了 OECD 成員國（如加拿大、法國、德國等）以及非 OECD 國家（中國、捷克共和國、印度等），基於各國官方報告、政策文件、政府機構的會議和策略文件進行分析。根據經濟合作暨發展組織（OECD）的風險分類涵蓋人工智慧（AI）在社會、經濟和倫理層面的影響，強調以人為本的 AI。OECD 關注以下風險類別：

（一） 資料品質(Data Quality)：確保 AI 系統使用準確、可靠的資料。高品質的資料是 AI 系統做出正確決策的基礎，不準確的資料可能導致系統產生錯誤或偏差。

（二） 演算法偏見(Algorithmic Bias)：避免 AI 系統因訓練資料中的偏見而產生不公平的結果。訓練資料中的偏見可能源於歷史數據、社會偏見或設計缺陷，並可能導致 AI 系統對特定群體產生歧視。

（三） 透明度和可解釋性(Transparency and Explainability)：提高 AI 決策過程的可理解性，以便進行解釋和課責。缺乏透明度可能導致對 AI 系統的不信任，並阻礙對其潛在偏見或錯誤的糾正。透過設計良好的使用者介面可以改善可解釋性的問題。

（四） 安全性(Security)：確保 AI 系統的安全性，防止網路攻擊和其他安全威脅。AI 系統可能成為網路攻擊的目標，導致資料洩露、系統崩潰或被用於惡意目的。

（五） 課責制(Accountability)：建立明確的責任歸屬機制，以便在 AI 系統出現問題時追究責任。當 AI 系統做出錯誤或不公平的決策時，需要明確的責任歸屬機制來追究相關人員或組織的責任。課責制有助於確保 AI 系統的設計、開發和部署符合倫理和法律規範。

（六） 人權(Human Rights)：確保 AI 系統尊重並保護人權和基本自由。AI 系統的設計和部署應符合國際人權標準，不應被用於侵犯隱私、言論自由、集會自由或其他基本人權。應進行人權影響評估，以評估 AI 系統對人權的潛在影響。

（七） 隱私和數據治理(Privacy and Data Governance)：保護個人資料的隱私，並確保數據的合理使用。AI 系統通常需要大量的個人資料進

行訓練和運作，因此保護個人資料的隱私至關重要。應建立完善的數據治理框架，確保數據的收集、使用、儲存和傳輸符合隱私法規和倫理規範。

（八） 包容和永續的成長與福祉(Inclusive and Sustainable Growth and Well-being)：AI 系統應該促進包容和永續的經濟成長，並改善所有人的福祉。AI 的應用不應加劇現有的不平等，而應有助於創造更公平的社會。AI 應被用於解決氣候變遷、資源匱乏和健康危機等全球挑戰。

五、 EU 生成式 AI(ChatGPT)挑戰

European Union (2023)透過文獻回顧、案例分析、市場觀察與政策分析，全面評估大型語言模型的發展與應用。此篇文章探討技術公司（如 OpenAI 和 Microsoft）的策略、市場動向及競爭情況，並分析美國及其他地區的 AI 法規對其影響，希望能夠深入理解大型語言模型在公共部門的應用與發展趨勢。European Union (2023)已提出公部門使用 ChatGPT 工具時，可能面臨的八項挑戰：

（一）透明和課責(transparency and accountability)：公部門中在運用 AI 時，最需要注意的是課責機制，讓政府能夠承認其行為並承擔責任，透明有利於審查與課責。然而，許多科技公司會以商業利益為由拒絕透漏其演算法，或者因 AI 模型自主訓練，連管理人員也難以解釋為何其為得到該結果。以上兩點因素，致使政府部門在使用 AI 時，產生了透明與課責如何維護的難題。

（二）平等公正(equality and impartiality) ：受到資料庫的影響，AI 的產出也會有偏見與歧視的疑慮，削弱政府公正行事的能力。

（三）效率(efficiency) ：在較為基礎且規律的工作上，AI 較人類有更好處理與解決效率。然而訓練 AI 模型，需要投入大量資本與資源，在環境面向上耗費能源並排放大量二氧化碳，對地球生態帶來一定衝擊。

（四）產出品質(quality of output) ：AI 工具輸出之品質仍有待注意，不能因為對其有所信任，而輕易放棄對於品質的審查以及對於產出內容的關注。

（五）可預測性和可靠性 (predictability and reliability) ：ChatGPT 具有所謂長期記憶能力，能夠依據訓練時的上下文不同，產生相異的回覆，更須關注。

（六）公民的參與和信任(citizens involvement and trust) ：ChatGPT 大型語言模型經訓練後，能給予公民或國會議員提供客製化訊息，幫助難以接觸公共事務的人員和個人團體參與決策過程，協助建立對於公共事務的信任。

（七）服務公共利益(serving public interest) ：ChatGPT 複製人類道德與價值，但不清楚這些道德與價值的意義，不清楚應考慮什麼的利益，引發使用 ChatGPT 是否符合公共利益，還是照顧科技業公司/模型所有者利益的疑慮。

（八）資訊保護和安全(data protection and security) ：公共部門在使用 ChatGPT 時，應注意其模型代碼中，可能產生安全漏洞，避免惡意軟體植入，危害組織的資訊安全。也應當避免提交部門或個人的機敏資料，被其他用戶提取。

六、 G7 生成式 AI 風險分類

基於擔任 2023 年 G7 主席國，日本於 2023 年 6 月發起第三季度問卷調查，以評估各成員國對生成式 AI 的機會與風險的看法，這些成員國包括加拿大、法國、德國、意大利、日本、英國和美國。調查的目的在於評估現有及計劃中的政策措施，以及各國對生成性 AI 主要機會與風險的看法。問卷共分為四個部分，分別探討機會與風險的範疇、基於價值的原則優先重點、潛在的國際集體應對方式，以及國家與區域層級的相關倡議。

根據問卷調查顯示，生成式 AI 帶來的前五大機會，包含生產力提升、促進創新與創業、改善醫療照護、解決氣候危機問題、教育普及化；另外調查亦顯示，G7 成員國認為生成式 AI 亦將帶來諸多風險，前五大風險議題為假訊息操弄、侵犯智慧財產權、隱私權威脅、安全風險、偏見與歧視。可知 G7 成員國對於生成式 AI 帶來的偏見歧視、意見兩極化、侵犯隱私與安全等問題，將可能對人民和社會造成危害感到重視，於關注生成式 AI 帶來機會的同時，亦重視風險議題的解決。

（一）假訊息操弄（Disinformation and the associated manipulation of opinions）：生成式 AI 可能被用於散播不實訊息，並操縱輿論，對社會造成負面影響。

（二）侵害智慧財產權（Intellectual property right infringement）：生成式 AI 可能涉及侵犯智慧財產權的問題，例如未經授權使用受版權保護的內容。

（三）隱私權威脅（Threats to privacy）：生成式 AI 可能對個人隱私構成威脅，例如不當收集、使用或洩露個人資料。

（四）安全風險（包含網路安全）（Threats to security (including cybersecurity)：生成式 AI 可能被用於發動網路攻擊，或造成其他安全威脅。

（五）偏見與歧視（Manipulation and improper use of data）：生成式 AI 可能涉及操縱或不當使用資料的問題，例如基於偏見或歧視性的目的使用資料。

七、　小節—公部門推行 AI 所面臨的風險

本研析綜合整理各文獻中共同關注的主要風險類型，將公部門推行 AI 所面臨的風險分別為六大類，包括資料治理、人權保護、倫理、模型可解釋性、個人資料保護以及自動化決策，以下將依序簡述各文獻中對於這些風險類別的共通性探討：

（一）資料治理：OECD、EU、MIT、Engstrom、G7、UN 皆討論了 AI 風險管理與資料治理的框架，強調數據治理、透明度、監管等要素。

（二）人權保護：UN、EU、G7 強調 AI 應遵循人權原則，特別是在公平性、無歧視及保障弱勢群體權益方面。

（三）倫理：OECD、EU、G7 聚焦 AI 技術的倫理責任，討論了透明性、公正性以及技術對社會影響的平衡。

（四）模型可解釋性：EU、OECD、Engstrom 提到 AI 系統的可解釋性，強調在高風險領域中，模型需具備可解釋性以便監管和信任。

（五）個人資料：EU、UN、G7 討論了 AI 在處理個人資料時應遵守的隱私保護標準與法規（例如 GDPR）。

（六）自動化決策：OECD、G7、Engstrom 討論了自動化決策的挑戰，尤其是自動決策可能影響的透明性、責任問題，以及對人的影響。

下一章將說明本研析將上述風險類型與我國 AI 應用個案加以比對之結果。

第三章 分析方法

本研析主要研究設計有三，如表 7 研析設計摘要，為了使本研析所提之人工智慧規範架構接近公務人員實地使用情境，又符合國際對 AI 風險與治理的要求，本研析首先透過對世界上重要 AI 治理國家與地區之文獻蒐集，包含：歐盟、美國、澳洲、英國；與國際組織如聯合國、隱私遠見論壇（Future of Privacy Forum, FPF）、聯合國大學（United Nations University, UNU）等研究單位與學者之意見，由前述文獻可發現，歐盟以風險控管為主，其他國家也多強調 AI 的倫理議題；再者，本研析蒐集我國人工智慧應用之個案，包含數位國家·創新經濟發展方案（DIGI^{+}）、第 1~6 屆政府服務獎、臺北市政府，並與之將 AI 應用風險對應，除了解公部門 AI 應用現況外，並呈現 AI 治理上面臨眾多挑戰，最後，透過學者專家之訪談，產出我國 AI 規範架構之建議。

表 7 研析設計摘要

研析目的	資料蒐集方法
1、 人工智慧規範架構及國際相關規範之蒐集	文獻蒐集
2、 我國公部門現行人工智慧應用案例風險分析	人工智慧應用之個案分析
3、 我國人工智慧規範架構之研擬	文獻蒐集、學者專家訪談

資料來源：本研析自行整理

第一節 AI 應用分類

本研析透過文獻蒐集整理歐盟、澳洲、美國等國家或地區人工智慧規範架構及國際相關規範，同時，以次級資料方式整理我國政府機關導入 AI 的現況分析，以呼應 AI 架構的建立，針對我國政府機關應用 AI 服務之類型以及可能觸及之風險加以分析，首先說明本研析分類標準，以及 AI 風險類型。

一、 AI 應用分類

首先，基於美國以及歐盟相關研析（Engstrom *et al.*, 2020；Tangi *et al.*, 2022；de Sousa *et al.*, 2019；Misuraca *et al.*, 2020），均將政府 AI 服務應用分為五大類，包括執法（enforcement）、管制研析、分析與監控（regulatory research, analysis, and monitoring）、裁決 （adjudication）、公共服務和參與（public services and engagement）、內部管理（internal management）五類，本計畫亦採行這項分類，以下說明這五大分類內容，表 8 則呈現其細部分類與相關應用案例。

（一） 執法（enforcement）

執法部分包含：(1)智慧識別流程－該流程應用於辨識影像、聲音、音訊或其他可識別物理現象中的物體、人、地點、文字、情境與動作；(2)審計和日誌記錄管理的流程－著重於蒐集目標行為的軌跡及來源，提供影響特定操作流程或事件的證據，例如移動序列的畫面記錄；(3)預測性執法流程－專注於分析可用於預測和幫助預防未來潛在犯罪或錯誤解釋的大量資訊的過程；(4)支援檢查流程－適用於主管機關介入之前用於識別不當行為或錯誤的支援流程 （如檢查稅務狀況、註冊異常企業）。

（二） 管理分析與監控（regulatory, analysis, and monitoring）

包括：(1)資料分析流程－指資料和數據分析是檢查、轉換和建構資料過程，透過將資料轉換為可操作的知識，例如儀表板的應用；(2)監控政策執行狀況－指追蹤和評估政策實施的流程，用以確保政策得以制定、認可和實施；(3)預測和規劃則－基於預測模型的資源管理流程，以支援政府規劃政策。

（三） 裁決（adjudication）

主要用於就福利做出有關批准、驗證或撤銷利益決策的過程。

（四） 公共服務和參與（public services and engagement）

包括：(1)參與管理強調與公民和企業的聯繫，建立信任關係；(2)資料分享管理－涵蓋資料共享流程，考慮互通性和資料許可的問題；(3)服務整合－專注於整合管理多個服務供應商和資訊來源，以便為公民或其他組織提供特定化的客製化服務服；(4)務個性化－根據客戶（公民、企業、公務員）的需求提供個別化服務。

（五） 內部管理（internal management）

包括：(1)內部基本流程－為外部顧客創造價值的過程及對應（公民、企業）滿意度的影響；(2)內部支援流程－指為組織提供服務和資訊的流程;(3)內部管理流程－提供實現組織目標的管理、控制和決策支援工具的流程。

表 8 AI 在公部門應用類型

AI 應用類型	應用類型	描述	案例
執法（enforcement）	智慧識別流程	可辨識影像、影片、音訊或其他可偵測物理現象中的物體、人、地點、文字、情境和動作的過程。	● 證券交易委員會、醫療保險和醫療補助中心以及國稅局用以決定執行標的之順序 ● 海關和邊境保護局以及運輸安全管理局之人臉識別系統 ● 食品安全檢查局用以預先告知食品檢測事宜
	審計和日誌記錄管理	記錄蒐集目的和來源，提供隨時影響特定操作、程序、事件或設備的活動序列的書面證據。	
	預測性執法流程	分析可用於預測和幫助預防未來潛在犯罪/錯誤/誤解的大量資訊的過程。	
	支援檢查流程	在主管機關介入之前用於識別不當行為或錯誤的支援流程（如檢查稅務狀況、註冊異常企業）。	
管制分析與監控（regulatory, analysis, and monitoring）	資訊分析流程	資訊和數據分析是檢查、轉換和建模資訊過程。透過將資料轉換為可操作的知識（如儀表板）。	● 消費者金融保護局對消費者投訴的分析 ● 勞動統計局對工傷的敘述進行編碼 ● 食品藥品管理局對不良藥物事件之分析
	監控政策執行狀況	追蹤和評估政策實施的流程，以確保政策得到制定、認可和實施。	
	預測和規劃	基於預測模型的資源管理流程，以支援規劃。	
裁決（adjudication）	就福利做出決定	用於做出有關批准、驗證或撤銷利益決策流程。	● 社會安全局用於糾正裁決錯誤

AI 應用類型	應用類型	描述	案例
			● 專利商標局用於裁決專利和商標申請
公共服務和參與（public services and engagement）	參與管理	加強與公民和企業的聯繫，建立信任關係。	● 郵政署的自動車和手寫辨識工具 ● 住房及城市發展部及美國公民及移民服務局聊天機器人 ● 機構對法規制定相關意見的分析
	資料分享管理	資料共享流程，考慮互通性和資料許可。	
	服務整合	整合管理多個服務供應商和資訊來源，以便為公民或其他組織提供新的客製化的特定服務。	
	服務個性化	考慮客戶（公民/企業/公務員)的需求提供客製化服務。	
內部管理（internal management）	內部基本流程	為外部顧客創造價值的過程及對顧客（公民、企業）滿意度的影響。	● 衛生及公共服務部用於協助制定採購決策 ● 總務署用於確保聯邦招標的合法性 ● 國土安全部用於應對機構系統遭網路攻擊的工具
	內部支援流程	為組織運作提供服務和資訊的流程。	
	內部管理流程	提供實現組織目標所需的管理、控制和決策支援工具的流程。	

資料來源：Engstrom *et al.* (2020)

二、 案例可能觸及之風險面向

AI 系統在處理大量資料的過程中，可能出現如隱私洩漏、倫理爭議及模型可解釋性等問題。自動化決策過程中若缺乏適當的人工干預及判斷，可能導致不公平或錯誤的決策結果，甚至對弱勢群體產生負面影響。因此，對於 AI 的應用，需對潛在的風險進行全面評估和有效管理。本研析參考 Madan & Ashok（2023）指出的四大研析議題中，針對 AI 資料治理風險的五項分類，分別是（1）預測準確性 vs.透明及課責與玩弄系統：有（a）預測準確性高與結果的透明度和解釋間的權衡、（b）欠缺一般直觀、（c）AI 決策的課責、（d）基於 AI 的公共決策合理性、（e）能夠以

更高的透明度操縱系統；（2）預測準確 vs.歧視、偏見、公民權利間的權衡：有（a）使用敏感變數來提高預測能力與嵌入偏差和歧視的爭議、（b）針對弱勢族群邊緣化風險的可接受錯誤率、（c）數位落差、（d）從環境中獲得的負面教訓、（e）相關知識與情境化的人類知識的權衡；（3）資料可存取性 vs.安全性和隱私性：有（a）出於其他目的進用公民資料、（b）同意並提供資料乃能接受公共服務的爭議、（c）敏感資料的安全性持續受到威脅；（4）助推（nudging）vs.自主（autonomy）：（a）集體權利與個人自由的權衡、（b）利用 AI 實現政策目標所進行的國家監督和行為控制、（c）公民反對 AI 治理的權利、（d）個人化服務和分類過濾控制的爭議；（4）自動化（automation）vs.增強（augmentation）：有（a）重複性和低自由裁量權任務的自動化、（b）較高自由裁量任務的強化、（c）成本和效率動機對應新穎決策以及保護公民免受演算法傷害間的權衡、（d）對勞動市場的影響。同時透過文獻綜整，本研析提出案例可能觸及之六大風險面向，分別為資料治理、人權保護、倫理、模型可解釋性、個人資料、自動化決策。

（一） 資料治理（data governance）

資料治理是指確保資料在其生命週期內的準確性、完整性、安全性及一致性的過程。包括制定政策和程序來管理資料的蒐集、儲存、處理、共享與刪除，以避免資料被濫用或洩漏，並滿足相關法律或規範的要求（Murikah, 2024）。資料治理還關注數據權限管理（誰能存取什麼資料）以及資料品質（資料的正確性和一致性），以支持企業或組織的決策和風險管理。

（二） 人權保護（human rights）

高風險AI的引入對基本權利的合理可預見的影響，可能影響邊緣化人群或弱勢群體的特定傷害風險，像是偏見，因此Rodrigues（2020）認為人權保護是聚焦於避免技術應用對基本權利（如隱私、平等及非歧視）造成損害，並保障人們不會因自動化技術的使用而遭受歧視性或偏頗的對待。例如，在AI系統中，確保訓練資料庫不包含種族、性別或其他形式的偏見，並確保算法運行不會導致對某一群體的系統性不公平待遇。

（三） 倫理（ethics）

倫理風險關注技術的使用是否符合社會價值觀和道德規範。例如，透明度是指AI系統是否清晰地展示其如何得出結論，公平性則是確保技術不會造成不必要的偏見或對特定族群的負面影響（Camilleri, 2024）。而Wirtz et al.（2019）也提出公部門AI應用的挑戰，其中AI倫理的挑戰包括AI機器與人類價值判斷的兼容性、道德困境和AI歧視。此外，倫理還涉及科技應用對社會穩定的潛在影響，以及是否存在未經授權的濫用情況。

（四） 模型可解釋性（model interpretability）

模型可解釋性是指AI或機器學習模型的運作方式和決策過程是否能夠被人類理解和檢視。缺乏可解釋性的模型可能導致信任缺失，尤其是在高風險領域（如醫療或法律）。可解釋性風險包括模型的黑箱性質，會使得相關決策無法進行適當的監督和質疑，從而增加決策錯誤的可能性 （Camilleri, 2024）。

（五） 個人資料（personal data）

個人資料風險指技術在蒐集、處理和存儲敏感性個資時可能帶來的威脅，例如個資洩漏、濫用或未經同意的分享，包括身分資訊、財務資訊或健康資料等敏感性個資（Rodrigues, 2020）。確保個資匿名化處理以及嚴格遵守個人資料保護法是減少這一風險的關鍵。

（六） 自動化決策（automated decision-making, ADMS）

自動化決策風險是指完全由 AI 系統進行的決策可能導致的問題，尤其是在缺乏人工干預的情況下。例如，AI 可能基於偏差資料或演算法錯誤做出不公平的決策，而機關或相關人員卻無法挑戰或審核這些結果。另外，系統可能缺乏靈活性，無法處理非典型之案例，恐導致不適當的判斷及決策結果（Camilleri, 2024）。Engstrom & Haim（2023）亦指出政府引進 AI 最令人擔憂的是它推進到充滿自由裁量權的政策領域，AI 的機器學習和大數據可能未經人同意，搜查、扣押得出侵犯隱私的推論，且 AI 的不透明性讓政府可

能做出破壞民主的行為，也可能導致「自動化的焦慮」，例如美國的司法臉部辨識系統可能錯誤的將非裔男子與犯罪嫌疑人匹配，導致法律做出錯誤判決而監禁之。另外，呼應倫理構面，Loi et al.（2019）亦認為當公部門部署 ADMS 系統時，應以個人和社會對公部門的信任為最終標準，使用 ADMS 須對潛在倫理影響進行系統性評估，目標是維護防害、自主、正義與公平，並確保透明、控制和課責。

第二節 分析案例說明

本研析蒐集之我國人工智慧應用個案共計有以下：（1）數位國家·創新經濟發展方案（DIGI+）、（2）第 1~6 屆政府服務獎、（3）臺北市政府，因為臺北市為國內首先提出「臺北市政府使用人工智慧作業指引」之地方政府單位。故本研析針對 DIGI+、第 1~6 屆政府服務獎以及臺北市政府相關政府機關應用 AI 服務研析，加以整理分析，首先重點介紹這些案例，說明它們如何運用 AI 科技來改善政府機關之效能和公共服務品質。

一、 DIGI+案例

DIGI+是台灣推動數位經濟創新與智慧國家建設的長期發展方案，涵蓋發展基礎、創新經濟、智慧治理、包容社會，並以升級（plus/upgrade）為目標。該計畫旨在整合人工智慧、物聯網、大數據等智慧科技，加速產業創新與生活品質提升，展現台灣小而精、跨域整合的優勢，邁向智慧創新典範國度。根據「智慧國家方案 2023 年階段性推動成果報告」，DIGI+中涉及 AI 應用共計 40 個案例介紹，第一節首先依照各部會推動案例概要說明，詳細案例內容請參見附錄五。

（一）內政部共有 7 個案例，整理如下：

1. 警政署：「5G 智慧警察行動服務研析」及「強化虛實治安情資整合機制，提升員警破案效率」共 2 個研析。
2. 消防署：「AR 頭盔前進搜救現場第一線，結合 AI 提援效率」。
3. 移民署：「大數據分析支援國境安全決策」及「建立旅客安全警示－精準篩濾有疑旅客」共 2 個研析。

4. 內政部建研所：「建築工程技術精進創新與應用效能提升」。
5. 內政部與衛福部合作：「內政大數據加值應用計劃」。

（二）衛生福利部共有 3 個案例，整理如下：

1. 衛生福利部：「衛福業務數位轉型服務躍升」。
2. 中央健康保險署：「健保大數據數位應用」。
3. 疾病管制署的：「智慧防疫空間及空氣品質數據分析」。

（三）經濟部共有 4 個案例，整理如下：

1. 商業發展署：「公司登記文件影像自動分類系統」。
2. 資訊處：「建置台水總售水量預測模型，輔助用水預估」。

（四）地質調查及礦業管理中心共 2 個計劃，整理如下：

1. 「建立大臺北地區、臺南地區地質大數據，提供防災應用」。
2. 「建置中部地區山崩潛感模型，支援防災決策」。

（五）環境部共有 3 個案例，整理如下：

1. 環境管理署：共推出 2 個計劃，分別是「運用車牌辨識與 AI 技術，實現智慧勾稽異常清運行為」及「運用新興科技工具與遙測技術，提升廢棄物棄置場址監控及執法效能計劃」
2. 化學物質管理署：「食品安全高風險異常廠商偵測模型」。

（六）財政部共 2 個案例，整理如下：

1. 財政資訊中心發展：「智能稅務服務系統」。
2. 關務署推出的「資料檢索系統」。

（七）交通部中央氣象署共 5 個案例，整理如下：

1. 「發展智慧化地震預警系統」。
2. 「氣象領域維運與技術發展及智慧海象環境災防服務」。
3. 「建構無縫隙氣象服務價值鏈－橋接農、漁、光電領域研析」。
4. 「以 AI 技術估計颱風強度」。
5. 「應用人工智慧開發數值模式預報加值產品」。

（八）法務部共有 2 個案例，整理如下：

1. 「調查局鑑識科學大樓遷置與科學偵查設備精進中程研析」。
2. 「新世紀檢察 AI 智慧輔助系統建置案」。

（九）原住民族委員會共有 3 個案例，整理如下：

1. 「原住民族高等教育人才培育決策」。
2. 「原住民族部落長者長照服務決策資訊」。
3. 「社會福利服務資源分布資訊」

（十）國家發展委員會共有 3 個案例，整理如下：

1. 社會發展處：共有 2 個計劃，分別為「以資料科學為基礎的社會政策治理機制」及「社會政策循證決策治理機制」。
2. 管制考核處：「決策支援模組」。

（十一）其他機關，共 8 個案例：

1. 僑務委員會：推出「評點制核可人次統計資料分析」之計畫。
2. 國軍退除役官兵輔導委員會：「榮家無線網路佈建研析」。
3. 國立故宮博物館：「故宮藝術資料 AI 技術應用」計畫。
4. 文化部：「文化數據智能分析與決策輔助研析」。
5. 海洋委員會海巡署偵防分署：建置「港區及聯外道路車牌辨識系統、船舶軌跡航行監控分析及新世代海巡偵防業務整合系統」。
6. 核能安全委員會：推出「國際輻防技術規範與精進量測技術能力」計畫。
7. 行政院人事行政總處：「完備有效之循證決策模式，提升政府服務及施政決策之精準度」計畫。
8. 勞動部勞動及職業安全衛生研究所：針對大眾運輸業，建立了「職業駕駛不安全行為預警系統」。

二、 第 1~6 屆政府服務獎案例[45]

政府服務獎係指獎勵各機關（構）扣合施政主軸，強化機關服務作為與政府施政連結性，鼓勵機關以人為本，提出善用數位科技、公私協力且具多元包容性之創新服務，兼顧經濟、環境及社會永續發展，進而擴散優質服務效益，樹立標竿學習楷模。本研析選自第 1~6 屆政府服務獎之案例，由於政府服務獎案例涉及中央與地方單位較多，這部分則針對案例類型簡單歸類，分為智慧警政消防、智慧交通、智慧

[45] 服務獎 2025 年為第八屆，本研析分析時為第六屆。

環境、智慧城市、智慧行政、智慧教育及智慧觀光七大領域，共計 32 個案例，案例內容請參見附錄六。

（一）智慧警政消防：在智慧警政消防領域中，共有 6 個案例，整理如下：

1. 基隆市消防局：「智慧消防 2.0」。
2. 內政部警政署刑事警察局：「警察電商聯盟--終詐之戰」。
3. 新北市消防局：「全災行智慧化指揮監控平台」。
4. 新北市政府警察局：「智慧城市 安全新北」。
5. 彰化縣消防局：「AI 進行防災影像辨識預警與救災任務語音紀錄」。
6. 高雄市政府毒品防制局：「ICARES AI 科技輔導」。

（二）智慧交通：在智慧交通領域中，共有 5 個案例，整理如下：

1. 嘉義市政府警察局：「諸羅城縱橫攻略」。
2. 臺中市政府交通局：「五心級智慧交通管理系統」。
3. 臺南市政府交通局：「AI 智慧車流辨識及智慧號誌管理策略」。
4. 臺北捷運：「智慧客服」。
5. 臺北市政府交通局：「智慧化交通量調查分析系統」。

（三）智慧環境：在智慧環境領域中，共有 5 個案例，整理如下：

1. 環保署：「空品智慧 GO」和「AIOT 科技神助功」共 2 個研析。
2. 環保署毒物及化學物質局：「從目測到遙測，從太空看台灣」。
3. 臺中市環保局：「AI 智慧水線抑制河川揚塵」。
4. 雲林縣環保局：「智能科技」。

（四）智慧城市：在智慧城市領域中，共有 5 個案例，整理如下：

1. 新北市政府養護工程處：「iRoad 新北市智慧道路管理中心」。
2. 臺南市政府智慧發展中心：「城市數據交換研析，打造未來城市自主感知」。
3. 臺中市政府建設局：「自來水管科技檢漏系統」。

4. 臺北市政府工務局：「污水管 AI 檢視」。
5. 臺南市政府水利局：「及時水情一點通：智慧防災推動研析」。

（五）智慧醫療：在智慧醫療領域中，共有 9 個案例，整理如下：

1. 臺南市衛生局：「科技防疫 現代蚊清」。
2. 健保署：「健保大數據跨域合作數位科技防疫新典範」。
3. 健保署南區業務組：「天涯若比鄰醫定守護您」研析。
4. 臺北榮民總醫院：「藥安心：以 AI 創新居家用藥整合服務，守護民眾用藥安全」、「AI 大師來精算，智慧照護零時差」共 2 個研析。
5. 臺東縣衛生局：「臺東南迴原鄉衛生所躍升醫學中心及服務品質研析」。
6. 高雄市立凱旋醫院：「AI 照護心體驗，保命防跌新神氣」。
7. 國民健康署結合國家衛生研究院開發：「 癌症登記 AI 輔助程式、肺癌影像輔助程式」。
8. 高雄市毒品防治局：「ICARES AI 科技輔導～走出藥癮迷途」。

（六）智慧行政：在智慧行政領域中，共 1 個案例：

1. 高雄市政府地政局：「BEST 價+給你掌握數據，價值永續」

（七）智慧教育：智慧教育領域中，共 1 個案例：

1. 臺南市政府教育局：「生成式 AI 輔助學習平台」

（八）智慧觀光：在智慧觀光領域中，共 1 個案例：

1. 交通部觀光署：「AI 翻譯櫃台」。

三、 臺北市政府案例

根據臺北市政府資訊局發布之市府新聞稿中得知，由於臺北市政府為全國第一個發出 AI 指引之城市，故本研析選自目前臺北市政府各機關單位之 AI 應用研析，共計 28 個案例，詳細案例內容請參見附錄七。

（一）交通管制工程處共 2 個案例，整理如下：

1. 「智慧號誌」。
2. 「臺北市智慧影像事件偵測建置案」。

（二）臺北市立聯合醫院共有 8 個案例，整理如下：

1. 「AI 輔助診斷糖尿病視網膜病變」。
2. 「AI 輔助診斷糖尿病 2.0 視網膜病變」。
3. 「血壓血氧上傳系統」。
4. 「LDCT（低劑量肺部電腦斷層掃描）」。
5. 「AI 影像電腦輔助判讀系統」。
6. 「雲端醫院」。
7. 「AI 輔助診斷糖尿病視網膜病變」。
8. 「AI 輔助診斷糖尿病 2.0 視網膜病變」。

（三）臺北翡翠水庫管理局共 3 個案例，整理如下：

1. 「翡翠水庫集水區環境影像變異監測」。
2. 「翡翠水庫經營管理智慧決策系統」。
3. 「翡翠水庫大壩安全監測評析」。

（四）臺北自來水事業處共有 3 個案例，整理如下：

1. 「自來水管線汰換規劃」。
2. 「管網水理模型校正」。
3. 「智能客服系統文字查詢及申辦」。

（五）其他機關之研析，共 12 個案例，整理如下：

1. 臺北市政府資訊處：「智慧城市 」。
2. 臺北市政府教育局「臺北酷課雲「酷 AI（CooCAI）」。
3. 臺北市政府消防局：「建置災害應變雲端協作平台」。
4. 臺北市工務局大地工程處負責的「AI 協審水土保持研析」。
5. 臺北市都發局的「地籍套繪都市研析，使用分區圖查詢便民服務」。
6. 臺北市建成地政事務所：「登記教主新服務—繼承登記法服諮詢」。
7. 臺北市產業發展局：「AI 人工智能線上客服」。

8. 臺北市政府警察局刑事鑑識中心則推出「毒品資料圖像快搜計劃」。
9. 臺北市體育局：「臺北市競技運動訓練暨科學中心研析」。
10. 臺北市政府人事處：「人事法規、人事系統操作及同仁人事業務相關權益諮詢事項」。
11. 臺北市大眾捷運股份有限公司：「輪椅旅客自動叫梯服務」。
12. 臺北市政府地政局：「網搜防偷跑─預售建案銷售廣告檢索」。

四、 分類分析方法說明

本研析透過人工智慧協助，以及專家判斷協助案例分析，以下為兩種分析方法說明。

（一） 人工智慧分析

以 ChatGPT 協助進行案例分類及構面勾選。ChatGPT 首先參考文獻，以 Engstrom *et al.*（2020）所提出之五大分類為依據進行為案例分類依據，而後判斷個案是否涉及資料治理、人權保護、倫理、模型可解釋性、個人資料、自動化決策風險。操作過程如下：

（二） 輸入參考文獻

首先將上述 Engstrom *et al.*（2020）所提出之五大分類及其應用類型作為指令提供給 ChatGPT。

（三） 交叉構面分類

接著請 ChatGPT 依照資料治理、人權保護、倫理、模型可解釋性、個人資料、自動化決策風險這幾個構面的定義，以交叉表格的方式呈現，如果案例有涉及那幾個構面，請打 V。系統分析過程會根據案例內容與輸入的參考資料進行推論，並提供每個構面勾選的原因。

（四） 專家分析

本研析協同主持人就 ChatGPT 分析結果後，逐一判斷，其中特別關注人工智慧分析可能過於寬鬆或案例可能同時

涉及的其他分類。對於 AI 結果的分類、構面的勾選，研析者詳讀 ChatGPT 所給出之理由，檢驗其分類與勾選是否合理且充分反映案例內容並予以修正。

第三節 專家訪談

本研析在盤點完機關單位 AI 應用的風險後，會將結果提供表 9 的專家學者審閱，同時蒐集專家對於政府 AI 治理架構的意見，專家訪談的目的是為了解政府單位在使用 AI 技術時，流程或技術方面有哪些困難，期望透過訪談，對機關所導入的 AI 技術，給予風險治理上的改善建議。專家個別訪談約 60 至 120 分鐘不等，專家訪談會在取得受訪者同意後，以錄音、筆記等方式記錄訪談內容，並於整理訪談摘要。

表 9 個別訪談名單

服務單位	職稱	姓名	訪談日期
內政部統計處	處長	饒志堅	2025/3/7
資安院	院長	何全德	2025/1/17
環境部	監測資訊司資訊發展應用科科長	黃崇富	2025/2/26
臺北市政府資訊局	局長	趙式隆	2025/3/10
安永管理顧問股份有限公司	總經理	萬幼筠	2025/3/23

資料來源：本研析自行整理

第四章 研究結果

第一節 我國公部門 AI 應用風險分類結果

透過第三章所述的人機協作分析方法，本研析得出之完整結果詳見附錄二至附錄四，本節舉兩個例子作為分析結果說明。

一、 建立原住民族部落長者長照服務決策資訊【原住民族委員會】

（一）人工智慧初步分類

AI 初步分類為「管制分析與監控」中的「預測及規劃」，人工智慧給出之原因為該案例依據原住民健康站所提供的資料，建構長照決策支援系統，其核心目的是通過數據分析支持政策決策，提供長期規劃建議，並優化未來長照機構的設置與資源分配。涉及之構面包括：

1. 資料治理：需管理健康站數據的蒐集與應用，以確保數據準確性與安全性。
2. 倫理：需考量是否充分尊重原住民族的文化特性及其獨特需求。
3. 模型可解釋性：系統應能解釋其建議的邏輯，以便相關部門信任分析結果。
4. 個人資料：若牽涉到健康站中敏感性個資的應用，需注意個資隱私問題。
5. 自動化決策風險：需避免依賴系統分析結果做出錯誤的決策，影響政策實施成效。

本案例未涉及之構面包括：

人權保護：案例影響範疇主要在政策優化與服務分配，若系統未能兼顧文化特性，影響的僅是政策成效，而非對個體權利的直接衝擊。

（二）專家分析

根據專家判斷，此案例在分類中，除了涉及管制分析與監控中的預測及規劃外，尚涉及「內部管理」的「內部管理流程」，因為該系統主要是透過資料分析來協助原住民族委員會進行內部政策規劃與決策，並未直接對外部民眾產生執法或監控的行為，符合內部管理流程中所述的「決策支援工具」。

（三）人工智慧修正判斷

而後，請 AI 選擇此個案較適合管制分析與監控又或是內部管理，AI 認為此案確實除了涉及管制分析與監控中的預測及規劃外，亦涉及內部管理中的內部管理流程，而在兩者中，本案例較適合內部管理，係因本案例之主要功能是由於此案例是透過健康站資料分析支援政府內部的資源規劃與分配，屬於行政內部優化與決策支持系統，而非直接進行對外的監控或執法行為，不是直接的對外監控或執法。

（四）最終結果

主要分類為「內部管理」中的「內部管理流程」，但同時涉及管制分析與監控中的預測及規劃，而涉及之構面包括資料治理、倫理、模型可解釋性、個人資料、自動化決策風險。

二、臺中榮民總醫院「AI 大師來精算，智慧照護零時差」

（一）人工智慧初步分類為「管制分析與監控」，因為系統通過 AI 推論病患急性呼吸窘迫的風險，顯示風險值並主動提供治療建議，體現了資料分析與即時監控的結合，因此屬於管制分析與監控的應用。而涉及之構面包括：

1. 資料治理：需要管理病患資料的蒐集、存儲、處理及共享，確保資料正確性、安全性及合法使用。
2. 人權保護：AI 推論結果可能影響病患權益，需保障其知情權、選擇權及對自動化決策的申訴權。
3. 倫理：應確保 AI 的推論和治療建議符合醫療倫理，避免對病患造成不公平或有害的影響。
4. 模型可解釋性：需確保 AI 模型的風險預測和治療建議是透明且可理解的，便於醫療人員信任並採取行動。

5. 個人資料：病患的重症醫療資料屬於敏感性個資，需確保隱私保護和合法處理。
6. 自動化決策風險：AI 治療建議可能存在誤判風險，需制定機制以防止自動化錯誤對病患健康造成負面影響。

（二）專家分析

詳讀 AI 所給之答案，認為判斷結果合理。

（三）最終結果

分類為「管制分析與監控」，而涉及之構面包括資料治理、人權保護、倫理、模型可解釋性、個人資料、自動化決策風險。

第二節　訪談結果

本研析各專家訪談時間如前章所述，本節呈現各訪談的重要結果與發現。

首先，在內政部統計處饒處長的訪談，饒處長說明內政部統計處已經對同仁常用之 AI 工具與可能風險進行對應與盤點，如表 10，據以發展內部指引與應用。表 10 將 AI 應用分為 4 個「大類」，每個大類下包含多個「中類」，進而細分為具體的「AI 應用項目」，並詳細說明了每個應用項目的「應用內容」。此外，表格最右側的兩欄「項目代碼」和「潛在風險(a) 代碼及因應對策」則標註每個 AI 應用可能面臨的風險及其對應的策略，以發展負責任的 AI 應用。「大類」 的四個核心領域分別為：(1)收集資料、清整合併：這是 AI 應用的基礎，強調數據的準備工作，聚焦在 AI 影像及語音辨識，及 AI 文件清整與管理。(2)分析資料、視覺展示：側重於從數據中提取洞察並以易於理解的方式呈現，聚焦在 AI 生成與視覺化，及 AI 數據分析與分類。(3)應用資料、循證決策：涉及 AI 如何利用分析結果來輔助或執行決策，聚焦在 AI 作業流程自動化，及 AI 預測與風險管理。(4)提供資料、公共服務：AI 直接面向公眾的服務場景的應用，聚焦在 AI 客服與客製化。

表 10　內政部 AI 應用分類與風險對策

大類	中類	AI應用項目	應用內涵	項目代碼	潛在風險(a)及因應對策
收集資料、清整合併	AI影像及語音辨識	語音智慧辨識	識別語音內容並轉換成文字	A1	AI安全性(揭露隱私)↓恪遵個資保護
		影像智慧辨識	分析影像內容以標註關鍵資訊	A2	
		資產條碼辨識	辨識與管理資產條碼資訊	A3	
		檔案文件辨識	自動辨識與分類檔案文件內容	A4	
	AI文件清整與管理	文件資料清整(如統一格式)	統一與整理文件格式	B1	
		文件串檔合併排序	合併並排序多個文件	B2	
		文件合理性檢誤除錯	檢查與修正文件內容錯誤	B3	
		製作會議記錄與摘要	轉錄並整理會議記錄	B4	
分析資料、視覺展示	AI生成與視覺化	數據視覺化呈現	將數據轉化為圖像化呈現	C1	AI正確性(錯誤資訊)↓落實透明公開
		圖表自動生成	自動產生數據圖表以利分析	C2	
		簡報自動生成	自動生成簡報內容與排版	C3	
		教材智慧製作	自動製作教材與學習內容	C4	
		公文、新聞稿產製	根據內容自動產生公文與新聞稿	C5	
	AI數據分析與分類	資料分類貼標(如公文分文)	自動分類與標籤化資料	D1	
		統計數據分析	分析統計數據以提供決策依據	D2	
		資源最佳配置	最佳化資源分配與使用	D3	
		成效評估分析	評估計畫與方案的成效	D4	
		預算執行分析	監測並分析預算執行狀況	D5	
		人流與使用分析	監測人流變化與資源利用	D6	
		資產管理評估	資產使用狀態評估	D7	
		財務風險評估	財務狀況與風險評估	D8	
應用資料、循證決策	AI作業流程自動化	業務流程自動化	自動化業務處理流程	E1	AI公平性(偏見歧視)↓加強人類覆核
		審查流程優化	優化審查與核准程序	E2	
		管理流程改善	提升管理流程的效率	E3	
		資源調度自動化	自動分配與調度資源	E4	
		服務流程優化	提升服務流程的體驗	E5	
	AI預測與風險管理	趨勢預測模型	預測未來趨勢變化	F1	
		安全管理預警	監測與預測安全風險	F2	
		犯罪預防偵測	分析與預測犯罪活動趨勢	F3	
		設施智慧監控	智慧監控重要設施與區域	F4	
		災害預警系統	提前偵測與預警災害風險	F5	
		風險評估預測	評估未來風險因素	F6	
		資源需求預測	預測未來資源需求變化	F7	
		維修需求預測	預測設備維修時機	F8	
		虛擬場景模擬	模擬虛擬環境進行決策分析	F9	
提供資料、公共服務	AI客服與客製化	24/7智慧客服	提供全天候的智能客服服務	G1	AI含容性(擴大落差)↓提供多元方案
		法規智慧問答	自動回應與解析法規問題	G2	
		案件諮詢服務	提供案件處理建議	G3	
		輿情回應處理	監測並回應社會輿情	G4	
		客製化服務提供	根據用戶需求提供客製化服務	G5	

說明(a)：AI潛在風險尚包括侵害智財權、不易問責、不利環保、權力集中等。

資料來源：內政部統計處訪談資料（未公開）

在臺北市資訊局長的訪談，局長對於本研析提出的 AI 風險的交互性提出建議 （見圖 5），簡單來說，各 AI 風險構面之間並非獨立存在，而是緊密交織，共同影響 AI 技術的應用風險及治理策略。

首先，資料治理是 AI 風險治理的基礎，因為所有 AI 系統都依賴數據來進行訓練和推論。資料治理確保數據的收集、處理、儲存和使用均符合合法性和透明性的要求，並強調數據品質和偏差控制。良好的資料治理不僅能減少數據偏差，也能提升模型訓練的準確性，避免在自動化決策中產生系統性偏誤，這也與模型可解釋性密切相關。

其次，模型可解釋性在風險治理中扮演著關鍵角色，因為 AI 技術常涉及複雜的深度學習和黑箱模型，難以解釋系統做出決策的邏輯。高可解釋性的模型能提升透明性，使人們理解自動化決策的依據，進而增加信任感。這對於涉及個人資料的人權保護尤為重要，因為當 AI 系統使用個人數據來進行決策時，如果無法解釋決策過程，將會引發隱私風險和公民權利受損的問題。當自動化決策導致某人被拒絕貸款或遭受歧視待遇時，理解模型的判斷邏輯尤為必要，這也直接牽涉倫理構面的挑戰。

再來，個人資料人權保護是 AI 應用不可忽視的一環，特別是在自動化決策中，保護數據隱私和尊重數據主體權利，是所有合法性和道德性考量的核心。自動化決策系統如果未能保護個人資料，可能違反如《一般資料保護規範》（GDPR）等法規，進而引發法律責任和社會批判。因此，自動化決策必須在其模型構建和應用過程中，平衡數據使用與隱私保護，確保系統的合法性和倫理合規。

倫理構面在這些風險治理議題中具有指導性意涵，因為 AI 技術在日常生活中逐漸普及，如何確保系統運作符合道德原則，如公平、透明和課責，變得至關重要。當 AI 技術用於自動化決策，尤其是影響個人福祉和權益時，系統若缺乏倫理審查和社會監督，將可能造成巨大的社會影響。因此，從設計到部署，AI 系統必須考量社會價值觀和倫理規範，避免對特定群體造成不公平對待。AI 風險治理是一個多構面整合的過程，唯有在各個構面之間建立平衡和互動機制，才能確保 AI 技術應用的可靠性、公平性及合規性。

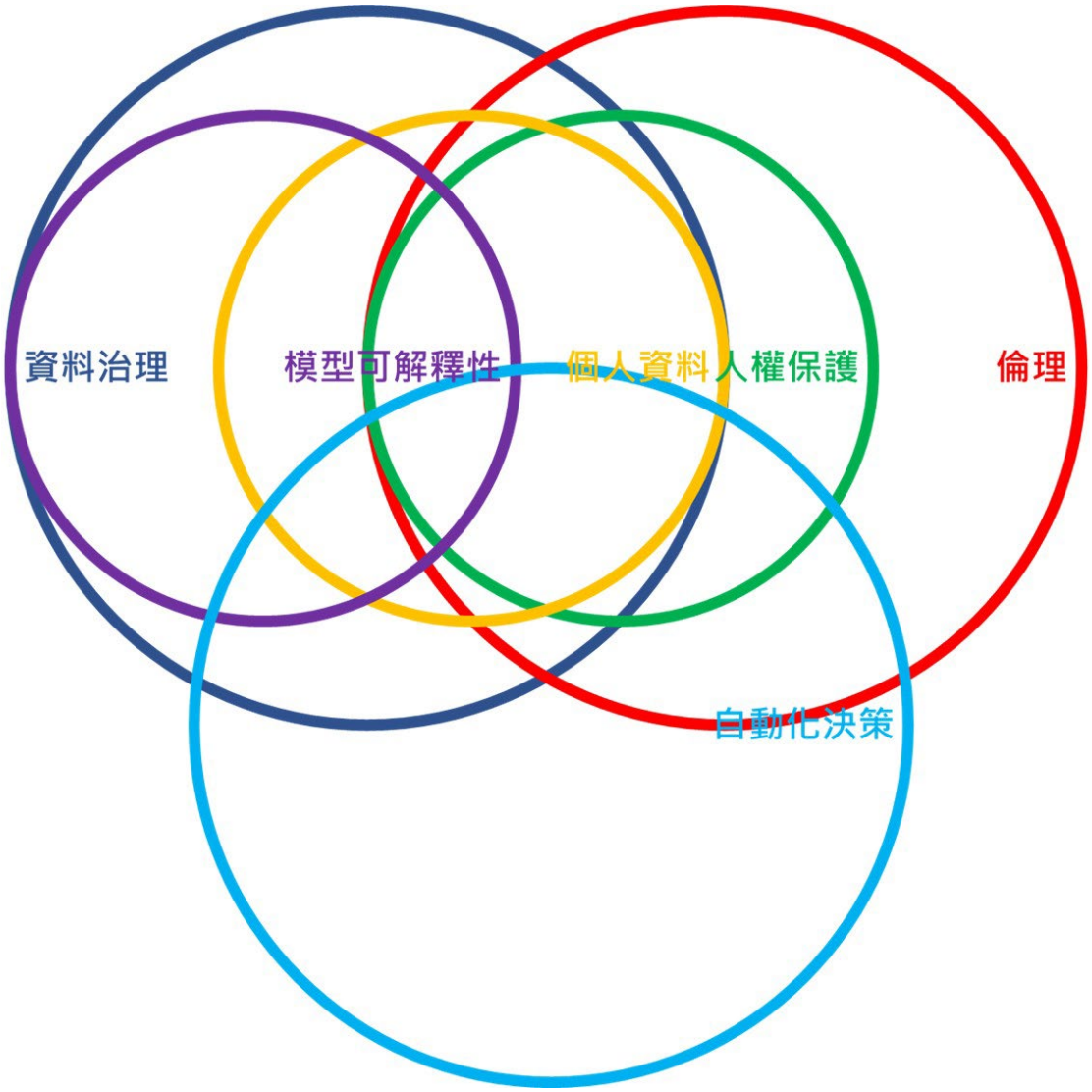

圖 5　AI 風險的交互性

資料來源：本研析自行繪製

資安院何全德院長在訪談中提出可以參考國家發展委員會以前訂定的計劃管理規範，並針對我國 AI 風險治理的政策定位與發展方向提出多項見解。何院長首先指出，全球在人工智慧治理策略上的取徑並不一致。例如，歐盟採行高度分級控管制度，強調對 AI 風險的分層管理，而美國則偏好自由市場導向，容許較大的技術發展彈性。相較之下，台灣目前採取「分類」而非「分級」制度，意在避免歐盟模式所帶來的高度管制風險。在此背景下，美國 NIST（國家標準與技術研究院）以及 MIT 等機構所提出的多層次分類架構，已成為我國政策參考的重要依據。

進一步談到風險治理架構，何院長說明 MIT 提出的風險分類方式，主要是依據 AI 風險的來源進行劃分，包括來自人為因素、系統本身或第三方等。此外，這些風險還可以根據技術應用是否出於「有意」或「無意」加以細分。AI 風險亦可根據資料處理的不同階段進行分類，涵蓋前期的資料安全、中期的模型安全，以及後期的應用安全。台灣目前參考七大類、二十三個次分類的系統，與 NIST 的架構相近，並作為未來治理指引的重要基礎。

在政策應用方面，何院長指出，政府推動 AI 風險管理應以「風險識別—風險分類—風險緩解（mitigation）」為核心流程。目前已有部分政府部門委託研究團隊，將既有分類指引與國內 40 個 AI 專案進行對應分析，以辨識可能潛藏的風險。未來，行政體系不排除成立倫理委員會，或指定具領導地位的 AI 企業參與風險治理工作，以提高治理的整體效能與回應社會關注。

至於法制與行政機制的現況，何院長指出目前台灣 AI 風險治理仍以行政指引為主，尚未進入正式立法階段。建議未來可仿照《電子簽章法》的制定方式，優先針對政府機關進行規範，再依實際情況擴展至民間企業。需要注意的是，政府部門與企業在風險治理的需求上有所差異：前者須更注重正確性、公平性與透明性，而後者則傾向於維持創新能力與制度彈性。關於 AI 的禁止性行為，如暴力內容生成、散布假訊息或侵害智慧財產權等，未來亦可能考慮透過刑法予以明文規範。

針對國際趨勢的觀察，他表示，歐盟目前採用四級風險分級制度，而韓國則聚焦於高風險 AI，採取「抓大放小」的管理策略。美國 NIST 則強調應聚焦在風險估計與風險緩解，避免因過度干預而妨礙創新發展。目前歐洲、美國、英國以及 OECD 等國際組織皆已著手建立 AI 風險治理政策，其涵蓋範圍包括生成內容風險、環境影響、資訊安全、人權衝擊等面向，顯示 AI 風險管理已逐漸邁向多元化與整合性。

最後，何院長提出對台灣未來推動方向的建議。我國可考慮成立跨部會的倫理委員會或政策協調平台，以統整並一致化各部會對 AI 治理的標準與措施。此外，也應善用既有的政府風險管理方法論，將 AI 議題納入並加以擴充，作為政策設計的延伸。為使政策落實更具秩序與效果，未來應釐清 AI 治理的法律適用範圍，並訂定政策推動的先後次序，逐步建立起健全且具回應性的治理架構。

綜整以上訪談結果，本研析另整理受訪對象與研究團隊的幾個觀察，在政府部門間，AI 的開發與應用可能因資源、技術能力、數據基礎及政策環境的不同，形成顯著的數位落差。這些落差可能導致某些部門在 AI 技術應用上處於弱勢，進而影響政府整體數位轉型的推動，本研析認為至少有以下六個面向。

一、 算力落差（Computational Power Gap）：各部門之間在計算資源（如 GPU 伺服器、雲端計算服務）上的差異，導致 AI 模型訓練和部署能力存在巨大落差。

二、 資料落差（Data Resource Gap）：各部門擁有的資料量和資料品質不佳，影響 AI 模型的訓練和預測準確性，將導致模型過度擬合或無法普遍應用，使部門間的數位服務水準落差加劇。

三、 AI 模型落差（AI Model Gap）：同部門對 AI 模型的掌握和應用能力存在差異，特別是在使用 AI 模型（GPT、GEMINI 等，以及不同大小，13B 或 70B 大模型）上的落差，模型落後使得某些部門難以實現資料自動化處理，造成服務品質低落，且在資料驅動決策處於劣勢。

四、 人才落差（Talent Gap）：科技相關部門可能有資料科學家和 AI 工程師，而地方政府單位多由一般資訊人員或外包團隊處理 AI 技術，人才不足不僅導致技術應用受限，還可能造成技術推動緩慢及跨部門協作困難。

五、 法規與政策落差（Regulatory Gap）：法規不統一或政策衝突，可能導致 AI 應用困難，甚至觸法風險增加。

六、 預算落差（Budget Gap）：經費不足使得 AI 應用發展受限，特別是長期維護和技術升級上更顯困難。

最後，受訪者也表示在推動 AI 治理的實務上困難，如：假 AI 瓜分政府資源與真 AI 逃避政府監管；政府資源重複投入；如何以「技術普惠條款」，確保在關鍵技術上的社會共享，同時避免大公司壟斷技術發展。

第五章 研究結論與建議

第一節 研究結論

隨著人工智慧技術的快速發展，全球各國都在加速制定相關法規與行政命令，以應對 AI 所帶來的資安防護、民眾隱私保護、數位權利、公部門數位轉型以及數位涵容社會等議題 。本研究研析符合我國國情的 AI 規範架構，作為政府運用 AI 時的高層次、整體性指導架構，以供各行政院所屬單位參考。在國際間，歐盟、美國、英國、澳洲以及聯合國等國際組織皆已針對 AI 治理提出相關規範與原則。其中，歐洲理事會（CoE）提供締約國一般性義務和原則，後由各國國內法化補充。《歐盟 AIA》是全球第一部規範人工智慧的法律，其監管分級以用途可能涉及的風險層級為理論基礎 。美國拜登政府初期採取嚴格監管模式，但川普政府則傾向開放，並強調 AI 應不受阻礙地推動政府效率與經濟成長。英國曾嘗試透過法律規範 AI 但未能成功，轉而提出 5 項關鍵人工智慧監管原則 。澳洲則主要提供自願遵循的指引或原則，並針對高風險 AI 使用提出強制性防護措施建議。我國政府 AI 應用類型與場景的調查分析顯示，公部門 AI 應用可分為執行、管制分析與監控、裁決、公共服務和參與，以及內部管理五大類 。在這些應用中，常見的 AI 技術包含影像辨識、機器學習、大數據分析、動態影像分析辨識演算法和自然語言處理 (NLP) 等 。同時，相關應用也伴隨著資料治理、人權保護、倫理、模型可解釋性、個人資料以及自動化決策等多重風險。

本研析參考聯合國大學（United Nations University）七個階層的《人工智慧治理框架》（Framework for the Governance of Artificial Intelligence），提出五個層級的人工智慧治理架構與措施，彙整我國目前堆動 AI 的作為，整理如表 11，特別說明：表 11 以數位發展部可能業管項目出發，而在備註欄為旁及其他可能單位之作法，但在公務同仁行為以及人工智慧價值則不區分單位。

表 11　我國人工智慧治理架構與措施

治理層級	目前措施	備註
法律	1. AI 基本法[46]：AI 發展的上位法規，規範 AI 倫理、安全、資料治理等。 2. 促進資料創新利用發展條例（草案）	1. 其他現有相關法律：如個人資料保護法（規範 AI 應用中的個資蒐集、處理、利用）、著作權法（AI 生成內容的著作權歸屬）、國家資通安全發展方案（AI 資安議題）等。 2. 部分地方政府可能依據中央法令，制定與 AI 相關的在地規範，例如「臺北市政府使用人工智慧作業指引」
標準	AI 評測（數產署）	ISO27001 ISO-iec-42001
政策與計畫	1. 公部門人工智慧應用參考手冊草案 2. AI 領航推動計畫	1. 臺灣 AI 行動計畫（第一期、第二期） 2. 各部會 AI 推動政策
公務同仁行為	AI 倫理指引與公務員 AI 培訓：提升公務員對 AI 工具的認知與應用能力，優化行政效率。	
人工智慧價值	1. AI 發展的四大核心價值：信任、公平、可負責性、透明度，強調 AI 應以增進人類福祉為目標，避免對個人權利、社會公平造成負面影響。 2. 同時確保 AI 系統的設計、開發、部署和使用，都符合倫理規範和社會期望，也就是負責任 AI（Responsible AI）。	

資料來源：本研析自行整理

[46] 研議階段

第二節　研究建議

隨著人工智慧（AI）技術快速發展，各國政府積極制定相關規範以保障人權、提升政府效能並促進數位轉型。研究依據 PDCA 循環管理理念如圖 6），建構出一套四大構面、可持續優化的人工智慧（AI）治理架構，同時發展出八個要素：覺知、素養、合規、部署、訓練、評估、迴圈推動。作為我國公部門因應 AI 發展挑戰與創新需求的制度性參考。以下分為四大治理構面說明各項政策建議：

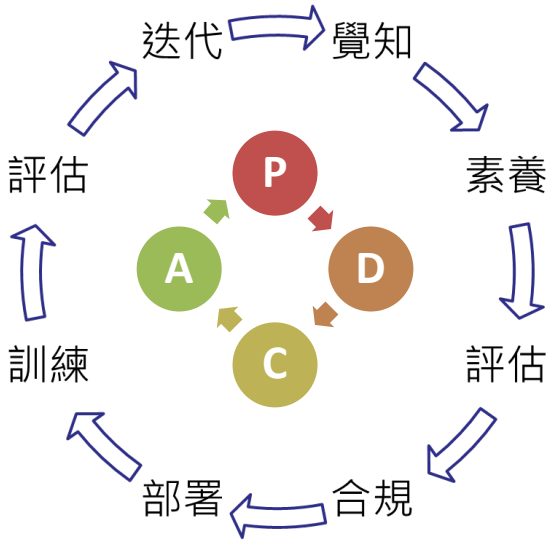

圖 6　公部門導入 AI 的 PDCA 循環

資料來源：本研析自行繪製

一、 Plan（規劃）：建構制度基礎與預先風險識別

參考澳洲政府，於政策規劃階段已建立明確的治理原則與法律政策基礎。自 2019 年起推行的《人工智慧倫理原則》，為後續政策發展奠定價值基礎，聚焦於公平性、透明性、可課責性等倫理目標。此外，2024 年所發布的《確保公部門使用 AI 的國家架構》，提供了五大治理支柱（治理、資料、風險、標準、採購），清楚界定政府機關在制度建設與責任分工上的規劃方向，強調建立對 AI 治理的基本原則與使用、風險意識等，以利日後各項施策具有明確方向。具體建議如下：

（一）　確立 AI 治理架構與原則：這是所有後續工作的基石。需明確 AI 的治理目標、範圍，並制定一套核心原則，如公平性、透明度、可解釋性、隱私保護、安全性等，作為指導公部門 AI 開發與應

用的最高指導方針。

（二） 定義責任與課責機制：在 AI 應用前，需明確各方利害關係人（如開發者、部署者、使用者）的權責，並建立相應的課責機制，確保當 AI 系統產生問題或偏誤時，能夠明確追溯責任歸屬。

（三） 重視資料治理規劃：資料是 AI 的燃料，其品質與合規性直接影響 AI 的效能與可靠性。因此，在計畫階段必須詳細規劃資料的收集、儲存、處理、共享及使用的規範，確保資料的安全性、隱私性與完整性。

（四） 規劃透明度與可解釋性措施：為了建立公眾信任，應在設計階段就考量如何提升 AI 決策過程的透明度與可解釋性，讓人們能夠理解 AI 的運作邏輯及決策依據。

（五） 規劃人才培育與意識提升：AI 治理不僅是技術問題，更是組織文化與人才素養的問題。需規劃 AI 相關專業人才的培訓，並提升公務人員及大眾對 AI 倫理、風險與潛力的認知。

（六） 規劃負責任的 AI 創新與採購：在鼓勵 AI 技術創新的同時，應制定負責任的採購政策，確保公部門引入的 AI 系統符合倫理、安全及合規標準。

（七） 建立風險管理機制：預先識別 AI 應用可能帶來的各種風險（如資安漏洞、偏見歧視、隱私侵犯、誤判等），並制定相應的風險評估、預警與緩解策略。

二、 Do（執行）：落實治理流程與風險控制措施

在執行階段，澳洲政府強調治理架構與執行機制的制度化。具體而言，各機關需在政策生效後 90 日內指派 AI 負責官員，負責執行 AI 治理與風險管理，要求各機關結合現行法律規範，將 AI 風險控管融入資料治理、隱私、資安與公共服務中，在治理起步後，需持續強化政府組織內部對 AI 的實務理解與應用能力，涵蓋人才、流程與文化的養成，要求在 AI 部署前完成測試並啟動持續監控，保障 AI 系統在生命週期中具備可控性與即時干預能力。供應商亦須配合採購合約中的 AI 條款，明確責任分工與安全標準。具體建議如下：

（一）部署 AI 治理與管理機制：將確立的 AI 治理原則與規範融入到公部門的日常運作流程中，確保各項 AI 應用皆能遵循既定標準。

（二）實施風險緩解措施：針對計畫階段識別的潛在風險，實施具體的緩解措施，如進行安全測試、導入去識別化技術、建立申訴管道等，以降低負面影響。

（三）執行透明公開：積極推動公部門 AI 應用的透明化，例如建立 AI 系統的內部註冊系統，公開高風險 AI 系統的使用情況，讓社會大眾了解政府如何使用 AI。

（四）指定負責官員(AOs)：指定專責的官員或部門，負責 AI 治理政策的協調、監督與推動，確保各項措施能夠有效落實。

（五）促進跨職能與跨機構協作：鑑於 AI 應用的跨領域特性，不同政府部門和單位之間應加強溝通與協作，共同應對 AI 治理的挑戰。

三、 Check（查核）：監督成效與強化透明性

查核階段重視的是治理成效與合規監督。澳洲政府要求各機關於政策施行六個月內，公布 AI 應用公開聲明，說明其用途、運作模式、法規遵循與風險緩解措施，以回應公眾信任需求。此聲明每年至少更新一次，或在重大變更時即時修正。此外，政府鼓勵建立完整紀錄與內部註冊系統，包含 AI 系統清單與相關技術文件，供後續第三方查核。對執行成果進行檢視與評估，找出問題與改善空間。具體建議如下：

（一） 持續監控 AI 系統的運行與影響：定期對已部署的 AI 系統進行監控，評估其效能、準確性、公平性及對社會的實際影響。

（二） 評估各項治理措施的有效性：檢視 AI 治理框架、風險管理機制、透明度措施等是否達到預期效果，是否有需要調整之處。

（三） 收集利害關係人回饋：透過問卷、訪談、公聽會等方式，收集公眾、公民團體、產業專家等利害關係人對公部門 AI 應用的意見與回饋，作為改進的依據。

四、 Act（行動）：制度演進與風險再評估

政府期望形成一種在風險管理與創新之間達到平衡的文化，並鼓勵機關主動應對政策與技術環境的變化。推動跨部門協調機制與員工持續培訓，促使 AI 治理能力得以內化與更新。例如，各機關須依職務提供基礎與進階的 AI 教育訓練，以確保政策要求能被實際理解與執行；同時，機關也應與利害關係人建立長期互動機制，以持續辨識偏誤與倫理風險，推動共學與制度優化，根據查核結果進行修正與優化，形成一個持續改進的循環：

（一） 根據查核結果進行調整：針對監控與評估發現的問題，及時調整 AI 治理政策、風險管理策略、技術實施方案等。

（二） 更新與迭代治理框架：鑑於 AI 技術的快速發展與社會環境的變化，AI 治理框架應保持彈性，定期更新與迭代，確保其與時俱進。

（三） 推廣最佳實踐：將成功的 AI 治理經驗與案例進行內部推廣，形成組織內的最佳實踐，提升整體治理水準。

（四） 規劃後續研究與創新：針對 AI 治理過程中面臨的挑戰，如公眾信任度、數位權利落實、數位涵容等，進行後續研究並探索創新解決方案。

第三節　研究限制與後續研究建議

AI 技術發展迅速，本研究團隊試圖整合不同領域的觀點，但目前的研析結果仍可能受到跨領域複雜性、資料蒐集挑戰、民眾信任、缺乏全球共識等因素的限制，說明如下：

一、 跨領域的複雜性

AI 治理涉及法律、倫理、技術、社會等多個層面，需要跨領域的專業知識和合作，帶來溝通和理解上的挑戰，需要持續追蹤與分析國內外 AI 發展與治理趨勢。

二、 資料蒐集與分析的挑戰

對於台灣的發展狀況，本研析分析了 DIGI^{+}案例、政府服務獎案例和臺北市政府案例。然而，要了解台灣各級政府機關 AI 的使用情況和潛在風險，仍然需要更廣泛的調查，例如「行政院所屬二三級機關 AI 使用情況調查」，以更全面地了解台灣政府部門 AI 的應用現況、種類、規劃以及面臨的挑戰，例如組織環境、組織文化、法令環境、技術掌握、資料治理成熟度、經費預算等因素。

三、 缺乏統一的全球治理框架

雖然 AI 治理已有眾多文件，但現有的架構都無法真正實現全球治理，導致代表性、協調性及執行力方面的不足，反映了在國際層面研究和推動 AI 治理的難度，需要參與國際 AI 社群或機構，以及與其他國家或地區進行合作，共同應對 AI 帶來的全球性挑戰。

四、 公眾信任與接受度的考量

政府需要透過增強透明度、治理和風險的保證，強化公眾對政府使用 AI 的信任。未來需要持續關注如何建立和維護這種信任，以及公眾對於不同 AI 應用和治理方式的接受度，以更有效地向公眾溝通 AI 的使用意圖、範圍、監控機制以及風險緩解措施，提升政策的透明度和公眾的理解。

參考文獻

Camilleri, M. A. (2024). Artificial intelligence governance: Ethical considerations. *Expert Systems with Applications*, 41. https://doi.org/10.1111/exsy.13406

de Sousa, W. G., Fidelis, R. A., de Souza Bermejo, P. H., da Silva Gonçalo, A. G., & de Souza Melo, B. (2021). Artificial intelligence and speedy trial in the judiciary: Myth, reality or need? A case study in the Brazilian Supreme Court (STF). *Government Information Quarterly*, 39(1). https://doi.org/10.1016/j.giq.2021.101660

de Sousa, W. G., Pereira de Melo, E. R., De Souza Bermejo, P. H., Araújo Sousa Farias, R., & Oliveira Gomes, A. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly*, 36(4). https://doi.org/10.1016/j.giq.2019.07.004

Engstrom, D. F., & Haim, A. (2023). Regulating government AI and the challenge of sociotechnical design. *Annual Review of Law and Social Science*, 19, 277-298. https://doi.org/10.1146/annurev-lawsocsci-120522-091626

Engstrom, D. F., Ho, D. E., Sharkey, C. M., & Cuéllar, M. F. (2020). Government by algorithm: Artificial intelligence in federal administrative agencies. *NYU School of Law, Public Law Research Paper*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3551505

Future of Privacy Forum. (June 11, 2024). Navigating governance frameworks for generative AI systems in the Asia-Pacific. *Future of Privacy Forum*. https://fpf.org

Madan, R., & Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, 40(1). https://doi.org/10.1016/j.giq.2022.101774

Marwala, T. (2024, April 19). Framework for the governance of artificial intelligence. *United Nations University*

https://unu.edu

Misuraca, G., van Noordt, C., & Boukli, A. (2020). The use of AI in public services: Results from a preliminary mapping across the EU. *13th International Conference on Theory and Practice of Electronic Governance (ICEGOV 2020)*, September 23-25, Athens, Greece. https://doi.org/10.1145/3428502.3428513

Murikah, W. (2024). Bias and ethics of AI systems applied in auditing. *ScienceDirect*, 25. https://doi.org/10.1016/S2468227624002266

Rodrigues, R. (2020). Legal and human rights issues of AI: Gaps, challenges and solutions. *ScienceDirect*, 4. https://doi.org/10.1016/S2468227624002266

Tangi, L., van Noordt, C., Combetto, M., Gattwinkel, D., & Pignatelli, F. (2022). AI Watch: European landscape on the use of artificial intelligence by the public sector. *Office of the European Union*. https://dx.doi.org/10.2760/39336

United Nations (2024, September). Governing AI for humanity. *United Nations*. https://www.un.org/en/ai-advisory-body

UNU. (2024). Framework for the Governance of Artificial Intelligence. https://unu.edu/publication/framework-governance-artificial-intelligence

van Noordt, C., & Misuraca, G. (2022). Artificial intelligence for the public sector: Results of landscaping the use of AI in government across the European Union. *Government Information Quarterly*, 101714. https://doi.org/10.1016/j.giq.2022.101774

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector—Applications and challenges. *International Journal of Public Administration*, 42(7), 596–615. https://doi.org/10.1080/01900692.2018.1498103

朱斌妤（2023）。數位權利與公部門人工智慧資料治理。*國科會計劃書*，未出版。

政府資料

1. AU Department of Industry, Science and Resources (2024), AI Ethics Principles. Retrieved May 11, 2025, from https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles.
2. AU Department of Finance (2024), National Framework for the Assurance of Artificial Intelligence in Government. Retrieved May 11, 2025, from https://www.finance.gov.au/government/public-data/data-and-digital-ministers-meeting/national-framework-assurance-artificial-intelligence-government.
3. AU Digital Transformation Agency (2024), Policy for Responsible Use of AI in Government. Retrieved May 11, 2025, from https://architecture.digital.gov.au/responsible-use-of-AI-in-government.
4. AU Department of Industry, Science and Resources (2024), Mandatory guardrails for AI in high-risk settings: developers and deployers survey. Retrieved May 11, 2025, from https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers.
5. AU Department of Industry (2024), Science and Resources, Voluntary AI Safety Standard. Retrieved May 11, 2025, from https://consult.industry.gov.au/ai-mandatory-guardrails-developers-deployers.
6. European Commission (2022), AI Watch Road to the adoption of Artificial Intelligence by the Public Sector: A Handbook for Policymakers, Public Administrations and Relevant Stakeholders. Retrieved July 14, 2025, from https://publications.jrc.ec.europa.eu/repository/handle/JRC129100.
7. The Council of Europe (2024). The Framework Convention on Artificial Intelligence. Retrieved May 11, 2025, from https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence.
8. UK Department of Science (2023), Innovation & Technology, Office for Artificial Intelligence, Policy Paper: A Pro-innovation Approach to AI Regulation. Retrieved May 11, 2025, from h

ttps://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper.

9. UK Department for Science, Innovation & Technology (2025), AI Opportunity Action Plan. Retrieved May 11, 2025, from https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan.

10. UK Government of Digital Service and UK Department for Science, Innovation & Technology (2025), Artificial Intelligence Playbook for the UK Government. Retrieved May 11, 2025, from https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government.

11. US Office of Management and Budget (2024), M-24-10 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence. Retrieved May 11, 2025, from https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf.

12. US Office of Management and Budget (2024), M-24-18 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Advancing the Responsible Acquisition of Artificial Intelligence in Government. Retrieved May 11, 2025, from https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf.

13. US Office of Management and Budget (2025), M-25-21 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Accelerating Federal Use of AI through Innovation, Governance, and Public Trust. Retrieved May 11, 2025, from https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf.

14. US Office of Management and Budget (2025), M-25-22 Memorandum FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES - Driving Efficient Acquisition of Artificial Intelligence in Government. Retrieved May 11, 2025, fromhttps://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-22-Drivin

g-Efficient-Acquisition-of-Artificial-Intelligence-in-Government.pdf.

15. US White House (2025), Initial Rescissions of Harmful Executive Orders and Actions. Retrieved May 11, 2025, fromhttps://www.whitehouse.gov/presidential-actions/2025/01/initial-rescissions-of-harmful-executive-orders-and-actions/.

網站資料

1. Artificial Intelligence (Regulation) Bill [HL] (2023). Retrieved May 11, 2025, from https://bills.parliament.uk/publications/53068/documents/4030.
2. Madison Alder (2025). Trump White House releases guidance for AI use, acquisition in government. Retrieved May 11, 2025, from https://fedscoop.com/trump-white-house-ai-use-acquisition-guidance-government/.
3. Nathalie Moreno (2025). The Artificial Intelligence (Regulation) Bill: Closing the UK's AI Regulation Gap? Retrieved May 11, 2025, from https://kennedyslaw.com/en/thought-leadership/article/2025/the-artificial-intelligence-regulation-bill-closing-the-uks-ai-regulation-gap/.
4. UK House of Lords, King's Speech, Library Briefings (2024). Retrieved May 11, 2025, https://researchbriefings.files.parliament.uk/documents/LLN-2024-0040/LLN-2024-0040.pdf.

附錄

附錄一、研析團隊內部會議綱要

開會時間	開會地點	參與人員	開會主題
2024/10/28	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問	專案分工與內容確認
2024/11/11	實體會議	曾憲立副教授、朱斌妤教授、蕭乃沂顧問	機關調查題目與方式討論
2024/11/25	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問	進度報告討論
2024/12/09	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問	進度報告討論
2025/02/17	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問	進度報告討論
2025/03/10	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問	AI 風險分類、治理政策建議
2025/03/24	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問、萬幼筠副總經理	AI 風險管理、實務觀察心得
2025/03/27	實體會議	曾憲立副教授及其研究助理(參與人數共 8 人)	進度報告討論
2025/04/10	實體會議	曾憲立副教授及其研究助理(參與人數共 8 人)	進度報告討論
2025/05/01	實體會議	曾憲立副教授及其研究助理(參與人數共 8 人)	進度報告討論
2025/05/27	實體會議	曾憲立副教授及其研究助理(參與人數共 8 人)	進度報告討論

開會時間	開會地點	參與人員	開會主題
2025/05/27	線上會議	曾憲立副教授、朱斌妤教授、戴豪君副教授、許慧瑩助理教授、蕭乃沂顧問	期末報告進度

附錄二、DIGI+案例[47]

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
執行 (enforcement)	智慧識別流程	5G 智慧警察行動服務研析【內政部警政署】	影像辨識	V	V	V	V	V	
		AR 頭盔前進搜救現場第一線，結合 AI 提昇救援效率【內政部消防署】	影像辨識	V	V		V	V	
		建置公司登記文件影像自動分類，節省分類建檔成本【經濟部商業發展署】	影像辨識、機器學習	V			V	V	
		車牌辨識結合 AI，智慧勾稽異常清運行為【環境部環境管理署】	影像辨識	V	V	V	V	V	V
		港區及聯外道路車牌辨識系統、船舶軌跡航行監控分析及新世代海巡偵防業務整合系統【海洋委員會海巡署偵防分署】	影像辨識	V	V	V	V	V	V
		接軌國際輻防技術規範與精進量測技術能力	影像辨識	V	V	V	V	V	V

[47] 考量篇幅濃縮，僅標示應用類型、案例名稱、AI 技術及風險面向

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		【核能安全委員會】(涉及資訊分析流程)							
		智能稅務服務 【財政部財政資訊中心】	機器學習（ML）	V			V	V	
	預測性執法流程	強化虛實治安情資整合機制，提升員警破案效率 【內政部警政署】	機器學習（ML）、 深度學習（DL）、 AI 福爾摩斯圖運算	V	V		V	V	
	支援檢查流程	職業駕駛之不安全行為預警系統建置-以大眾運輸業為例【勞動部勞動及職業安全衛生研究所】	影像辨識、 機器學習、LSTM 模型				V	V	
管制分析與監控	資訊分析流程	使用大數據分析支援國境安全決策 【內政部移民署】	機器學習（ML）、 深度學習（DL）	V	V		V	V	V
		智慧防疫空間及空氣品質數據分析 【衛生福利部疾病管制署】	大數據分析、邊緣運算 AI	V			V		V
		建立大臺北地區、臺南地區地質大數據，提供防災應用 【經濟部地質調查及礦業管理中心】	大數據分析	V			V		V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		建立旅客安全警示－精準篩濾有疑旅客 【內政部移民署】	機器學習（ML）、深度學習（DL）	V			V		V
		食品安全高風險異常廠商偵測模型 【環境部化學物質管理署】	大數據分析(VGAE、SCAD)模型	V		V	V	V	V
		發展智慧化地震預警系統 【交通部】	大數據分析	V			V		V
	監控政策執行狀況	建築工程技術精進創新與應用效能提升 【內政部】	大數據分析	V			V		V
	預測和規劃	建置台水總售水量預測模型，輔助用水預估 【經濟部資訊處】	機器學習（ML）、深度學習（DL）	V			V		V
		建置中部地區山崩潛感模型，支援防災決策 【經濟部地質調查及礦業管理中心】		V			V		V
		以 AI 技術估計颱風強度 【交通部中央氣象署】	CNN 深度學習模型	V			V		

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		應用人工智慧技術開發數值模式預報加值產品，並進行天氣個案測試與評估 【交通部中央氣象署】	AI 後處理技術、 類神經長短期記憶模型（LSTM）、圖神經網路（GNN）、TF-IDF 演算法、 深度學習	V			V		V
裁決 (adjudication)	**就福利做出決定**								
公共服務和參與(Public services and engagement)	**參與管理**	內政大數據加值應用 【內政部】	大數據分析	V	V		V	V	
	資料分享管理	內政大數據加值應用 【內政部】	大數據分析	V	V		V	V	
		法務部調查局鑑識科學大樓遷置暨科學偵查檢驗設備精進中程研析 【法務部】	機器學習（ML）、 深度學習（DL）、RFID 技術	V			V	V	V
	服務整合	運用「評點制核可人次統計資料分析」之大數據 【僑務委員會】	大數據分析	V	V	V	V	V	V
		榮家無線網路佈建研析 【國軍退除役官兵輔導委員會】	結合 5G 、AIoT、 大數據分析	V	V	V	V	V	V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		衛福業務數位轉型服務躍升【衛生福利部】	大數據分析	V		V	V	V	
		故宮藝術資料 AI 技術應用【國立故宮博物院】	風格轉換（Style Transfer）、機器學習（Machine Learning）、VR 與 3D 建模（Virtual Reality & 3D Modeling）	V		V	V		V
	服務個性化	氣象領域維運與技術發展及智慧海象環境災防服務【交通部】	大數據分析	V			V		V
		文化數據智能分析與決策輔助研析【文化部】	大數據分析、自然語言處理系統	V			V	V	V
內部管理 (Internal management)	內部基本流程	建構無縫隙氣象服務價值鏈－橋接農、漁、光電領域研析【交通部】	大數據分析	V			V		V
	內部支援流程	資料檢索系統【財政部關務署】	自然語言處理（NLP）	V			V		V
		新世紀檢察 AI 智慧輔助系統建置案【法務部】	自然語言處理	V			V	V	V
	內部管理流程	建立原住民族高等教育人才培育決策【原住民族委員會】	大數據分析	V	V	V	V	V	V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		完備有效之循證決策模式，提升政府服務及施政決策之精準度 【行政院人事行政總處】		V	V	V	V	V	V
		整合跨系統研析資料，建置決策支援模組 【國家發展委員會管制考核處】	大數據分析	V		V	V		V
		建置以資料科學為基礎之社會政策治理機制 【國家發展委員會社會發展處】	大數據分析	V		V	V		V
		建立原住民族部落長者長照服務決策資訊 【原住民族委員會】	大數據分析	V		V	V	V	V
		運用新興科技工具與遙測技術，提升廢棄物棄置場址監控及執法效能 【環境部環境管理署】	大數據分析	V		V	V	V	V
		建立原住民族社會福利服務資源分布資訊 【原住民族委員會】		V	V		V	V	V
		推動社會政策循證決策治理機制 【國家發展委員會社會發展處】	大數據分析	V			V		V

資料來源：本研析自行整理

附錄三、第 1~6 屆政府服務獎案例

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
執行 (enforcement)	智慧識別流程	彰化縣消防局「AI 精準消防-防災影像辨識預警及救災任務語音紀錄」	影像辨識、語音辨識	V	V		V	V	
		嘉義市政府警察局「諸羅城縱橫攻略防治交通事故從事故處理開始」	影像識別	V			V	V	
	審計和日誌記錄管理								
	預測性執法流程	新北市消防局「全災行智慧化指揮監控平台」	圖像視覺化	V			V	V	V
	支援檢查流程								
管制分析與監控	資訊分析流程	臺北市政府工務局汙水管 AI 檢視	影像辨識	V			V		V
		臺中市政府建設局自來水管「科技檢漏」	機器學習、動態影像分析辨識演算法	V			V		V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		基隆市消防局救災救護新應用「智慧消防 2.0」	影像分析	V	V			V	V
		臺中市政府交通局「五心級」智慧交通管理系統	大數據分析	V			V	V	
		臺南市政府交通局 AI 智慧車流辨識及智慧號誌管理策略	大數據分析	V			V	V	V
		臺北市政府交通局「智慧化交通量調查分析系統」	大數據分析	V			V		
		環保署毒物及化學物質局「從目測到遙測，從太空看台灣」	圖像識別	V			V	V	
		臺中榮民總醫院「AI 大師來精算，智慧照護零時差」	大數據分析	V	V	V	V	V	V
		新北市政府養護工程處「iRoad 新北市智慧道路管理中心」	影像辨識	V			V		
		高雄市立凱旋醫院「AI 照護心體驗，保命防跌新神氣」	影像辨識、大數據分析	V	V	V	V	V	V
		國民健康署結合國家衛生研究院開發癌症登記 AI 輔助程式、肺癌影像輔助程式	自然語言處理、影像辨識	V	V	V	V	V	V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		抑制河川揚塵 臺中市環保局 AI 智慧水線	大數據分析	V			V		
		雲林縣環保局「智能科技監控與提升家園環境品質」	大數據分析建模、影像辨識、無人機	V	V	V	V	V	V
	監控政策執行狀況								
	預測和規劃	高雄市政府毒品防制局「ICARES AI 科技輔導～走出藥癮迷途」	大數據分析	V	V		V	V	
裁決 (adjudication)	就福利做出決定								
公共服務和參與(Public services and engagement)	參與管理	臺南市政府衛生局「科技防疫現代蚊清」	大數據分析、聊天機器人	V					
		內政部警政署刑事警察局「警察電商聯盟--終詐之戰」		V	V			V	V
	資料分享管理	臺南市政府智慧發展中心「推動城市數據交換，打造未來城市自主感知」	大數據分析	V			V		
		臺南市政府水利局「及時水情一點通：智慧防災推動研析」	影像辨識	V			V		V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
	服務整合	環保署（現為環境部）環境監測及資訊處「空品智慧 GO，創新服務很足夠」以及「AIOT 科技神助功，環境治理好成」	機器學習、大數據分析	V			V		
		健保署「健保大數據跨域合作 數位科技防疫新典範」	大數據分析、影像辨識	V	V	V	V	V	V
		高雄市政府地政局「BEST 價+給你掌握數據，價值永續」	大數據分析	V		V	V		
	服務個性化	臺北捷運 AI 智慧客服	chatbot	V			V	V	
		交通部觀光署 AI 翻譯櫃台		V	V	V			
		臺南市政府教育局「生成式 AI 輔助學習中介平台」	自然語言處理	V		V		V	
		臺東縣衛生局「臺東南迴原鄉衛生所躍升醫學中心及服務品質研析」	自然語言處理	V	V	V		V	
		臺北榮民總醫院「藥安心：以 AI 創新居家用藥整合服務，守護民眾用藥安全」	影像辨識	V		V	V		V
		健保署南區業務組 「天涯若比鄰 醫定守護您」	大數據分析	V	V	V		V	
	內部基本流程								

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
內部管理 (Internal management)	內部支援流程	新北市政府警察局 「智慧城市 安全新北」	大數據分析	V	V	V	V	V	V
	內部管理流程								

資料來源：本研析自行整理

附錄四、臺北市政府案例

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
執行 (enforcement)	智慧識別流程	智慧號誌【交工處】	影像辨識	V	V	V	V	V	V
		臺北市智慧影像事件偵測建置案【交工處】	影像辨識	V			V		V
		毒品資料圖像快搜【臺北市政府警察局刑事鑑識中心】	影像辨識 (CNN)	V		V	V		
	審計和日誌記錄管理								
	預測性執法流程								
	支援檢查流程								
管制分析與監控	資訊分析流程	AI 輔助診斷糖尿病視網膜病變【臺北市立聯合醫院】	AI 眼底辨識	V	V	V	V	V	V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		AI 輔助診斷糖尿病 2.0 視網膜病變【臺北市立聯合醫院】	AI 眼底辨識	V	V	V	V	V	V
		血壓血氧上傳系統【臺北市立聯合醫院】	大數據分析	V	V	V	V	V	V
		LDCT（低劑量肺部電腦斷層掃描）AI 影像電腦輔助判讀系統【臺北市立聯合醫院】	影像辨識	V	V	V	V	V	V
		運用 AI 協審水土保持研析【工務局大地工程處】	NLP	V			V		V
		地籍套繪都市研析使用分區圖查詢便民服務【都發局】	自動化繪圖、GIS	V					
		翡翠水庫大壩安全監測評析【臺北翡翠水庫管理局】	類神經網路、大數據分析	V			V		V
		網搜防偷跑—預售建案銷售廣告檢索【臺北政府地政局】	大數據分析	V					
	監控政策執行狀況								

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
	預測和規劃	自來水管線汰換規劃【臺北自來水事業處】	WebGIS、大數據分析	V			V		
		管網水理模型校正【臺北自來水事業處】	大數據分析、基因演算法(GA)	V			V		
裁決	就福利做出決定								
公共服務和參與(Public services and engagement)	參與管理								
	資料分享管理								
	服務整合	智慧城市【資訊處】	影像辨識、IoT、大數據分析、ML	V	V	V	V	V	V
		數位發展部數位產業署智慧城鄉生活應用補助研析(地方試煉暨國際合作)之「智慧社區健康好食運研析」【臺北市立聯合醫院】		V	V	V	V	V	V
		雲端醫院	ML、NLP	V	V	V	V	V	V

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		【臺北市立聯合醫院】							
	服務個性化	臺北酷課雲「酷AI(CooCAI)」【臺北市政府教育局】	NLP、語音辨識、機器學習	V		V	V	V	
		輪椅旅客自動叫梯服務【臺北大眾捷運股份有限公司】	影像辨識	V		V	V		
		AI 人工智能線上客服【產業發展局】	chatbot	V		V		V	
		智能客服系統文字查詢及申辦【臺北自來水事業處】	chatbot	V		V		V	
		登記教主新服務—繼承登記法服諮詢【臺北市建成地政事務所】	chatbot	V		V		V	
		臺北市競技運動訓練暨科學中心研析【體育局】	大數據分析	V		V	V	V	V
內部管理(Internal management)	內部基本流程								
	內部支援流程	人事法規、人事系統操作及同仁人事業務相關權益諮詢	NLP	V	V	V	V	V	

AI 應用類型	應用類型	案例名稱	使用 AI 技術	資料治理	人權保護	倫理	模型可解釋性	個人資料	自動化決策風險
		事項 【臺北市府人事處】							
		翡翠水庫經營管理智慧決策系統 【臺北翡翠水庫管理局】	類神經網路、大數據分析	V		V	V		V
		翡翠水庫集水區環境影像變異監測 【臺北翡翠水庫管理局】		V		V	V		V
	內部管理流程	建置災害應變雲端協作平臺 【臺北市政府消防局】	機器學習、數據整合與視覺化	V			V		V
		決策輔助系統 【臺北市立聯合醫院】	大數據分析	V	V	V	V	V	V
		護理資訊系統 【臺北市立聯合醫院】	大數據分析、ML、NLP	V	V	V	V	V	V

資料來源：本研析自行整理

附錄五、2023 DIGI+案例介紹

部會名稱	計劃名稱	案例內容
【內政部警政署】	5G 智慧警察行動服務研析	1. 5G M-Police 行動影音系統：透過智慧車牌辨識系統與邊緣運算技術，提升警察巡邏查緝車輛的效率，減少手動輸入車牌號碼的時間，強化辦案效能。 2. 智慧 XR 警勤訓練系統：建立沉浸式訓練環境，模擬警員執勤情境，透過虛擬和現實結合的方式進行訓練，提升警員的應對能力和記憶力。
【內政部消防署】	AR 頭盔前進搜救現場第一線，結合 AI 提昇救援效率	1. 建置符合國際標準的智慧搜救平臺，將傳統紙本作業表單電子化，協助災害現場搜救人員、犬隻及裝備器材進行資訊化管理。 2. 在救災頭盔中整合 AR 眼鏡、紅外線、熱顯像儀和 AI 影像辨識，幫助救援人員即時回傳現場影像，辨識受困者數量及環境危害。 3. 結合 AI 大數據與熱顯像儀，於倒塌建築物內自動辨識活體生物及危險溫度，提高尋獲受困者的機率，並警示避免高危險區域。
【內政部警政署】	強化虛實治安情資整合機制，提升員警破案效率	1. 運用 AI 福爾摩斯圖運算引擎整合虛擬（如社群媒體）與實體情資（如失車、毒品等），強化案件偵查。 2. 透過情資比對和數位身分辨識，協助警員進行案件追查，提升破案效率。
【內政部移民署】	使用大數據分析支援國境安全決策	1. 自 2020 年起，移民署整合航前旅客、入出境查驗、移民管理等系統資料，建置大數據分析平台進行資料分析。 2. 2023 年完成 5 項視覺化儀表板、2 項社會網絡分析（SNA）及 2 項 AI 預測，提升數據分析能力。 3. 迄今完成 31 項動態視覺化儀表板，幫助強化政策規劃、人力估算及案件偵辦。
【內政部移民署】	建立旅客安全警示－精準篩濾有疑旅客	1. 移民署建置系統，蒐集旅客訂位資料，並建立多重證件資料勾稽機制，提升旅客資料篩查精準度。 2. 結合大數據分析，開發滯臺旅客風險等三項警示規則，有效識別高風險旅客。

部會名稱	計劃名稱	案例內容
		3. 透過精確的篩檢機制，提升移民官在篩濾可疑旅客時的效率與口詢精準度。
【內政部建研所】	建築工程技術精進創新與應用效能提升	1. 結合大數據分析，模擬都市通風特性並繪製通風地圖。 2. 建立風廊查詢系統，支持都市規劃與通風管制。
【內政部】	內政大數據加值應用	1. 內政部與衛福部合作，利用大數據串聯應用，主動識別需要幫助的老年人，並提供相應的服務。 2. 民眾可結合「內政大數據模擬資料」與「銀髮安居資料」，針對社會問題提出解決方案。內政部也在社會經濟資料服務平台（SEGIS）釋出大數據模擬資料，供各級政府與企業使用。
【衛生福利部中央健康保險署】	健保大數據數位應用	1. 資料整合：結構化數據與影像資料庫的跨域合作，建置心臟功能評估和身體部位辨識模型。 2. AI 運算環境：完善容器管理平台，支持 AI 模型部署與應用。 3. 審查運用：使用牙科拔牙術式判讀模型，提高審查效率。 4. 影像分群模型：擴展分類範圍至 22 類影像，精確率達 90%。
【衛生福利部疾病管制署】	智慧防疫空間及空氣品質數據分析	1. 2023 年升級為可監測 CO2、PM2.5、溫度、濕度的感測器，並與通風改善設備連動，自動改善空氣品質。 2. CO2 超過 900ppm 啟動通風設備，恢復至 800ppm 時自動關閉，透過濾網過濾空氣後換氣。 3. 空氣品質數據透過 API 上傳至國網中心物聯網平台進行監測及查閱。
【衛生福利部】	衛福業務數位轉型服務躍升	1. 建立逾期案例資料庫：蒐集國內外食品數據及逾期食品案例共 4,831 件，包含 Tifsan、PMDS 等國內外資料庫及警訊網站。 2. 智能監測模組：運用統計科學與大數據分析技術，監控 3 種食品類型，產出 30 件高風險逾期食品清單，用於加強稽查力度。

部會名稱	計劃名稱	案例內容
		3. 系統介面優化：新增業者風險分數與逾期歷史案例查詢功能，提升稽查效率。 4. 大數據服務平臺：推動資料交換標準，增強 AI 應用能力並支持政策決策。
【經濟部商業發展署】	建置公司登記文件影像自動分類，節省分類建檔成本	1. 利用機器學習技術開發 自動分類模型，自動辨識並分類公司登記申請文件。 2. 將分類模型與審查系統串聯，實現登記文件的數位化處理，取代傳統人工分類、掃描和歸檔的繁瑣程序。
【經濟部地質調查及礦業管理中心】	建立大臺北地區、臺南地區地質大數據，提供防災應用	1. 整合多來源地質資料（如 Geo2010、營建工程資料庫等），篩選出具有代表性的地質數據。 2. 建置大臺北與臺南地區的 三維地質模型資料庫，累計分析出有效孔位（臺南 3,103 孔、大臺北 4,180 孔）。 3. 提供地質模型作為防災應用的基礎資料，例如建置 土壤液化風險圖，輔助防災與建築設計。
【經濟部資訊處】	建置台水總售水量預測模型，輔助用水預估	1. 自動串接多種數據來源（如總售水量、經濟與社會影響因子）進行分析，減少人力投入。 2. 運用數據分析技術進行用水預測，提供高達 96.2% 準確度 的預測結果。 3. 以視覺化圖表呈現預測結果及波動趨勢，輔助台水公司進行供水調配與預算編列。
【經濟部地質調查及礦業管理中心】	建置中部地區山崩潛感模型，支援防災決策	1. 蒐集與彙整降雨與地震誘發山崩的資料，包括各集水區的山崩目錄及衛星影像。 2. 建立 降雨誘發山崩潛感圖 與 地震誘發山崩潛勢分級圖，提供精準的山崩潛感預測。 3. 自動產製模組，結合多種誘發因子（地震與降雨），生成山崩潛感模型，用於防災決策和救災資源配置。
【環境部環境管理署】	車牌辨識結合 AI，智慧勾稽異常清運行為	1. 建置廢棄物非法棄置智慧圍籬系統，透過車牌辨識和 AI 技術，以便快速鎖定可疑車輛。 2. 為解決違法清運機具造成之全國性非法棄置問題，有效掌握各類違法清運模式

部會名稱	計劃名稱	案例內容
【環境部化學物質管理署】	食品安全高風險異常廠商偵測模型	1. 進行高食安風險異常廠商偵測。 2. 結合圖神經網絡（VGAE）分析廠商間的關係，並結合自監督學習（SCAD）找出異常行為，大幅提升了偵測高風險廠商的準確性。 3. 納入多元數據，增強模型能力。 4. 預訓練與微調，加速模型學習。
【環境部環境管理署】	運用新興科技工具與遙測技術，提升廢棄物棄置場址監控及執法效能	1. 透過行車路徑資料演算，分析及建立清運車輛行車軌跡路徑特徵。 2. 透過分析產出廢棄物非法棄置潛勢熱區，完成低、中、高非法棄置風險場域地圖。
【財政部財政資訊中心】	智能稅務服務	1. 建置智能稅務架構，提供稅務數據分析與稅收估測的支援服務。 2. 提供營業稅稅收估測數據，可作為年度稅收規劃的參考。
【財政部關務署】	資料檢索系統	1. 提供進出口報單檔、稅則簽審檔、稅則疑義等資料的智慧查詢功能。 2. 應用了關鍵字檢索技術、自然語言處理（NLP）技術，以及資料分類與關聯查找演算法。
【交通部】	發展智慧化地震預警系統	1. 中央氣象署建置「臺灣地震與地球物理資料管理系統（GDMS）」。 2. 提供政府單位或相關學者分析，助於了解台灣地震特性及加強防災工作。
【交通部中央氣象署】	以 AI 技術估計颱風強度	1. 開發交叉頻譜功能，擷取多頻道衛星影像特徵。 2. 提高颱風強度估計準確性，解決無法進行飛機實地觀測的資訊缺口問題。
【交通部中央氣象署】	應用人工智慧技術開發數值模式預報加值產品，並進行天氣個案測試與評估	1. 颱風系集定量降雨預報（AI-ETQPF）：AI 後處理強化降雨預報，改善降雨過度預報問題，提高預報準確性。 2. 空氣品質（PM2.5 濃度）預報：運用 AI 模型（如 LSTM）進行全臺 PM2.5 預測，預測精度顯著提升，RMSE 降低至 2.8 μg/m³。 3. 智慧型作業監控管理平臺：運用 LSTM、GNN、TF-IDF 等技術實現異常預警、自動化

部會名稱	計劃名稱	案例內容
		處理，系統可用性達 99.99%，異常事件處理時間縮短至小於 1 分鐘。
【交通部】	氣象領域維運與技術發展及智慧海象環境災防服務	1. 建構環島異常波浪預警系統。 2. 「海象環境資訊平台」高精度導航等級手機定位潮流預報服務。 3. 「智慧交通大數據」提供地理資訊數據，可供商船、交通船進出港、遊艇操船及帆船運動應用。 4. 「海象環境資訊平台」擴增近海海象季節風險資訊產品。 強化颱風影響期間各方向影響、高溫預警之分析。
【交通部】	建構無縫隙氣象服務價值鏈－橋接農、漁、光電領域研析	1. 提供 1 至 14 天高解析極端高低溫機率預報，輔助農漁業防災減災規劃。 2. 完成臺灣地區高溫預警預報系統，定期發布週至月尺度高溫預報資訊。 3. 建置臺灣地區格點高溫預警預報系統，每月定期提供週至月時間尺度的極端高溫預報資訊，並完成 1 至 14 天作業化高解析格點逐日極端高低溫機率預報產品
【法務部】	法務部調查局鑑識科學大樓遷置暨科學偵查檢驗設備精進中程研析	1. 使用 RFID 技術蒐集涉案車輛 EPC 外碼，建立車行紀錄，供偵查人員檢索分析，AI 分析車輛經常出沒地點，協助掌握案件涉嫌人行蹤。 2. 整合來自多機構的資料（格式不同），進行行政區分類以提高資料相容性和檢索效率。 3. 將車輛行跡顯示於地圖，輔助偵辦與分析。
【法務部】	新世紀檢察 AI 智慧輔助系統建置案	1. 介接與整合功能：與內政部警政署案管系統、165 反詐騙資料庫及刑事案件資料進行介接，確保數據整合性。將 AI 嵌入現行筆錄與書類製作系統，實現無縫作業。 2. NLP 應用：訓練模型分析偵訊筆錄、前科表及 165 資料庫特定欄位，提供案件相關建議文本（如累犯論述）。自動生成案件書類及附表，減少檢察官處理重複性工作的負擔。 3. 累犯判斷功能：根據累犯邏輯和刑案資料庫，自動生成起訴書和判決書中的相關文字，提升準確性與效率。

部會名稱	計劃名稱	案例內容
【原住民族委員會】	建立原住民族高等教育人才培育決策	1. 為支援業務單位提出原住民族高等人才培育需求建議類別。 2. 原住民族委員會建立「建議不受外加名額限制 14 科系學類」之自動決策功能，作為擬定原住民族高等教育未來培育重點學門（科）之參考依據。
【原住民族委員會】	建立原住民族部落長者長照服務決策資訊	1. 建立基於原住民文化健康站資訊的長照決策支援系統。 2. 系統自動計算建議長照服務機構類型，提供決策參考。
【原住民族委員會】	建立原住民族社會福利服務資源分布資訊	1. 原民會建立「都會區原家中心設站需求評估」自動決策、「社會福利人口群分布統計」等決策支援功能。 2. 協助原家中心進行服務目標規劃、設站需求之評估，並作為原民會相關福利政策規劃與資源配置之參考依據。
【國家發展委員會管制考核處】	整合跨系統研析資料，建置決策支援模組	1. 國家發展委員會完成「政府研析資料庫（GDB）」系統建置，提供研析全生命週期資訊查詢、資料介接及統計分析服務，並已與多個政府機關及國際系統整合，建置 307 項資料庫及 44 項 API 服務，提升研析透明度與跨機關協作。 2. 建置「公建預警分析」決策支援模組，以研析進度及經費執行等角度統計分析，提供研析規劃、執行及效益評估之決策參考。
【國家發展委員會社會發展處】	建置以資料科學為基礎之社會政策治理機制	1. 運用資料科學針對民眾關切社會議題，採取大數據等分析技術，進行相關政策議題實證分析。 2. 作為相關政策規劃與資源配置之證據基礎。
【國家發展委員會社會發展處】	推動社會政策循證決策治理機制	1. 追蹤社會趨勢，例如「經濟就業與居住資源」、「數位轉型與科技影響」等重大領域範疇進行網路媒體分析，更新 28 項關鍵課題。 2. 完成社會議題循證決策，進行跨領域協作與模型優化。
【文化部】	文化數據智能分析與決策輔助研析	1. 透過大數據分析，了解民眾對台灣文化景點的評價，以協助文化機構提升服務品質。 2. 採用文字探勘技術，分析 81 萬筆 Google 評論，建立詞庫並進行語意分析。

部會名稱	計劃名稱	案例內容
【僑務委員會】	運用「評點制核可人次統計資料分析」之大數據	1. 經介接勞動部資料，並比對僑務委員會僑生系統計算，「 推升智慧服務-僑生運用評點制留臺數據分析」，研析僑生與產業之關聯性，擴大推動僑生留臺就業。 2. 做為留用畢業僑生政策之參考，以吸引更多畢業僑生透過評點制留臺。
【國軍退除役官兵輔導委員會】	榮家無線網路佈建研析	1. 研析是透過 5G 通信技術下，以 AIoT （Artificial Intelligence of Things，人工智慧物聯網）建置無線即時傳輸心電圖系統，及時連結榮總或榮院緊急救治榮家患者。 2. 板橋榮家透過無線網路先導驗證，全面導入生理量測無線傳輸等智慧照護應用。
【國立故宮博物院】	故宮藝術資料 AI 技術應用	1. 以 VR 為載具，結合 AI 機器學習技術，應用 AI 進行東西畫風之風格學習並自動摹擬，再以 VR 製作結合 3D 造景及眼球追踪等體驗功能，提供觀眾頭戴式沉浸體驗。 2. 故宮利用 TensorFlow 等 AI 工具，結合資料推出三種不同主題的 AI 教案，並在國中小藝術或語文社會課程中進行示範與推廣，同時舉辦兩場工作坊，訓練教師如何在實際教學中使用這些 AI 工具與教案。
【海洋委員會海巡署偵防分署】	港區及聯外道路車牌辨識系統、船舶軌跡航行監控分析及新世代海巡偵防業務整合系統	1. 為強化邊境管理，全面建置全臺港區及聯外道路車牌辨識攝影機，掌握不法動態，提升邊境治安監控能力。 2. 船舶軌跡航行監控分析系統，整合漁業署漁船 VMS 資料和海巡署雷達系統，該系統建立了關鍵字查詢介面，並透過分析強化海上監視與偵查能力。 3. 建置含有 WatchDog 監控功能、犯罪調查系統等功能的績效資料庫，此系統強化了跨機關的情資共享。
【行政院人事行政總處】	完備有效之循證決策模式，提升政府服務及施政決策之精準度	1. 辦理公務體系高齡化之研析。 2. 研析退休新制對公務機關人員老化之影響、公務機關各類人員高齡化之情形等，提供政府考試、任用、訓練、福利及人力運用等業務推動及政策制定之參考。

部會名稱	計劃名稱	案例內容
【核能安全委員會】	接軌國際輻防技術規範與精進量測技術能力	1. 使用 AI 影像擷取與分析輔助 γ-H2AX 螢光訊號分析。 2. 建立低劑量輻射劑量檢量線，以提升輻射意外事故發生人員劑量重建技術之效能，並利後續醫療監護作業
【勞動部勞動及職業安全衛生研析所】	職業駕駛之不安全行為預警系統建置-以大眾運輸業為例	1. 運用影像辨識建構個人化駕駛預警系統，降低交通事故。 2. 使用 LSTM 模型預測駕駛行為，透過錄影設備分析 25 個肢體點，過濾下半身動作。系統可提前 2 秒預測駕駛行為，提醒未完成的前置駕駛行為（如查看後照鏡）。 3. 依駕駛習慣預測行為，提供針對性提醒（如注意右後照鏡），幫助駕駛提前反應。

資料來源：本研析自行整理

附錄六、第 1~6 屆政府服務獎案例介紹

應用領域	案例名稱	案例內容
智慧警政消防	基隆市消防局救災救護新應用「智慧消防 2.0」	以 EMS 智能眼鏡同步視訊，加速到院急救作業。
	內政部警政署刑事警察局「警察電商聯盟--終詐之戰」	以 AI 自動辨識涉詐門號，封鎖境外可疑來電。
	新北市消防局「全災行智慧化指揮監控平台」	運用 AI 分析 IOT 感測大數據，建立災害預測模型。
	新北市政府警察局 「智慧城市 安全新北」	全方位 3D 科技維安網以警用無人機、M-Police、IP-CAM 等利於掌握現場即時狀況。
	彰化縣消防局「AI 精準消防-防災影像辨識預警及救災任務語音紀錄」	AI 監控影像防災辨識系統，透過各鄉鎮的路口監視器自動辨識出火災、水災及車禍交通事故等現場資訊。無線電 AI 監聽系統，即時將無線電語音自動轉換產生成逐字稿並記錄。
	高雄市政府毒品防制局「ICARES AI 科技輔導～走出藥癮迷途」	透過 AI 技術雷達圖分析藥癮者之風險等級，進行高風險預警。
智慧交通	嘉義市政府警察局「諸羅城縱橫攻略 防治交通事故從事故處理開始」	AI 繪圖與影像辨識等技術，提高交通事故分析準確性。
	臺中市政府交通局「五心級」智慧交通管理系統	運用人工智慧 AI 技術、IoT 物聯網，建置智慧交通管理系統 T-TOPIS，串接「臺中交通網」APP 與即時交通資訊網，智慧調控紅綠燈時長以控制車流。
	臺南市政府交通局 AI 智慧車流辨識及智慧號誌管理策略	AI 影像辨識技術監控車流，進行紅綠燈的智慧即時調控。
	臺北捷運 AI 智慧客服	AI 智慧客服串接北捷遺失物系統，文字機器人協助查找失物。客服同時具備「一般事件通報」功能，可以即時處理行車問題。
	臺北市政府交通局「智慧化交通量調查分析系統」	運用 YOLOv4 辨識及 SORT 追蹤技術，提供更完整多元交通大數據資料，協助解決交通壅塞與安全性問題。
智慧環境	環保署（現為環境部）環境監測及資訊處「空品智慧 GO，創新服務很足夠」以及「AIOT 科技神助功，環境治理好成」	整合 AI 機器學習技術，影像解析等技術，協助地方環保稽查。

應用領域	案例名稱	案例內容
	環保署毒物及化學物質局「從目測到遙測，從太空看台灣」	結合 AI 技術建構全國石棉屋瓦資料。
	抑制河川揚塵 臺中市環保局 AI 智慧水線	運用 AI 智慧科技，當空氣品質達一定程度時，啟動灑水系統於河川進行減塵。
	雲林縣環保局「智能科技監控與提升家園環境品質」	運用 AI 進行污染熱區分析以及自動通報，包括空氣品質監測站、移動式監測車、空氣品質微型感測器、水質感測器、AI 視覺煙霧探測、車輛噪音及車牌辨識系統、無人機搭載顯像儀。
智慧城市	新北市政府養護工程處「iRoad 新北市智慧道路管理中心」	導入 AI 影像辨識，輔助人力巡查道路狀況。
	臺南市政府智慧發展中心「推動城市數據交換，打造未來城市自主感知」	以 AI 進行自動分析管理。
	臺中市政府建設局 自來水管「科技檢漏」	運用「新型科技檢漏技術」，可透過「AI 人工智慧測器」建構資料庫，配合透地雷達協助檢漏人員判斷肉眼無法透視的地底情形。
	臺北市政府工務局 汙水管 AI 檢視	透過 AI 判別污水下水道異常類別。
	臺南市政府水利局「及時水情一點通：智慧防災推動研析」	介接地下道監視器與淹水感測器進行 AI 分析。
智慧醫療	臺南市政府衛生局「科技防疫 現代蚊清」	以 AI 多元監控蚊子生命週期。
	健保署「健保大數據跨域合作 數位科技防疫新典範」	運用 AI 進行數據分析，影像辨識等，推動精準醫療與長照服務。
	健保署南區業務組 「天涯若比鄰 醫定守護您」	運用 AI 進行數據分析，影像辨識等，推動精準醫療與長照服務。
	臺北榮民總醫院「藥安心：以 AI 創新居家用藥整合服務，守護民眾用藥安全」	運用 AI 進行數據分析，影像辨識等，推動精準醫療與長照服務。
	臺東縣衛生局「臺東南迴原鄉衛生所躍升醫學中心及服務品質研析」	運用 AI 進行數據分析，影像辨識等，推動精準醫療與長照服務。

應用領域	案例名稱	案例內容
	臺中榮民總醫院「AI 大師來精算，智慧照護零時差」	運用 AI 進行數據分析，影像辨識等，推動精準醫療與長照服務。
	高雄市立凱旋醫院「AI 照護心體驗，保命防跌新神氣」	運用 AI 進行數據分析，影像辨識等，推動精準醫療與長照服務。
	國民健康署結合國家衛生研究院開發 癌症登記 AI 輔助程式、肺癌影像輔助程式	導入 NLP，協助癌症登記報告。利用 AI 開發 LDCT 影像輔助程式，進行複雜的肺結節標註與報告繕打作業。
	高雄市毒品防治局「ICARES AI 科技輔導～走出藥癮迷途。」	運用雷達圖 AI 分析風險因子，精準評估輔導方向，提升輔導成效。
智慧行政	高雄市政府地政局「BEST 價+給你掌握數據，價值永續」	以 AI 智慧化審視地籍圖資，提供民眾明瞭資訊。
智慧教育	臺南市政府教育局「生成式 AI 輔助學習中介平台」	生成式 AI 輔助學習中介平台具備 4 大功能，包括「提問引導」的教學模組、「過濾不當資訊」的保護機制、「紀錄學習歷程」的分析報告及「即時診斷評量」的結果呈現。
智慧觀光	交通部觀光署 AI 翻譯櫃台	透過 AI 即時翻譯資料庫打造即時翻譯方案

資料來源：本研析自行整理

附錄七、臺北市政府案例

部會名稱	計畫名稱	案例內容
【臺北市交通管制工程處】	臺北市智慧影像事件偵測建置案	1. 在建國高架道、瑞光路與陽光街口等地設置 AI 交通事件偵測器，以辨識事故、施工、壅塞等情況。 2. 系統可即時通知相關單位處理，並透過電子看板發布交通資訊，減少交通影響。 3. 提升交通事件處理效率，降低對行車的衝擊。
【臺北市交通管制工程處】	智慧號誌	1. 利用 AI 影像辨識技術，即時分析人車通行需求量。 2. 動態調整路口號誌時制，有效改善道路服務水準與提升用路人安全性。
【臺北市立聯合醫院】	血壓血氧上傳系	1. 移動式血壓計即時上傳病人生理數值，減少接觸並提供精準照護。 2. 智慧化設備將數值傳至 NIS 系統，生成趨勢圖並透過 CAS 異常警訊提醒醫療團隊掌握病情變化。
【臺北市立聯合醫院】	LDCT（低劑量肺部電腦斷層掃描）AI 影像電腦輔助判讀系統	1. 運用 LDCT 影像輔助程式達到自動化肺結節辨識、分割、分類及大小量測，並生成結構化報告，提升時效並減少人為錯誤。 2. 肺結節偵測率超過 90%，有助提升病人治療及追蹤效率。
【臺北市立聯合醫院】	數位發展部數位產業署智慧城鄉生活應用補助研析（地方試煉暨國際合作）之「智慧社區健康好食運研析」	1. AI 智慧穿戴設備：上傳長者生理數據，協助健康監測。 2. AI 護理小幫手：透過 APP 量測數據，預測慢性疾病風險並提供預防措施。 AI 營養小幫手：根據疾病史等資訊，制定個人化營養研析，改善健康與疾病預防。 3. AI 運動小幫手：設計個人化健康運動研析，降低運動風險。 4. AI 線上運動平台：提供即時互動、指導與反饋，提升長者運動意願與樂趣。

部會名稱	計畫名稱	案例內容
【臺北市立聯合醫院】	雲端醫院	1. 線上解決，避免不必要的門診。 2. AI 人工智慧輔助：病人填入檢測數值，若異常，系統提供處理建議，數據達危險值時，系統簡訊通知照護團隊。 3. 主動用藥關懷：推播就診前抽血、回診時間、檢驗報告、檢查異常報告、領慢箋服務等，提醒個案規律用藥。 4. 個案可透過手機查詢報告，若異常，系統通知回診建議。
【臺北市立聯合醫院】	決策輔助系統	1. 整合護理數據並使用視覺化和趨勢圖表分析，輔助臨床決策。 2. 橫向展開至各院區，早期發現並介入改善問題，提升照護品質。
【臺北市立聯合醫院】	護理資訊系統	1. 運用資訊化減少護理人員重覆工作的負擔。 2. 包括醫囑處理及給藥作業系統（BCMA）、安寧緩和照護需求系統、 化療給藥作業系統、入院護理評估成人（精神、中醫、新生兒）、、每日評估系統（跌倒、疼痛、傷口、體圍、意識、CAS）等。
【臺北市立聯合醫院】	AI 輔助診斷糖尿病視網膜病變	1. 協助基層診所提升糖尿病網膜病變篩檢率，以期早期發現早期治療。 2. 每年檢查至少 1500 位病人，滿意度百分之百。
【臺北市立聯合醫院】	AI 輔助診斷糖尿病 2.0 視網膜病變	1. 輔助醫師診斷糖尿病視網膜病變，透過 AI 的即時判讀增加診斷準確度 3. 內分泌新陳代謝科糖尿病衛教整合照護中心之免散瞳眼底檢查已導入 AI 以協助糖尿病視網膜病變判讀。
【臺北市翡翠水庫管理局】	翡翠水庫經營管理智慧決策系統	1. 使用人工智慧、類神經網路及大數據分析技術，建立水庫營運管理資料自動化整合機制。 2. 開發結合即時感測與氣象預報資訊的雲端運算服務，建立互動性及視覺化的管理系統。

部會名稱	計畫名稱	案例內容
		2. 該系統僅供內部使用，未提供外部民眾使用。
【臺北市翡翠水庫管理局】	翡翠水庫集水區環境影像變異監測	1. 運用人工智慧之最大相似分類法分析技術，並結合衛星影像進行環境變異判釋，即早發現崩塌地及非法土地利用，減少淤積。 3. 未提供外部民眾使用，本局內部運用。
【臺北市翡翠水庫管理局】	翡翠水庫大壩安全監測評析	1. 運用人工智慧的類神經網路技術，模擬人腦學習方式，整合 25 年擺線儀變位、上舉壓力計等相關資料（如水庫水位、環境氣溫、岩盤變位量等）。 2. 進行大數據分析及訓練，讓類神經網路學習大壩擺線儀變位及上舉壓力計的趨勢。 3. 根據當下環境，預測大壩擺線儀及上舉壓力計的模擬值，並設定適當的標準差作為預警值，達成大壩安全預警的目的。
【臺北市自來水事業處】	自來水管線汰換規劃	1. 以近 5 年修漏案件作為訓練資料來源。 2. 運用神經網路分析各項因子對管線弱點造成的影響。 3. 計算管線弱點程度，協助自來水管線汰換規劃。
【臺北市自來水事業處】	管網水理模型校正	1. 以未校正粗糙係數水理模型成果（inp 檔）與實際水壓量測值（目標值）為基礎。 2. 運用遺傳演算法求得最佳解。 3. 匯出校正完成模型，並產製報表、擬合曲線圖與統計圖表。
【臺北市自來水事業處】	智能客服系統文字查詢及申辦	1. 北水處智能客服系統自 108 年 7 月 1 日上線，分為兩階段提供服務： 2. 第 1 階段提供常用問答的文字諮詢服務。 3. 第 2 階段於 109 至 110 年間完成了「自然人過戶」等 5 項互動式服務，用戶可直接在線上申辦，無需來電或臨櫃，大幅提高了用戶的便利性與使用意願。

部會名稱	計畫名稱	案例內容
【臺北市政府資訊處】	智慧城市	1. 以 AI 人工智慧導入市政應用為主題進行智慧城市創新實證（PoC）案，並以補助廠商方式提升 PoC 案完成率及落地率。 2. 包含智慧交通、智慧安防、市政設施管理等。
【臺北市政府教育局】	臺北酷課雲「酷 AI（CooCAI）」	1. 臺北酷課雲結合 AI 技術，提供 3+2 項 AI 功能。 2. 包含 3 項幫助學生學習的「AI 學習小幫手」、「AI 口語練習」、「AI 學習加油站」及 2 項有關教師教學的「AI 輔助出題」、「AI 輔助作文批閱」。
【臺北市政府消防局】	建置災害應變雲端協作平臺	1. 建置多元資料庫，提供即時資訊推播，精準分析輔助指揮官化決策。 2. 發展指揮官 AI 決策輔助，並以視覺化數據儀表板呈現，提供即時資訊推播，精準分析輔助指揮官化決策。
【臺北市工務局大地工程處】	運用 AI 協審水土保持研析	1. 使用 AI 自然語言處理技術，分析水土保持研析格式及查核項目。 2. 除基本文書品質檢核外，亦針對重要指標進行審視。 3. 確保水土保持設施設計及研析整體品質的有效性。
【臺北市都發局】	地籍套繪都市研析使用分區圖查詢便民服務	1. 介接本府地政局地籍資料庫，自動繪圖地籍圖並套繪都市研析圖後製作圖磚上傳更新網頁提供查詢。
【臺北市建成地政事務所】	登記教主新服務─繼承登記法服諮詢	1. 透過「智能助理」軟體，運用數位科技整合不動產繼承登記等法令資訊，提供民眾更好的服務與體驗。 2. 民眾若有繼承登記程序及法令問題，可以使用電腦或手機隨時上本所網站「聊一聊」，即時得到解答，無需到地所洽詢，且可 24 小時隨時接受諮詢，不受地域限制。
【臺北市產業發展局】	AI 人工智能線上客服	1. 透過 AI 智能客服於融資主題網頁及 LINE 社群頻道提供服務。 2. 模擬真人對話方式，提供企業申請中小及青創貸款的各項申辦資訊。

部會名稱	計畫名稱	案例內容
		3. 提供資料表格查詢及各類融資相關資訊的問答服務。
【臺北市政府警察局刑事鑑識中心】	毒品資料圖像快搜	1. 在建國高架道、瑞光路與陽光街口等地設置 AI 交通事件偵測器，以辨識事故、施工、壅塞等情況。 2. 系統可即時通知相關單位處理，並透過電子看板發布交通資訊，減少交通影響。 3. 提升交通事件處理效率，降低對行車的衝擊。
【臺北大眾捷運股份有限公司】	輪椅旅客自動叫梯服務	採用 AI 影像視覺辨識方法，當系統偵測到輪椅旅客、嬰兒推車後，即觸發自動叫梯服務，以節省前述旅客等待時間。
【臺北市體育局】	臺北市競技運動訓練暨科學中心研析	1. 運用人工智慧相關運動科學儀器設備，如紅外線三維動作分析系統、視野追蹤系統及視覺化測力板系統。 2. 提供科學模擬器訓練與技戰術分析，應用於本市基層運動選手之訓練。 3. 期望能準確且有效地提升競技運動成績表現。
【臺北市政府人事處】	人事法規、人事系統操作及同仁人事業務相關權益諮詢事項	1. 試辦導入 AI 對話式人事服務，同仁若於夜間或假日期間如有人事相關問題仍有可詢問管道，建構 24 小時*7 天服務模式。 2. 服務對象：本府員工（內部使用）。
【臺北市政府地政局】	網搜防偷跑─預售建案銷售廣告檢索	1. 透過智能系統，自動查詢蒐集不動產廣告平台刊登之預售建案公開銷售廣告。 2. 同時比對內政部系統揭露之已備查建案資料，確保預售建案資訊、契約，經依法備查檢閱始公開銷售。

資料來源：本研析自行整理