

出國報告（出國類別：研究）

## 探究運用 AI 實現智慧國家之課責制度

服務機關：審計部

姓名職稱：邱仲晃審計兼科長

派赴國家：荷蘭

出國期間：113年8月22日至113年11月16日

報告日期：114年1月2日

## 摘 要

行政院自 2017 年起推動「數位國家·創新經濟發展方案(2017 至 2025 年)」(DIGI+ 方案)，選定 AI (Artificial Intelligence) 為我國下世代的發展主軸，行政院規劃以臺灣領先全球的 IC 產業優勢為基礎，打造由人才、技術、場域及產業構築而成的 AI 創新生態圈，引導臺灣成為 AI 發展重鎮，進而孕育 AI 新興產業應用發展。於 DIGI+ 方案架構下，行政院推出「臺灣 AI 行動計畫」，執行期間為 2018 年至 2021 年，嗣以「臺灣 AI 行動計畫」為基礎，並因應國際挑戰及國內情勢邁向「臺灣 AI 行動計畫 2.0」，執行期間為 2023 年至 2026 年，規劃「人才優化與擴增」、「技術深耕與產業發展」、「完善運作環境」、「提升國際影響力」及「回應人文社會議題」等 5 個主軸執行。有關臺灣 AI 行動計畫及其 2.0 之執行良窳，攸關我國國內產業發展及國際競爭力，實值審計機關予以查核。又，審計機關人少事繁，殊值借鏡國外先進國家審計機關作法，研究如何運用 AI 輔助執行審計業務，並運用創新技術與方法執行查核工作，以達事半功倍之效。

本次赴荷蘭研究該國及歐盟等有關 AI 及其課責制度發展，另參採聯合國及經濟暨合作發展組織 (OECD) 相關研究結果，據以研提建議意見。有關行政機關部分，計有：1. 及早制定我國 AI 法案，健全 AI 發展及運用法制架構；2. 加強投資 AI 及完備基礎建設並鼓勵民間積極參與，促進公私部門合作；3. 研擬歐盟 AI 法案生效後，對我國企業衝擊之因應措施；4. 秉持不遺漏任何人 (Leave no one behind) 的精神，深根教育國民及企業，因應 AI 時代來臨；5. 允宜爭取加入 AI 相關國際專業組織，以掌握 AI 國際最新發展等 5 點意見；審計機關部分，則有：1. 關注演算法查核技術方法之國際發展趨勢，以發展演算法審計模式；2. 賡續執行公平性審計，減緩 AI 落差惡性循環，促使社會永續發展；3. 研酌配合審計人員任用條例修正，加強進用具備 AI 相關專業知識人才；4. 持續與國際專業機關 (組織) 交流，汲取國際最新 AI 發展及運用新知；5. 賡續加強審計人員 AI 相關訓練，增進審計人員專業培力等 5 點意見。希有助國內臺灣 AI 行動計畫之賡續推動，並提供審計機關執行審計業務之參考。

# 目 次

第一章 前言 .....	1
第一節 研究動機與目的 .....	1
第二節 研究過程 .....	2
第二章 我國運用 AI 實現智慧國家之政策規劃及相關規範 .....	5
第一節 政策內容 .....	5
第二節 法制架構 .....	11
第三章 歐盟、經濟合作暨發展組織（OECD）及荷蘭政府發展 AI 情形及相關規範 .....	13
第一節 AI 的定義及組成內容 .....	13
第二節 AI 倫理意涵 .....	15
第三節 AI 發展及運用等監理機制及規範 .....	21
第四章 歐盟審計院、荷蘭審計院運用 AI 於審計業務概況及查核報告 .	47
第一節 歐盟審計院、荷蘭審計院查核歐盟、荷蘭政府發展及運用 AI 情形 .....	47
第二節 歐盟審計院、荷蘭審計院發展及運用 AI 於審計業務情形 . . . . .	57
第三節 公平性審計（Leave No One Behind） .....	68
第五章 結論及建議意見 .....	72
第一節 結論 .....	72
第二節 建議意見 .....	74
一、行政機關部分 .....	74
二、審計機關部分 .....	77

## 附錄

- 一、拜會荷蘭審計院訪談題綱
- 二、荷蘭中央政府會計帳戶法案 2016
- 三、訪談 Sarah Giest 教授題綱
- 四、於萊登大學公共行政管理系研討會簡報資料
- 五、聯合國教育、科學及文化組織 (UNESCO)準備狀態評估法：AI 倫理建議的工具
- 六、聯合國教育、科學及文化組織 (UNESCO)倫理影響評估法：AI 倫理建議的工具
- 七、歐洲審計院 (ECA) 部署 AI 的一般風險描述、影響及其因應對策
- 八、歐洲審計院 (ECA) 部署 AI 的業務風險描述、影響及其因應對策

## 參考文獻

## 圖 次

圖 1	與荷蘭審計院訪談者合照.....	3
圖 2	與 Sarah Giest 教授合照.....	4
圖 3	研討會簡報實況.....	5
圖 4	研討會與會者合照.....	5
圖 5	臺灣 AI 行動計畫推動架構.....	6
圖 6	臺灣 AI 行動計畫 2.0 推動架構.....	8
圖 7	臺灣 AI 行動計畫臺灣 AI 行動計畫 2.0 差異比較 .....	10
圖 8	AI 的組成內容 .....	14
圖 9	OECD 根據傷害嚴重程度對 AI 事件和危害提出的分類 .....	18
圖 10	歐盟 AI 創新計畫運用領域.....	26
圖 11	OECD AI 事件監測機制 .....	36
圖 12	AI 運用於審計業務之分析面向 .....	47
圖 13	荷蘭中央機關發展 AI 遇到之障礙態樣.....	52
圖 14	歐盟 SAI AI 推行階段自我評估結果.....	58
圖 15	歐盟 SAI 正使用或規劃使用特定 AI 技術情形.....	58
圖 16	SAI 推行 AI 遭遇挑戰的類型 .....	58

## 表 次

表 1	臺灣 AI 行動計畫主軸及其部會分工.....	7
表 2	臺灣 AI 行動計畫 2.0 主軸及其部會分工.....	9
表 3	UNESCO 所提以人權為中心之 AI 倫理 10 項核心原則及其內涵 .....	15
表 4	荷蘭 AI 策略行動計畫路徑及其內涵.....	27
表 5	AI 國際規範 12 項重要發展歷程 .....	33
表 6	OECD AI 5 項原則及其建議 .....	35
表 7	查核 AI 需要考慮的要素.....	48
表 8	ECA AI 溝通計畫 .....	62
表 9	增強提供的數據服務項目.....	63
表 10	與國際機構合作帶來的預期益處.....	66
表 11	雲端與地端比較.....	67



# 探究運用 AI 實現智慧國家之課責制度

## 第一章 前言

### 第一節 研究動機與目的

行政院自 2017 年<sup>1</sup>起推動「數位國家・創新經濟發展方案(2017 至 2025 年)」(DIGI+ 方案)，作為引領數位發展、帶動創新的施政藍圖，期加速我國產業及生活融入 AI、物聯網、大數據等智慧科技，同時發揮臺灣小而精、跨域整合快的優勢，讓臺灣成為智慧創新的典範國度。在知識經濟的時代，科技創新是帶動經濟成長和國家進步的主要動力，特別是 AI (Artificial Intelligence, AI) 科技正在改變全球的產業發展，成為銳不可擋的重要趨勢。臺灣為與世界科技發展脈動同步，亦已選定 AI 為我國下世代的發展主軸，行政院規劃以臺灣領先全球的 IC 產業優勢為基礎，打造由人才、技術、場域及產業構築而成的 AI 創新生態圈，引導臺灣成為 AI 發展重鎮，進而孕育 AI 新興產業應用發展，於 DIGI+ 方案架構下，行政院推出了「臺灣 AI 行動計畫」，執行期間為 2018 年至 2021 年，期藉由「AI 人才衝刺」、「AI 領航推動」、「建構國際 AI 創新樞紐」、「場域與法規開放」、「產業 AI 化」等五項重點工作，讓臺灣在下一波的智慧革命中取得機會與優勢。嗣以「臺灣 AI 行動計畫」為基礎，並因應國際挑戰及國內情勢邁向「臺灣 AI 行動計畫 2.0」，執行期間為 2023 年至 2026 年，規劃以「人才優化與擴增」、「技術深耕與產業發展」、「完善運作環境」、「提升國際影響力」、以及「回應人文社會議題」等五個主軸執行。有關臺灣 AI 行動計畫及其 2.0 之執行良窳，攸關我國國內產業發展及國際競爭力，實值審計機關予以查核，惟囿於 AI 發展屬新興議題且方興未艾，如何查核行政部門執行 AI 計畫，有待汲取國外先進國家作法及國際專業組織研究，作為本部執行審計業務參考。

又，審計機關人少事繁，殊值借鏡世界先進大國審計機關作法，研究如何運用 AI 輔助執行審計業務，並以創新技術與方法執行查核工作，已達事半功倍之效，為符合世界審計業務潮流，亦有赴國外研究先進國家審計機關作法之必要。

依據全球負責任 AI 指數 (Global Index Responsible AI, GIRAI) 評估 138 個國家 AI 發展及運用情形，蒐集超過 200 萬筆數據後予以分析，於 2024 年公布調查結果，荷

---

<sup>1</sup> 為統一國內、外年份表達用語，本研究均以西元年表示之。



蘭綜合成績為世界第一<sup>2</sup>，為有效汲取國外先進國家審計機關有關查核行政部門發展 AI 及本身發展、運用 AI 等經驗，經擇選荷蘭為研究國家，並藉地利之便，蒐集歐盟、經濟合作暨發展組織（OECD）及荷蘭之相關資料，以為我國參考。

## 第二節 研究過程

經奉准出國進行專題研究後，即對上開議題蒐集相關資訊，並洽荷蘭學術研究機構，嗣經荷蘭萊登大學（Leiden University）公共行政管理系（Institute of Public Administration, Leiden University）系主任 Bernard Steunenbergh 博士同意，以訪問研究者（Visiting Researcher）身份，由 Hsini Huang（黃心怡）博士擔任指導教授，於該校從事本次專題研究。謹臚列研究方式如下：

### 一、資料蒐集

AI 發展及其課責制度研究於近年方興未艾，國際組織（聯合國、OECD、國際最高審計機關組織、歐盟等）及各國最高審計機關均投入相當人力及經費，致力於研擬發展、監管措施，期使人類於享受 AI 帶來的好處之外，也掌控 AI 衍生的風險。本次專題研究先蒐集我國政府發展 AI 政策，並廣泛蒐集及研閱荷蘭政府、荷蘭審計院及歐盟、歐盟審計院、聯合國、OECD 等國際組織等發布相關研究（查核）報告，以為研究成果奠定基礎。

### 二、赴荷蘭審計院（Netherlands Court of Audit, NCA）<sup>3</sup>訪談

萊登大學為理論結合實務，聘請荷蘭審計院策略顧問 Sjoerd Keulen 博士，於該校教授審計、公共政策評估等課程。為瞭解 NCA 如何審核荷蘭政府推行 AI 情形及其技術與方法、NCA 發展及運用 AI 情形等，以電子郵件邀請 Sjoerd Keulen 博士於 2024 年 9 月 16 日下午會面，並協請洽拜會 NCA 事宜，嗣於 2024 年 10 月 21 日偕同指導教授 Hsini Huang 博士赴該院進行 1 小時半訪談，訪談題綱如附錄 1。NCA 訪談參與者，係 Ruud Wissenburg 先生（負責 IT 審計，多年從事 IT 相關工作）、Yara van der Laan 女士（AI 學程碩士，IT 審計專案經理人）及 Colin van der Noordt 先生（博士，AI 專家，負責

---

<sup>2</sup> 荷蘭總分為 86.16，我國總分為 33.86，於 138 個國家中排名第 34。

<sup>3</sup> 有關荷蘭審計院任務、權力及法律定位，規範於該國「中央政府會計帳戶法案 2016（Government Account Act 2016）」，詳附錄 2。

審核荷蘭政府使用 AI 及其演算法，評估其是否以負責任的態度使用），合照詳圖 1。

NCA 訪談參與者分享 NCA 目前運用 AI 情形、審核政府 AI 演算法情形、審計人員訓練機制與近 3 年國際交流 AI 議題情形等，茲摘陳如次：

(一) 荷蘭中央政府目前考量使用規範、資安規定及結果可能產生偏差等因素，尚未使用生成式 AI 於業務上。

圖 1 與荷蘭審計院訪談者合照

(二) NCA 目前尚未有 AI 發展整體策略，惟刻正探索運用 AI 於審計工作的可行性，規劃於 2025 年啟動生成式 AI 的先導運用，於網路安全、大數據處理、數據挖掘、撰寫摘要、將之前紙本資料轉換為數位化後進行全面搜尋 (holistic search) 等，未來將視成效再據以決定整體策略內容。



(三) NCA 認為荷蘭行政部門使用之 AI 演算法，其良窳是行政機關的責任，最終目標是行政機關必須對 NCA 證明演算法是妥適的、是否有存偏差 (bias)。演算法第三方驗證雖不失為一種機制，惟仍須克服第三方機構的專業性、是否願意對外提供其認證標準、程序、範圍等資訊。

(四) NCA 數位人才隸屬於各業務單位，執行審計工作尚需要數位人才協助，則以任務編組方式進行。近年來 NCA 預算充裕，為因應發展 AI 及查核演算法，招募一些具有 AI 背景人才進入 NCA 服務。此外，NCA 訓練單位建置於人力資源部門，並未有數位人才培訓計畫，惟單位主管認為有訓練需求，則可以將需求提供給該部門據以開課；NCA 有多種訓練管道，例如：NCA 內部訓練實體課程、線上學習課程、荷蘭政府開設課程、鼓勵員工到學校進修等。

(五) NCA 重視與國際組織、世界最高審計機關之交流，經常透過實體、視訊會議、電話或是電子郵件相互交流審計經驗，對話重點集中於如何審核 AI 及其演算法，大多仍處理探索階段，尚未形成共識，也難以掌握 AI 審計未來發展趨勢，惟交流日益頻繁熱絡，也有部分國家調整 NCA 發布之演算法審計框架後使用於查核工作。NCA 進一步表示願與本部持續保持聯繫，雙方亦可透過電子郵件往來、視訊會議，交流 AI 審計經驗。

### 三、訪談萊登大學學者

為瞭解發展及運用 AI，對特定族群產生不利影響及衍生公平性審計等議題，洽詢萊登大學公共管理學院相關研究專長教授予以訪談。萊登大學公共管理學院 Sarah Giest 教授，專注於創新和永續發展，以及如何使用政策視角研究塑造永續社會的技術、環境和社會解決方案。Sarah Giest 教授取得加拿大西門菲沙大學（Simon Fraser University）博士，目前擔任國際公共政策協會（IPPA）的副主席以及「數據與政策（Data & Policy）」和「政策設計與實踐（Policy Design and Practice）」等編輯委員會成員，並擔任荷蘭部會、OECD、歐盟執行委員會（European Commission）及聯合國等外部專家。

經洽 Sarah Giest 教授於 2024 年 10 月 24 日中午 12 點至 12 點半，進行半小時訪談，就聯合國一直倡議不遺漏任何人（Leave No One Beyond），政府於發展 AI 同時，應如何保護弱勢團體、審計機關在執行公平性審計時應注意哪些面向（關鍵要素）等議題就教（訪談題綱如附錄 3）。Sarah Giest 教授表示，政府可以考慮對於學齡兒童從小施予 AI 教育，從基礎教育著手，使大家有平等機會接觸，減少未來競爭力的差距；對於目前低技術的勞工階層，則可透過在職訓練，輔導成為 AI 基本操作員，不必塑造每個人都成為 AI 的編碼員；此外，審計機關在執行公平性審計時，允宜注意政府 AI 系統使用數據資料時，該資料內涵之偏差造成決策失誤，以及由於 AI 是依據投入數據資料，經過演算法產生通常是 0 與 1 兩個絕對的決策，但是福利發放等行政行為，有時不是如此絕對，尚須考量個案特殊因素及其配套措施，所以政府發展 AI 時，是否將人性因素納入考量（Human-in-the-loop）<sup>4</sup>，亦係審計機關得以查核重點之一。

圖 2 與 Sarah Giest 教授合照



<sup>4</sup> Human-in-the-loop（人在環內，或譯為：人在迴路、人機迴圈、人機共生），主要意涵是機器系統有從收集資料、判讀情勢到做出決策的迴圈，而人必須參與其中，衍生發展 AI 時，應考量人性因素。

## 四、於萊登大學公共行政管理系研討會進行簡報

萊登大學公共行政管理系於開學期間約每個月會舉辦一場研討會，針對數位治理、永續發展等特定公共議題，由系上老師或博士研究生進行專題簡報。2024 年 9 月開學後第一場研討會於 9 月 16 日中午舉行，經受邀於該研討會介紹「臺灣 AI 及其課責制度發展現況」。經搜尋我國行政院、行政院智慧國家推動小組、臺灣 AI 卓越中心等全球資訊網及本部內網有關案例分享資料，予以彙整，簡報內容包括我國政府體制、中央政府預算課責制度、本部於中央政府體制所扮演角色、本部組織職掌、近年來審計業務發展重點，我國 AI 發展政策概況及本部審核情形、本部發展 AuditGPT 及運用 ChatGPT 情形之實際案例等（簡報詳附錄 4），歷時約 1 小時。與會者針對本部發展 AuditGPT、運用 ChatGPT 情形及實際案例，深感興趣，詢問我國政府如何整合公部門、私部門及民間團體人才以發展 AI，本部是否鼓勵審計人員運用 ChatGPT 協助審計工作、本部進行專家諮詢及強化公民參與的做法，並建議本部使用 ChatGPT 協助審計工作，允宜注意隱私權保障及結果產生偏差的可能性（會議照片，圖 3、圖 4）。

圖 4 研討會簡報實況



圖 3 研討會與會者合照



## 第二章 我國運用 AI 實現智慧國家之政策規劃及相關規範

### 第一節 政策內容

#### 一、臺灣 AI 行動計畫

##### （一）計畫緣由

為促進數位經濟創新發展、提高國人生活品質，邁向「智慧國家」，並帶動 5+2 產

業創新及加值應用，行政院自 2017 年起推動「數位國家・創新經濟發展方案（2017 至 2025 年）」(DIGI+方案)，作為引領數位發展、帶動創新的施政藍圖，期加速我國產業及生活融入 AI、物聯網、大數據等智慧科技，同時發揮臺灣小而精、跨域整合快的優勢，讓臺灣成為智慧創新的典範國度。「DIGI+方案」中的 D 是指 Development（發展），發展堅固基磐；I 是指 Innovation（創新），創新數位經濟；G 指 Governance（治理），治理智慧國家；最後一個 I 則是指 Inclusion（包容），包涵容納公民社會。至最後的加號，是 plus、也是 upgrade（升級），希望臺灣在推動本方案後，國家數位基礎建設、經濟與社會創新各個層面均能夠向上提升。

而 AI 無疑將是下一波智慧革命的重要關鍵，因此，行政院推出了「臺灣 AI 行動計畫」，執行期間為 2018 年至 2021 年，以「創新體驗為先、軟硬攜手發展、激發產業最大動能」為願景，以法規鬆綁、場域及資料開放，以及加速投資動能的基本思維，藉由「AI 人才衝刺」、「AI 領航推動」、「建構國際 AI 創新樞紐」、「場域與法規開放」、「產業 AI 化」等五項重點工作，讓臺灣在下一波的智慧革命中取得機會與優勢。

## （二）願景

臺灣 AI 行動計畫將以實現「創新體驗為先，軟硬攜手發展，激發產業最大動能」為願景，從需求端出發，發展應用導向的 AI 前瞻技術，強化軟、硬體整合的系統技術，並提供科技創新所需的環境建構，包括育才及留才的環境、相關法規的調適、公共資料及場域的開放、技術研發、產業聚落等面向的統合推動，以激發產業最大的動能，創造我國智慧科技發展的經濟榮景。

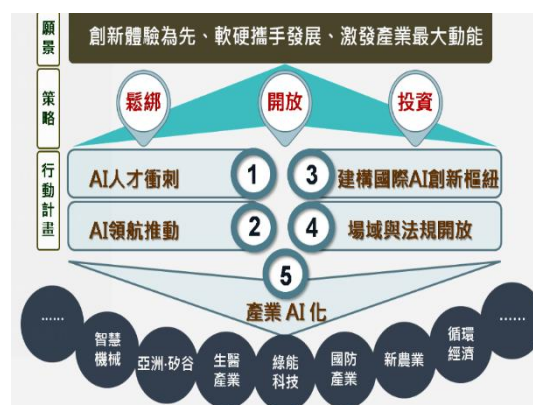
## （三）推動架構

在落實推動上，規劃以 AI 人才衝刺、AI 領航推動、建構國際 AI 創新樞紐、場域與法規開放、以及產業 AI 化等五個主軸計畫來實現，並與 5+2 產業創新方案扣合，搭配前瞻基礎建設及智慧城鄉計畫，共同推動普及智慧科技應用服務（整體推動架構，詳圖 5）。

## （四）主軸行動計畫

臺灣 AI 行動計畫以法規鬆綁、場域與資料開放、加速投資動能的基本思維，規劃 AI 人才衝刺、AI 領航推動、場域與法規開放、建構國際 AI 創新樞紐及產業 AI 化等五

圖 5 臺灣 AI 行動計畫推動架構



項行動主軸；據此整合政府相關計畫及產業資源，鏈結國際夥伴能量，希望建立 5+2 產業 AI 創新的完整布局，並期塑造臺灣成為全球智慧科技創新重要 樞紐。本行動計畫在各部會分工上，科技會報辦公室將負責全國整體科技發展與政策的盤點、分工、推動、與協調；各主軸之共同推動部會，在 AI 人才衝刺主軸有 科技部、經濟部、教育部、勞動部等；AI 領航推動有科技會報辦公室、經濟部、科技部、資安處、教育部等；建構國際 AI 創新樞紐為經濟部及科技部共同推動；場域與法規開放有經濟部、科技部、環保署、交通部、內政部等，法規調適包括國發會、科技會報辦公室及相關部會共同研析；產業 AI 化為經濟部、科技部、國發會、農委會、教育部、勞動部等共同推動（有關臺灣 AI 行動計畫主軸及其部會分工，詳表 1）。

表1 臺灣 AI 行動計畫主軸及其部會分工

行動計畫主軸	子項名稱	相關部會
AI人才衝刺	智慧科技菁英	科技部、教育部、經濟部
	智慧應用先鋒	經濟部、教育部、科技部、勞動部
	吸引全球AI人才	經濟部、科技部
AI領航推動	聚焦研究主題	科技會報辦公室、經濟部
	發展國家級AI前瞻研究網絡	科技部、經濟部、資安處、教育部
建構國際AI創新樞紐	扶植百家AI新創事業	經濟部、科技部
	發展國際級AI創新聚落	經濟部、科技部
場域與法規開放	實證場域與資料開放	經濟部、科技部、環保署、交通部、內政部
	AI 相關法規議題研析	國發會、科技會報辦公室、各部會
產業AI化	鏈結5+2產業創新與AI人才媒合	經濟部、科技部、國發會、農委會、教育部、勞動部
	完善產業AI化環境，帶動中小企業AI創新	經濟部、科技部

## 二、臺灣 AI 行動計畫 2.0

### （一）概述

「臺灣 AI 行動計畫」從「AI 人才衝刺」、「AI 領航推動」、「建構國際 AI 創新樞紐」、「場域與法規開放」及「產業 AI 化」等五大主軸推動，自 2018 至 2021 年在相關部會執行之下，獲得階段性成果。AI 仍在快速演進，且受全球政經局勢影響已成國家戰略性科技，先進國家持續投入資源 發展 AI 以鞏固在全球 AI 地位，我國亦以「臺灣 AI 行動計畫」為基礎，並因應國際挑戰及國內情勢邁向「臺灣 AI 行動計畫 2.0」，執行期間為 2023 年至 2026 年。

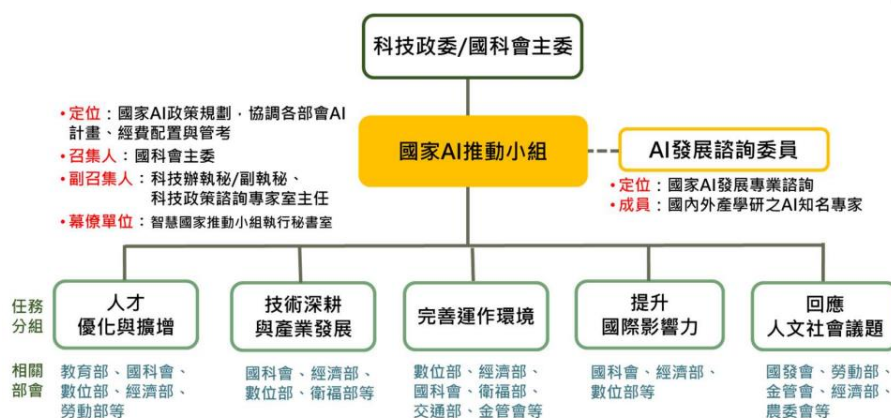
### （二）願景

臺灣 AI 行動方案 2.0 將以實現「以 AI 帶動產業轉型升級、以 AI 協助增進社會福祉、讓臺灣成為全球 AI 新銳」為願景。從產業端出發，透過深耕 AI 技術與發展 AI 產業及產業應用 AI，帶動我國整體產業轉型升級。並建構兼顧科技創新及風險治理的可信任 AI 發展環境，包括人才優化及留才攬才、重視 AI 倫理法制、推動資料治理及流通。再回歸到社會端，因應 AI 衍生的各項社會衝擊，並以 AI 科技發展具包容性的數位經濟，協助解決社會面臨重大挑戰，增進全民福祉。由此，從人才、技術、產業等厚植臺灣 AI 國力，公私協力提升臺灣 AI 在國際影響力，讓臺灣成為全球 AI 新銳。

### (三) 推動架構

將透過國家 AI 推動小組來統籌整體推動作業，包括國家 AI 政策規劃，協調各部會 AI 計畫及經費配置與管考等。另設立 AI 發展諮詢委員，由國內外產學研之 AI 知名專家組成，提供國家 AI 發展專業諮詢。在落實推動上，規劃以「人才優化與擴增」、「技術深耕與產業發展」、「完善運作環境」、「提升國際影響力」、「回應人文社會議題」等五個主軸任務來實現，並與 5+2 產業創新方案扣合，搭配前瞻基礎建設計畫，共同推動普及智慧科技應用服務（推動架構，詳圖 6）。

圖 6 臺灣 AI 行動計畫 2.0 推動架構



### (四) 主軸行動計畫

為能因應 AI 發展所帶來的效益與衝擊，我國 AI 政策於規劃時應同時予以兼顧，期望一方面善用 AI 科技帶動經濟繁榮發展，另一方面則運用 AI 科技促進邁向包容、安全、信任、公平的社會，實踐創新、包容、永續的智慧國家。本行動計畫在各部會分工上，國科會科技辦公室<sup>5</sup>負責統籌與協調。各主軸之共同推動部會，在人才優化與擴增主軸有教育部、國科會、數位部、經濟部、勞動部等；技術深耕與產業發展有國科會、經濟部、數位部、衛福部等；完善運作環境有數位部、經濟部、國科會、衛福部、交通部、金管會等；提升國際影響力有國科會、經濟部、數位部、教育部、外交部等，回應人文社會議題包括國發會、勞動部、金管會、經濟部、農委會等（有關臺灣 AI 行動計畫 2.0 主軸及其部會分工，詳表 2）。

5 配合行政院組織調整，行政院科技會報辦公室於 2022 年 7 月 27 日改制為國家科學及技術委員會科技辦公室，簡稱國科會科技辦公室。

表 2 臺灣 AI 行動計畫 2.0 主軸及其部會分工

行動計畫主軸	子項名稱	相關部會
人才優化與擴增	高等教育	國科會、教育部
	國民教育	教育部、數位部
	在職/就業培訓	數位部、經濟部、勞動部
技術深耕與產業發展群	布局 AI 軟體與片核心技術	國科會、數位部、經濟部
	加速 AI 軟硬體產業發展	數位部、經濟部
	優勢產業應用 AI，成為國際領先	製造業：經濟部 醫療業：衛福部、經濟部
	強化中小企業導入 AI 轉型升級	數位部
完善運作環境	推動資料治理，促進資料流通	數位部、國科會(科研資料)，相關部會協助資料跨域流通
	AI 法制推動	通用領域：國科會、數位部 個別領域：衛福部、交通部、金管會等
	成立 AI 產品/系統評測中心，推動與國際介接的 AI 規範與標準	數位部主責，相關部會協助
	易於取得高效能運算資源	國科會、交通部
提升國際影響力	參與國際 AI 相關組織	國科會、數位部
	推動 AI 領域國際合作	國科會、經濟部
	以臺灣 AI 能量貢獻國際社會	國科會、經濟部
回應人文社會議題	研析 AI 對社會衝擊及因應準備	國發會、勞動部、金管會
	以 AI 協助解決國家社會面臨挑戰	農委會、經濟部、相關部會

### (五) 臺灣 AI 行動計畫與臺灣 AI 行動計畫 2.0 之差異

「臺灣 AI 行動計畫」從「AI 人才衝刺」、「AI 領航推動」、「建構國際 AI 創新樞紐」、「場域與法規開放」及「產業 AI 化」等五大主軸推動，自 2018 至 2021 年在相關部會執行之下，獲得階段性成果。AI 仍在快速演進，且受全球政經局勢影響已成國家戰略性科技，先進國家持續投入資源發展 AI 以鞏固在全球 AI 地位，我國亦以「臺灣 AI 行動計畫」為基礎，並因應國際挑戰及國內情勢邁向「臺灣 AI 行動計畫 2.0」，以下就「人才優化與擴增」、「技術深耕與產業發展」、「提升國際影響力」、「完善運作環境」及「回應人文社會議題」等五大主軸說明之。



首先，為厚植臺灣 AI 實力，「臺灣 AI 行動計畫 2.0」將持續培育人才及發展技術與產業。人才培育可謂臺灣發展 AI 之核心，將以前期「AI 人才衝刺」為基礎持續精進，主要包括提升教學品質以優化人才能力，並配合企業對 AI 專業人才及跨領域人才之殷切需求，擴大培育規模，此即「人才優化與擴增」。在技術與產業方面，我國將 AI 視為重點科技之投入期間尚短，需持續挹注資源助其從萌芽邁入成長，將整合前期「AI 領航推動」及「產業 AI 化」成果，持續推動「技術深耕與產業發展」。除深度布局 AI 軟體／硬體（晶片）前瞻技術，並強化產學研合作及新創培育以促進相關產業成長，另在國內大型企業多已運用 AI 提升運作效率或創新產品服務之下，後續將以協助資源相對不足的中小企業導入 AI 為主。再者，國際合作與法規環境仍為「臺灣 AI 行動計畫 2.0」兩個重要推動構面，但重點有所調整。國際合作方面，臺灣 AI 行動計畫「建構國際創新樞紐」主要目的為吸引國際大廠來臺研發，將國際資源拉進來以提升臺灣 AI 創新發展能量，在 Google、微軟等相繼來臺、達成階段性目標後，後續將憑藉臺灣 AI 行動計畫推動經驗與成果走出去，積極參與國際組織，與全球 AI 領先國家建立合作關係，以「提升臺灣在國際 AI 影響力」。法規環境方面，新興科技發展初期以開放、鼓勵創新為原則，故臺灣 AI 行動計畫推動「場域與法規開放」，如建構自駕車測試場域及訂定無人載具科技創新實驗條例等。而隨著 AI 應用擴大，其引發風險如假訊息、隱私侵害、偏見歧視及安全性等已然浮現，先進國家將 AI 倫理及法制等納入 AI 政策，臺灣 AI 行動計畫 2.0 亦與國際潮流對齊，重視 AI 倫理法制，推動與國際介接的規範與標準，以期「完善可信任 AI 運作環境」。最後，有鑑於 AI 應用範疇日趨多元且逐步深化，對臺灣社會之影響無論善惡皆是日益明顯，故臺灣 AI 行動計畫 2.0 新增「回應人文社會議題」主軸，重點包括關注及研析 AI 對社會造成的衝擊(如工作變遷等)，以利研擬因應對策，另一方面將善用 AI 科技以協助解決如勞動力短缺、超高齡社會、淨零排放等社會面臨重大挑戰，讓全民受益於 AI (臺灣 AI 行動計畫臺灣 AI 行動計畫 2.0 差異比較，如圖 7)。

圖 7 臺灣 AI 行動計畫臺灣 AI 行動計畫 2.0 差異比較



### 三、智能革新數位領航計畫

預計執行期間為 2026 年到 2030 年，目前刻正由數位部規劃研擬執行內容。

## 第二節 法制架構

有感 AI 科技研發可能帶來的創新、優勢與衝擊，前科技部（已於 2022 年改制為國家科學及技術委員會），於 2009 年發布 AI 科研發展指引<sup>6</sup>，完善我國 AI 科研發展環境，強調「以人為本」（AI 科研應遵循以人為本之價值，以提升人類生活、增進人類福祉為宗旨，構築尊重人性尊嚴、自由與基本人權的 AI 社會）、「永續發展」（AI 科研應追求經濟成長、社會進步與環境保護間之利益平衡，以人類、社會、環境間的共存共榮為目標，創造永續願景）及「多元包容」（AI 科研應以創建及包容多元價值觀與背景之 AI 社會為發展目標，並且積極啟動跨領域對話機制，普惠全民對 AI 的理解與認知。）三大核心價值，從而延伸出八項指引，包括「共榮共利」、「公平性與非歧視性」、「自主權與控制權」、「安全性」、「個人隱私與數據治理」、「透明性與可追溯性」、「可解釋性」及「問責與溝通」等，提供我國 AI 科研人員在學術自由及研究創新發展前提下可依循的方向，期望 AI 科研人員能善盡能力，適時注意與利害關係人的互動對話，也有助於研究成果被接納及正確的擴散運用，開創符合普世價值、安全的 AI 社會。

國家發展委員會於 2018 年提出「AI 之相關法規國際發展趨勢與因應」報告，針對「AI 對於著作權法制之適用疑義」、「AI 時代下的個資保護與合理利用」、「智慧載具發展及侵權責任歸屬」、「AI 及大數據於公權力行政之運用」、「AI 與金融監理沙盒之應用」、「AI 運用於醫療服務之相關法律疑義」等六大議題提出法規調適及因應建議。

行政院於 2023 年 4 月成立「數位政策法制協調專案會議」，由 3 位政務委員（科技政務委員、法制政務委員、政務委員兼國發會主委）共同主持，跨部會專案蒐整、研析及協調處理數位法制議題，並依議題特性分為三個分組，分別為個人資料保護分組（個人資料保護委員會籌備處負責）、資料創新法制分組（數位發展部負責）以及 AI 法制分組（國科會負責）。其中 AI 法制分組主要任務：（一）關注國際及區域性 AI 法制動態對我國之影響，提出 AI 法制框架。（二）蒐整各部會法制議題，督促部會依照框架法規調適。（三）檢視各部會法制文件，確保法制政策一致性及完整性。<sup>6</sup>

隨著生成式 AI 爆發，相關應用發展涉及各行各業與民生社會，國家整體佈局仍滾動調適中。考量生成式 AI 變革速度快速，「數位政策法制協調專案會議」督促各部會密切關注國際發展，並就其業管範圍提出法規調適規劃，除了修訂法規以外，也須視需求先行釋出公部門指引。自 2023 年 4 月至今，AI 法制分組已透過跨部會專案會議督促各

---

<sup>6</sup> 資料來源：國家科學及技術委員會（2024），「AI（AI）推動現況與未來方向」專題報告。

部會依照法制框架提出各種法規或指引：

## （一）法規：

AI 基本法作為 AI 發展的基本行政方針隨著 AI 應用範圍日漸擴大，國際社會對於 AI 快速發展可能產生的衝擊與影響相當關切。歐盟 AI 法案（AI Act，下稱 AI 法案）經理事會於 2024 年 5 月 21 日通過，在公告 20 日後將分階段施行。該法案監管 AI 的開發、進入市場、提供服務與使用，認為應依不同風險等級，給予不同的義務，同時監管與裁罰併行，才能確保技術快速變化下符合歐盟基本權利憲章規定之健康、安全與基本權保障。由於其罰鍰金額可最高到 3,500 萬歐元或其全球收入之 7%，故各國與國際企業皆密切關注該法案後續實際執行情形。美國、英國、日本則未訂有專法，多由行政部門訂定指引方針供民眾參考。例如 2023 年 10 月美國發布安全與值得信賴 AI 行政命令（Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence），訂定優先議題及推動事項，要求各行政機關於期限內訂定指引或調查報告。透過全面性的行政命令，引領各領域因應 AI 風險，包括制定 AI 安全標準、隱私保護、促進公平與公民權利、維護消費者權益、支援勞工、促進創新與競爭等具體行動目標。我國發展 AI 亦應衡平創新發展與可能風險。為確立我國推動發展之方向與作法，建構可信任的運作環境，國科會已於 2024 年 7 月 15 日預告 AI 基本法草案，並至 9 月 13 日止，持續蒐集各界意見，該草案揭示永續發展、人類自主、隱私保護、資安與安全、透明可解釋、公平不歧視及問責等 7 大基本原則，以及創新合作及人才培育、風險管理及應用負責、權益保障及資料利用、法規調適及業務檢視之 4 大推動重點，做為引導我國各機關發展與促進 AI 應用之原則。

## （二）指引：

1. 一般性規範：行政院於 2023 年 10 月 3 日函頒「行政院及所屬機關(構)使用生成式 AI 參考指引」，規範行政院及所屬機關（構）秉持負責任及可信賴的態度運用生成式 AI，具政策宣示效應及示範作用，可引導各界養成對生成式 AI 的正確觀念，降低可能帶來的風險。
2. 個別領域規範：
  - (1) 金管會於 2023 年 10 月公布「金融業運用人工智慧（AI）之核心原則與相關推動政策」，引導金融業在兼顧消費者權益、金融市場秩序及社會責任下，投入科技創新，促進金融服務升級。金管會於 2024 年 6 月發布「金融業運用 AI 指引」，做為金融機構導入、使用及管理 AI 的參考，計分總則及 6 大章節，其中

總則主要說明 AI 相關定義、AI 生命週期、風險評估考量因素、以風險為基礎落實核心原則的方式、第三方業者的監督管理等共通事項；6 大章節則分別說明金融業在落實 6 項核心原則時，依 AI 生命週期及所評估的風險，宜關注的重點以及可採行的措施，包括目的、主要概念，以及各原則相應的注意事項、落實方式或採行措施等。<sup>7</sup>

- (2) 交通部於 2024 年 1 月發布「自駕公車實驗運行安全指引」草案，確保即將上路運行的自駕公車行車安全。數位發展部於 2024 年 3 月預告「AI 產品與系統評測參考指引」草案，針對語言模型建立初步評測項目，確保語言模型的可解釋性、公平、準確、透明等 10 項要求。
- (3) 國科會後續將持續追蹤各部會的法規修訂進度，並依照「數位政策法制協調專案會議」2024 年 5 月 3 日會議決議，要求重點法制議題之主管部會，包括：經濟部（AI 智慧財產權）、勞動部（AI 影響下勞動權益保護）、交通部（自駕車上路的道安、車安法規）、衛生福利部（醫療器材），優先調整相關規範。

## 第三章 歐盟、經濟合作暨發展組織（OECD）及荷蘭政府發展 AI 情形及相關規範

### 第一節 AI 的定義及組成內容

#### 一、AI 的定義

有關 OECD 與歐盟對 AI 的定義，僅差異於用字遣詞及語句編排方面，兩者在意涵上大致一致。

##### （一）OECD 對 AI 系統定義：

一個基於機器的系統，為了明顯或隱喻目的，根據接收到的資料中，推斷如何生成如預測、內容、建議或可能影響實體或虛擬環境的決策等輸出。不同的 AI 於部署後的

---

<sup>7</sup> 資料來源：金管會全球資訊網，[https://www.fsc.gov.tw/ch/home.jsp?id=96&parentpath=0,2&mcustomize=news\\_view.jsp&dataserno=202406200001&dttable=News](https://www.fsc.gov.tw/ch/home.jsp?id=96&parentpath=0,2&mcustomize=news_view.jsp&dataserno=202406200001&dttable=News)

自主性及適應性程度各不相同。<sup>8</sup>

## （二）歐盟對 AI 定義：

AI 是指基於機器的系統，其設計為以不同程度的自主性運行，並且在部署後可能表現出適應性，並且對於明確或隱含的目標，根據其收到的輸入推斷如何產生可以影響物理或虛擬環境的輸出，例如預測、內容、建議或決策。<sup>9</sup>

## （三）荷蘭對 AI 定義：

係遵循歐盟上述對 AI 的定義。

## 二、AI 的組成內容

一般認為 AI 由智能聊天機器人(Intelligent Chatbot)、機器流程自動化(Robotic Process Automation, RPA)、區塊鏈技術(Blockchain)、自然語言生成(Natural Language Generation)、雲端運算(Cloud Computing)、預測建模(Predictive Modeling)及智能內容識別(Intelligent Content Recognition)等 7 部分(詳圖 8)。

### （一）智能聊天機器人

(Intelligent Chatbot)：透過自然語言處理技術，智能聊天機器人能夠模擬人類對話，提供自動化的客服和信息回應。

圖 8 AI 的組成內容



（二）機器流程自動化 (Robotic Process Automation, RPA)：RPA 使用軟體機器人來自動執行基於規則的業務流程，減少重複性任務和人為錯誤。

（三）智能內容識別 (Intelligent Content Recognition)：這種技術能夠自動檢測和

<sup>8</sup> 資料來源：OECD (2024), EXPLANATORY MEMORANDUM ON THE UPDATED OECD DEFINITION OF AN AI SYSTEM.

<sup>9</sup> 資料來源：歐盟 AI 法案 (2024)。

分類不同類型的數字內容，如圖像、文件或語音。

(四) 區塊鏈技術 (Blockchain): 區塊鏈提供去中心化的安全技術，尤其在資料的驗證、保護和追蹤方面應用廣泛，為 AI 系統增添透明性和安全性。

(五) 自然語言生成 (Natural Language Generation): 這是一種 AI 技術，可以將數據轉換為可讀的語言文本，用於自動生成報告或解釋複雜的數據集。

(六) 雲端運算 (Cloud Computing): AI 系統需要大量的運算資源，雲端運算提供可擴展的計算能力，使 AI 模型可以在多個設備上運行。

(七) 預測模型 (Predictive Modeling): AI 透過分析歷史數據來建立模型，預測未來的趨勢或結果，在決策支持和風險管理中發揮重要作用。

## 第二節 AI 倫理意涵

### 一、AI 倫理概述

聯合國教育、科學及文化組織 (UNESCO) 於 2021 年發布「AI 倫理問題建議書」，提出以人權為中心之 AI 倫理 10 項核心原則，詳表 3。

表 3 UNESCO 所提以人權為中心之 AI 倫理 10 項核心原則及其內涵

原則	內涵
相稱性及不損害	AI 的使用不得超出實現合法目標必要的範圍。應運用風險評估以防止此類用途可能造成的危害。
安全及保護	AI 參與者應避免並應對可能的傷害 (安全風險) 以及潛在的攻擊漏洞 (安全保護風險)。
隱私權和數據保護	必須在 AI 的整個生命週期保護隱私，並應建立適當的數據保護機制。
多利益攷關方與適應性治理和協作	使用數據時必須尊重國際法和國家主權。此外，包容性的 AI 治理方法應有各利害關係人的參與。
責任和問責	AI 應該是可審計和可追溯的。應建立監督、影響評估、審計和盡職調查機制，以避免與人權規範發生衝突以及對環境福祉造成威脅。
透明度和可解釋性	AI 部署取決於其透明度和可解釋性。
人類的監督和決定	成員國應確保 AI 不會取代人類的最終責任和問責制度。
永續發展	應根據 AI 對“永續發展”的影響對其進行評估。
公平和非歧視	AI 參與者應促進社會正義、公平和不歧視，同時採取包容性方法，確保所有人都可享受 AI 所帶來的益處。

## 二、AI 倫理之挑戰與省思

### (一) 聯合國教育、科學及文化組織 (UNESCO) 報告

UNESCO「AI 倫理問題建議書」，提及 AI 對人類帶來很多助益並惠及所有國家，但也會引發根本性的倫理問題，例如：AI 可能加劇偏見、導致歧視、不平等、數位鴻溝和排斥，並對文化、社會和生物多樣性構成威脅，造成社會或經濟鴻溝；演算法運作模式和演算法訓練數據應具有透明度和可理解性；AI 對於多方面的潛在影響，包括但不限於人的尊嚴、人權和基本自由、性別平等、民主、社會、經濟、政治和文化演進、科學和工程實務、動物福利以及環境和生態系統。

AI 會加深世界各地國家內部和國家之間現有的鴻溝和不平等，必須維護正義、信任和公平，以便在公平獲取 AI、享受這些技術帶來的益處和避免受其負面影響，同時承認各國國情不同，並尊重一部分人不參與所有技術發展的意願。

聯合國建議，倫理價值觀和原則可以透過指引制定和實施基於權利的政策措​​施和法律規範，以期加快技術發展步伐。此外，會員國應研擬影響評估（例如倫理影響評估）框架，以確定和評估 AI 所帶來益處、關切和風險，此影響評估應根據建議書提出的價值觀和原則，確定對人權和基本自由（特別是但不限於邊緣化和弱勢群體或處境脆弱群體的權利、勞工權利）、環境和生態系統產生的影響以及倫理和社會影響，並促進公民參與。各國政府應採用監管框架，特別針對公共管理部門提出 AI 倫理影響評估程序，以預測後果、減少風險、避免有害後果、促進公民參與和應對社會挑戰。評估還應確立能夠對演算法、數據和設計流程進行評估的適當監督機制，並包括對 AI 的外部審查，確保其具有可審計性、可追溯性和可解釋性。倫理影響評估應透明，並適時向公眾開放。此類評估還應具備多學科、多利害關係人、多文化、多元化和包容等特性。應要求公共管理部門透過引入適當的機制和工具，監測這些部門所實施和/或部署的 AI。

### (二) 學者研究

有關 AI 倫理相關議題眾多，謹列舉下列 3 項議題之學者研究結果：

#### 1. 倫理價值觀嵌入 AI ？

「如何知道一個所謂的『人工意識』機器人是否真的有意識，而不是僅僅表現為有意識的樣子呢？」(Hoffman & Hahn, 2018)。醫療機器人在設計上似乎是幫助改善老年人的健康狀況，雖然目的是良善的，程式設計也不會傷害人類，但是讓機器人變成受照

顧者每日生活不可缺少的模式，甚至因必須聽從機器人的命令而失去人類的尊嚴。這類機器人就像是「參與具有道德後果的行為，但由於缺乏自主指導的意圖而不能承擔道德責任……它們充當著沒有道德責任的道德行動主體者」(Weber, 2018)。AI 伴隨著機器學習演算法和專家系統，可能取代許多人類在做的事情，如：醫療診斷甚至是運動賽事的報導，讓人們不免質疑人類還有什麼剩餘的優勢，又如何維護人性和人道行為的觀念 (Kanuck, 2019)。

## 2. 開發具有倫理思考的自主武器？

有學者認為，戰爭時軍事人員和平民百姓的傷害比例難以主觀認定及估計，若能透過 AI 演算將可較精準的推估，因此，贊成應發展有道德的致命自主武器，至於其他的不具倫理道德判斷的致命自主武器都應該被禁止 (Umbrello、Torres 與 De Bellis, 2019)。

Solovyeva 與 Hynek (2018) 則探討自主武器系統的六個難題，包括：自主武器系統性能的可預測與不可預測性、殺戮決定的非人性化、將敵方戰鬥與非戰鬥人員去除其人格、協調操作中的人機關係、戰略考量以及自主武器系統運作的法律範疇，從是非優劣的正反兩方觀點，分析自主武器系統研發所引發的倫理、法律、政治、戰略和科學，尚待後續針對上述難題進行診斷和尋求解方。

## 3. 偵測 AI 歧視、偏見和犯罪？

人類社會依據人與人互動經驗所累積建立的倫理關係，制定歧視、偏見和犯罪的判斷準則。而當 AI 機器成為代理人，有可能複製人類的意志，或透過學習（超級智能），發生對人的歧視、偏見和犯罪行為時，所需的防範處置即為人機互動倫理的重要議題。例如：人臉辨識系統對於不同膚色人種影像的錯誤標記（例如：某種膚色種族特別容易犯罪）、語音辨識能力獨厚男性的聲音、搜尋引擎中的資訊偏差等 (Howard 與 Borenstein, 2018)。

# 三、使用 AI 可能伴隨而帶來的傷害（風險）

隨著 AI 使用的成長，其好處和風險也隨之增加。這些風險可能導致實際傷害、AI 事件，或潛在危險、AI 危害。明確定義風險及傷害，對於管理和預防這些風險及傷害至關重要。



## （一）OECD1.AI（嚴重）事件的定義<sup>10</sup>

OECD 針對真實風險、潛在風險發生之嚴重性，將 AI 事件及危害區分為 AI 嚴重災難等 5 類（詳圖 9），目的將 AI 風險先予分類，再研擬因應措施。

圖 9 OECD 根據傷害嚴重程度對 AI 事件和危害提出的分類

嚴重性 ↑	潛在風險	真實風險
	AI 潛在嚴重災難	AI 嚴重災難
		AI 嚴重事件
	AI 潛在災難	AI 事件

- (1) AI 嚴重災難，是一種嚴重的 AI 事件，它會擾亂社區或社會的運作，並可能測試或超越其利用自身資源的應對能力，其影響可能是直接的、局部的，也可能是廣泛的、持續很長一段時間的。
- (2) AI 嚴重事件，是指一個或多個 AI 的開發、使用或故障，直接或間接導致以下任何傷害的事件、情況或一系列事件：A. 人員死亡或人員嚴重傷害一個人或一群人的健康；B. 關鍵基礎設施的管理和運作受到嚴重且不可逆轉的破壞；C. 嚴重侵犯人權或嚴重違反旨在保護基本權利、勞工權利和智慧財產權的適用法律規定的義務；D. 對財產、社區或環境造成嚴重損害。
- (3) AI 事件，是指一個或多個 AI 的開發、使用或故障，直接或間接導致以下任何危害的事件、情況或一系列事件：A. 對個人或群體的健康造成傷害或損害；B. 關鍵基礎設施的管理和營運中斷；C. 侵犯人權或違反保護基本權利、勞工權利和智慧財產權的法律規定；D. 對財產、社區或環境造成損害。
- (4) AI 潛在嚴重災難，是指一個或多個 AI 的開發、使用或故障可能導致嚴重 AI 事件或 AI 災難的事件、情況或一系列事件，即以下任何危害：A. 某人死亡或某人或某群人的健康受到嚴重損害；B. 關鍵基礎設施的管理和運作受到嚴重且不可逆轉的破壞；C. 嚴重侵犯人權或嚴重違反旨在保護基本權利、勞工權利和智慧財產權的適用法律規定的義務；D. 對財產、社區或環境造成嚴重損害；E. 社區或社會的運作受到破壞，這可能會考驗或超越其利用自身資源應對的能力。

<sup>10</sup> 資料來源：OECD（2024），DEFINING AI INCIDENTS AND RELATED TERMS.

(5) AI 潛在災難，是指一個或多 AI 的開發、使用或故障可能導致 AI 事件的事件、情況或一系列事件，即以下任何危害：A. 對健康造成傷害或損害一個人或一群人的；B. 關鍵基礎設施的管理和營運中斷；C. 侵犯人權或違反保護基本權利、勞工權利和智慧財產權的適用法律規定的義務；D. 對財產、社區或環境造成損害。

## (二) 歐盟

歐盟 AI 法案立法核心採取「風險基礎論」(risk based approach) 來監管 AI，並依照 AI 可能對人的基本權利產生威脅之等級分類，將 AI 劃分為不同的風險級別：不可接受的風險 (Unacceptable risk)、高風險 (High Risk)、有限風險 (limited risk) 和最低風險 (Minimal risk)。謹將各該風險概述如下：

### 1. 不可接受的風險

依據法案第 5 條規定，對人們基本權利構成明顯威脅之 AI 將被禁止，如忽視使用者自由意志並操縱使用者行為之 AI 或應用程序，包括使用語音輔助鼓勵未成年人危險行為之玩具、政府或公司可進行「社會評分」之系統，此外，生物識別系統之即時用途將被禁止，例如在工作場所使用情緒識別系統、對自然人進行分類之 AI、及在公共場合將遠程生物識別系統用於執法等應用，惟倘生物識別系統用於尋找失蹤者、預防對自然人之生命威脅、識別犯罪嫌疑人等情況，可准予使用。

### 2. 高風險

依據法案第 6 條規定，倘 AI 作為產品之安全零組件，或該 AI 本身即為獨立產品且適用於法案附錄所列舉之法規清單，即可視為高風險 AI。此外，法案附錄 3 所敘明有關生物辨識、關鍵基礎設施、教育或職業培訓、評估使用公共或私人服務資格等 AI 與應用，倘對自然人之健康、安全及基本權利構成重大損害，亦將被視為高風險 AI，例如用於招募、評估貸款資格等 AI，依據法案第 8 至 15 條相關規定，高風險 AI 需要遵守相關要求，包括具備風險管理系統、高品質之資料集(data set)、活動記錄、明確的使用者資訊、人工監督和高水準的準確性及網路安全性等條件。

### 3. 有限風險

在使用 AI 過程中，使用者面臨的主要風險來自 AI 缺乏透明度，爰法案第 50 條制訂透明化相關規定，如必須能夠讓使用者瞭解其正與 AI 互動；AI 產生之內容必須被標記為人工生成；須讓使用者知悉其正處於生物識別分類或情緒辨識系統之使用環境等。

## 4. 最低風險

大多數 AI 皆屬最小風險，例如支援 AI 之推薦系統和垃圾郵件過濾器，由於該等系統對人們之權利與安全風險極小，無須承擔法案規定之義務。

## 三、AI 倫理評估架構

聯合國教育、科學及文化組織 (UNESCO) 為因應 AI 發展及運用，可能帶給人類倫理方面的議題，於 2023 年發布「準備狀態評估法：AI 倫理建議的工具 (Readiness Assessment Methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence)」及「倫理影響評估：AI 倫理建議的工具 (Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence)」，提供各國制定政策參考。

### (一) 準備狀態評估法 (Readiness Assessment Methodology, RAM)

RAM 包括一系列定量和定性問題，目的係收集與國家 AI 生態系統相關法律和監管、社會和文化、經濟、科學和教育以及技術和基礎設施等面向的訊息，作為國家推行 AI 參考，幫助各國瞭解在道德上和負責任地為所有公民實施 AI 的準備程度，從而強調需要進行哪些制度和監管變革 (RAM 內容，詳附錄 5)。

### (二) 倫理影響評估法 (Ethical Impact Assessment Methodology, EIA)

EIA 透過一系列的問題，經過政府部門自我檢視後，可幫助確保所購買的 AI 符合建議中規定的道德標準。私部門和其他機構的開發人員亦可使用 EIA，以促進 AI 的道德設計、開發和部署 (EIA 內容，詳附錄 6)。

## 四、建立值得信賴 AI 的倫理指引

歐盟委員會 (European Commission) 於 2019 年發布「可信任 AI 倫理指引 (Ethics Guidelines for Trustworthy AI)」，說明值得信賴的 AI 4 項道德原則，分別係：

### (一) 尊重人類自主權

與 AI 互動的人類必須能夠保持充分有效的自決權，並且能夠參與其演進。AI 不應該無理地服從、脅迫、欺騙、操縱、制約或群體化人類。相反，它們應該被設計增強、補充和增強人類認知、社會和文化技能。在工作環境，AI 應協助人類，並執行有意義的工作。

## （二）避免產生傷害

AI 系統不應造成或加劇傷害或對人類產生其他不利影響。這需要保護人性尊嚴以及身心健全。AI 系統及其環境必須安全可靠，必須在技術上穩健，並且應確保不會容易被惡意使用。弱勢群體應得到更多關注並參與發展，AI 系統的部署和使用。也必須特別注意 AI 系統可能導致的情況或因權力或資訊不對稱，對審計人員、企業和消費者或政府和公民加劇不利影響。

## （三）維持公平性

AI 系統的開發、部署和使用必須公平，包含實質上及程序上的公平。實質公平方面，係指確保利益和成本的平等和公正分配，個人和團體免受不公平的偏見、歧視和污名化。如果存在不公平的偏見可以避免，AI 系統甚至可以增加社會公平。受教育機會均等，也應促進商品、服務和技術的發展。此外，AI 系統的使用絕不應該導致人們的選擇自由受到欺騙或無理損害。此外，公平意味著 AI 從業者應尊重手段與目的相稱的原則，並仔細考慮如何平衡相互競爭的利益和目標。針對 AI 系統及其操作人員所做的決策尋求有效補救。負責決策的實體必須是可識別的，並且決策過程應該是可以解釋的。

## （四）具有可解釋性

可解釋性對於建立和維持使用者對 AI 系統的信任至關重要。這意味著流程需求透明，公開溝通 AI 系統的能力和目的，以及可以向直接和間接受影響的人解釋。

# 第三節 AI 發展及運用等監理機制及規範

## 一、AI 發展

### （一）OECD

OECD 近年來致力於整合世界各國力量發展 AI，不論是舉辦論壇，或是與世界組織合作共同研究，促進朝向符合 OECD AI 原則目標發展，有關 OECD 行動，謹擇要摘列如下：

#### 1. 結合 AI 全球夥伴聯盟 (Global Partnership on Artificial Intelligence, GPAI) 共同合作

為釐清應如何針對資料隱私權進行保護、演算法應公開透明度，同時避免 AI 及其

帶來之社會價值的詮釋權掌握在少數人手中。作為現今 AI（深度學習）發源地之加拿大及法國於 2018 年起，即不斷推動、倡議建議獨立之國際性專家平臺，以強化民眾對於科技之信賴，同時避免科技遭到誤用。加拿大與法國於於 2018 年 6 月加拿大 G1 峰會期間首度發布「法、加 AI 聯合宣言」，呼籲成立 AI 專家平臺，兩國復於 2019 年 8 月法國 G7 峰會時提出成立 GPAI 倡議，續於 2020 年 6 月 15 日公布 GPAI 共同聲明，目前有 29 個成員國。OECD 和 GPAI 聯手建立綜合夥伴關係，利用 GPAI 和 OECD AI 工作計畫之間的協同作用和互補性，實施中體現以人為本、安全、可靠和值得信賴的 AI，促進治理和專家層面的包容性參與，並促進更有效率的流程，減少成本和重複。目前合作夥伴有 44 個國家。

## 2. 成立 OECD AI 工作小組和專家網絡 (OECD Working Party and Network of Experts on AI) 共同合作

OECD 與 GPAI 建立綜合夥伴關係之後，成立 OECD AI 工作小組和專家網絡，下分健康專家組、AI、數據和隱私專家組、OECD AI 指數專家組、AI 風險與責任專家組、AI 未來專家組、AI 事件專家組、計算專家組等 7 個組。謹簡要介紹如次：

### (1) 健康專家組 (Expert Group on Health)

AI 為面臨多重危機的衛生系統帶來希望，但其在醫療保健中的使用引發了人們對敏感資料隱私和使用黑盒演算法的擔憂。OECD 於 2024 年 5 月成立健康專家組，提供交流和相互學習的論壇，並以人為本、負責任、公平和可持續的方式應對大規模實施 AI 的挑戰。該專家組由來自 27 個國家和 13 個國際組織、倡導團體和患者代表的 60 多名成員組成。這意味著利用 AI 透過數據幫助世界 80 億人口實現期望的健康結果，邁出重要一步。健康專家小組的工作將專注於制定兼容的政策，以跨越孤立框架開發、實施和擴展 AI，使其在全球範圍內變得可及且有用。該小組將運用洞察力及齊力決有關法規相容性、標準統一、技術信任、激勵機制以及數據治理等挑戰，幫助醫療保健提供者運用 AI。OECD 期待政府成員、國際組織、學術界、民間社會和患者代表共同努力，交流知識和經驗，以建立一個共同框架，以建立由數據和 AI 驅動的最先進、高效和公平的衛生系統。

### (2) AI、數據和隱私專家組 (Expert Group on AI, Data and Privacy)

AI 依賴於越來越多的訓練數據，包括透過各種方法獲取的個人數據。在 OECD，AI、隱私和數據保護政策社群傳統上是分開運作的，分別依據 OECD 隱私指導方針中的既定原則進行運作。然而，隨著 AI 能力的進步和近年來 AI 應用的興起，越來越多的聲音呼籲加強 AI、隱私和數據保護社群之間的協同作用。OECD AI、數據與隱私專家小組與 AI

治理工作小組（AIGO）合作應對這些挑戰，制定能夠最大化 AI 創新潛力並減輕其隱私風險的政策框架，要求達到前所未有的跨領域合作。AI、數據與隱私專家組包括來自數據保護機構、政策制定者、產業、民間社會和學術界等。該組為各國政府提供知識和工具，制定符合未來需求的政策。專家小組的目標包括：

- A. 短期目標：確定相關的隱私原則和實踐，以支持可信賴的 AI 發展；界定 AI 商業行為的隱私和治理議題；盤點 AI、數據和隱私之間的協同作用與國際合作領域；提供 AI、數據和隱私的國家政策發展最佳實務；
- B. 長期目標：促進多方利害關係人對 AI、數據和隱私領域未來政策挑戰的討論；制定框架和指導方針，以確保 OECD 國家在 AI、數據和隱私政策上的全球一致性。

### （3）OECD AI 指數專家組（Expert Group on OECD AI Index）

OECD AI 指數目的在設計和實施一個關於可信賴 AI 的綜合衡量框架，該框架依據 OECD 10 項可信賴 AI 原則。OECD AI 指數滿足政策制定者對基於高質量指標的全面且具權威性的 AI 指數的需求。在制定 AI 政策時，政策制定者需要一個可靠的框架，評估 AI 政策和實踐的創新性及其可信度、所帶來的益處。OECD AI 指數的特點及附加價值如下：

- **全面性**：涵蓋 AI 生態系統的各個方面以及 OECD AI 原則，並基於 OECD 自 2018 年以來的 AI 指標和衡量工作。
- **廣泛認可**：在開發過程中，將諮詢來自所有利益相關群體的 AI 專家以及全球統計機構的意見。
- **權威性**：採用政府認可的方法論。
- **創新性**：捕捉當前未在其他地方體現的重要 AI 促進因素，如 AI 政策與信任、AI 計算與數據、以及 AI 的環境影響。
- **易於使用**：提供簡單的高層次數據和可視化圖表，便於理解，並深入探討具體的國家或行業數據。
- **全球與包容性**：涵蓋盡可能多的國家，儘管對某些國家可能僅提供部分指標。
- **為 OECD AI 原則的 2024 年審查提供參考**：OECD 理事會將在 2024 年對 OECD AI 原則的實施情況進行審查。
- **實用性**：包括具體的使用案例。

#### (4) AI 風險與責任專家組 (Expert Group on AI Risk& Accountability)

儘管 AI 帶來了巨大的好處，但也帶來偏見和歧視、意見兩極化、侵害隱私等實際風險及在某些國家出現的大規模監控，已開始對人類和社會造成實際的傷害。為開發「可信賴」、「負責任」或「道德」的 AI 系統，需要對其影響進行評估並管理 AI 風險，這包括生成式 AI 的應用背景。在過去的幾年裡，全球逐漸趨向於採用風險為基礎的方式和影響評估，幫助規範 AI 的運作。OECD 與世界其他組織或國家、專家進行交流，以確定評估可信賴 AI 的共同指標，從而評估 AI 的風險和影響。其目標是促進全球一致性，幫助實施有效且可問責的可信賴 AI 系統。該組工作包括下列 5 項：

- 研擬現有和正在開發的 AI 設計的核心標準、框架和指導方針，包括 AI 影響、合規和風險評估，以及 AI 風險管理。
- 盤點各項倡議間概念和術語的共同點和差異，並進行差距分析，如有必要，提出可能的術語。
- 將分析結果轉化為在 AI 整個系統生命週期中，進行負責任商業使用。
- 研究和分析 OECD 負責任商業行為 (RBC) 和 AI 標準的符合情況。
- 開發一個互動式線上工具，幫助組織和利害關係者比較框架，並瀏覽現有的方法、工具和良好實務，以識別、評估、處理和管理 AI 風險。

#### (5) AI 未來專家組 (Expert Group on AI Futures)

OECD AI 未來專家小組透過提供對 AI 未來的洞察，並為各國政府提供必要的知識和工具，以制定前瞻性的 AI 政策來應對這些挑戰。小組提供的見解包括：可能的先進 AI 系統未來發展的關鍵里程碑；與每個里程碑相關的各種風險；AI 安全的最佳實務；有關 AI 未來的情景；建議的政策方法，力求享受 AI 帶來的益處的同時減輕未來風險。

#### (6) AI 事件專家組 (Expert Group on AI Futures)

儘管 AI 帶來了巨大的好處，但它也帶來了風險。其中一些風險已經對人類和社會造成了實際的傷害，例如偏見和歧視、意見極化、隱私侵害以及安全和安全問題。這些傷害通常被稱為“AI 事件”。隨著 AI 在各個經濟和社會領域的持續應用，AI 事件的增加將無法避免。為開發可信賴且有益的 AI 系統，必須處理這些風險，首先要理解 AI 事件。監測 AI 事件需要全球一致努力，以便 AI 系統操作員和政策制定者能夠從其他國家見賢思齊。OECD 已經開始著手制定一個報告框架，且 OECD 開發全球 AI 事件監測器(AIM)

及時追蹤實際的 AI 事件。偶見國際媒體報導的 AI 事件，但許多其他事件未公開披露，現有數據顯示 AI 事件呈指數增長。即時透過監測 AI 事件來告知政策和監管選擇的需求日益增加，透過記錄過去或現在已造成或促成實際事件或近乎事件的「高風險」AI 系統的特徵。OECD 正與政策制定者、專家和所有利益相關群體的合作夥伴共同開發 AI 事件報告的共同框架。

### (7) 計算專家組 (Expert Group on Compute & Climate)

除了數據和演算法外，AI 計算能力係推動 AI 及相關經濟增長和競爭力的關鍵因素。OECD 計算專家組幫助 OECD 建立一個基本框架，以便按國家和地區瞭解、衡量和基準測試該國 AI 計算能力。該專家組持續與主要的 AI 計算參與者合作，並已開始一項數據收集工作。

## (二) 歐盟

### 1. 成立 AI 辦公室 (EU AI Office)

歐盟於 AI 法案第 56 條規定成立 AI 辦公室，以在促進 AI 的未來開發、部署和使用，促進社會和經濟效益和創新，同時降低風險。該辦公室將特別是在通用 AI 模型方面。致力於促進可信賴 AI 的研究和創新，並將歐盟定位為國際討論的領導者。

AI 辦公室內部組成及業務，概述如次：

- **監管和合規部門**：與成員國密切合作，協調監管方法，促進整個聯盟 AI 法案的統一應用和執行，並協助調查可能的侵權行為並實施制裁。
- **AI 安全部門**：專注於識別通用模型的系統性風險、可能的緩解措施以及評估和測試方法。
- **卓越 AI 和機器人部門**：支援和資助研究和開發，以培育卓越的生態系統。它協調 GenAI4EU 計畫，刺激模型的開發並將其整合到創新應用中。
- **AI 造福社會單位**：負責設計和實施 AI 辦公室在 AI 造福領域的國際參與，例如天氣建模及癌症診斷等。
- **AI 創新和政策協調部門**：負責監督歐盟 AI 戰略的執行，監測趨勢和投資，透過歐洲數位创新中心網路和 AI 工廠的建立以刺激 AI 的採用，並支援監管沙箱培育創新生態系統。



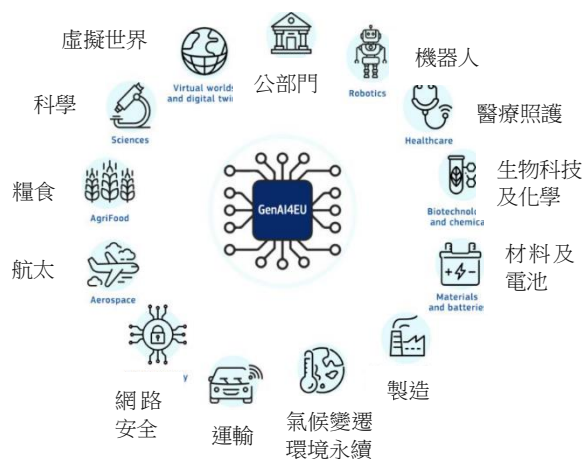
## 2. 成立 AI 委員會 (European Artificial Intelligence Board)

歐盟於 AI 法案第 65 條規定，成立 AI 委員會，由歐盟委員會和所有歐盟成員國的高級代表組成，作為 AI 監管機構的論壇，包括 AI 辦公室、國家當局和歐洲資料保護監管機構 (EDPS)，可以協調 AI 法案的一致應用。該委員會將關注歐盟 AI 政策的更新和討論，AI 法案實施情形，交流 AI 治理和 AI 法案實施的最佳實務及國家方法。

## 3. 執行 AI 創新計畫 (AI Innovation Package)

2024 年 1 月，歐盟委員會啟動 AI 創新計畫，支持新創企業和中小企業開發符合歐盟價值觀和規則的值得信賴的 AI，致力於開發歐洲 14 個工業生態系統以及公共部門的新穎用例和新興應用程式。應用領域包括機器人、健康、生物技術、製造、移動、氣候和虛擬世界等 (詳圖 10)。

圖 10 歐盟 AI 創新計畫運用領域



## 4. 推出數位歐洲計畫 (The Digital Europe Programme)

歐盟資助計畫，總預算超過 79 億歐元，塑造歐洲社會和經濟的數位轉型，致力於將數位技術帶給企業、公民和公共管理部門，提供策略性資金來應對這些挑戰，支援超級運算、AI、網路安全、先進數位技能等關鍵能力領域的項目，並確保數位技術在整個經濟和社會中的廣泛使用。它透過強化的歐洲數位創新中心 (EDIH) 網路支持工業、中小企業 (SME) 和公共管理部門的數位轉型。

### (二) 荷蘭

荷蘭政府致力於運用 AI 於公共服務，荷蘭應用科學研究組織 (TNO) 於 2024 年 8 月發布的研究顯示，荷蘭政府對 AI 應用程式的使用，相較 2021 年增加 1.5 倍以上。總計 266 個已確定的 AI 應用程式中，有 105 個 (39%) 是由該國地方政府機關所部署，此外，荷蘭政府戮力提高 AI 及其演算法的透明度。謹將荷蘭政府近年來有關 AI 政策，臚列如次：

### 1. 全球負責任 AI 指數 (Global Index Responsible AI, GIRAI) 評估結果

(1) GIRAI 評估 138 個國家 AI 發展及運用情形，由當地研究人員回答 1,800 多個問題，收集合計超過 200 萬筆數據後予以分析，於 2024 年在華盛頓公布調查結果，荷蘭綜合成績為世界第一。該調查得出下列結論：

- A. 國際合作對於負責任的 AI 的部署和使用扮演重要角色。
- B. AI 的安全性、可靠性有待提升。
- C. 工人沒有得到充分的保護。
- D. 大學和民間社會在促進負責任的 AI 方面發揮著重要的作用。
- E. 要在全球範圍內實現足夠水平的負責任的 AI，還有很長的路要走。

(2) 荷蘭非常重視負責任 AI 的發展與運用，例如：

- A. ELSA (道德法律社會) 實驗室和 ICAI (AI 創新中心) 實驗室與企業、政府、學術界和民間社會組織合作開發具體實施方案。
- B. 許多聯盟、基金會和社區致力於提高人民對於 AI 認知等教育活動。
- C. 戮力於 AI 和演算法的使用開發框架和影響評估。
- D. 學術界非常熱絡於研究演算法和 AI 等相關議題。

## 2. AI 策略行動計畫 (Strategic Action Plan for Artificial Intelligence, SAPAI)

荷蘭政府考量 AI 正改變世界，為利用 AI 促進荷蘭的經濟增長、繁榮和福祉，幫助解決老齡化、氣候變化、食品安全、健康和護理等社會問題，同時，又正視保護隱私、不歧視和自主權等基本權利的挑戰，於 2019 年 10 月頒布 AI 策略行動計畫 (SAPAI)，希望善用世界級網絡、數據中心及主機提供商等良好基礎條件，發展及運用 AI。於 SAPAI，荷蘭政府研擬 3 項路徑，分別係利用社會和經濟機會、創造適當的條件及加強基礎建設 (如表 4)，謹分述如下：

表 4 荷蘭 AI 策略行動計畫路徑及其內涵

路徑	內涵
利用社會和經濟機會	AI 提供社會問題的解決方案
	政府運用 AI 執行公共任務
	鼓勵 AI 創業
創造適當的條件	荷蘭的 AI 研究和創新具有高質量，並在

	歐洲領先
	荷蘭擁有更多可用數據，以實現更好的 AI 開發
	荷蘭在歐洲具有高質量的數據、智能連接能力和計算能力，為有效的 AI 應用奠定了基礎
加強基礎建設	公共價值和人權得到保護
	使每個人都能信任 AI
	市場開放、競爭激烈並為消費者提供良好的保護
	保護荷蘭公民、企業和政府的安全

### (1) 路徑一：利用社會和經濟機會

AI 提供巨大的社會和經濟機會。首先，在解決社會挑戰方面存在機會，例如：政府利用 AI 提高偵查和執法的效率，提供新的預防、診斷和治療方法，預測交通擁堵。AI 還有很大的潛力在公共任務執行中發揮作用，有助於改進政府組織的工作流程。此外，政府鼓勵企業開發 AI 應用。

#### A. AI 提供社會挑戰的解決方案

AI 可以幫助解決許多社會挑戰，如安全、健康和醫療、農業和食品、能源轉型和可持續性。政府特別重視荷蘭獨特的優勢：卓越的公私合作（PPS）。

#### (A) 安全方面

AI 技術在司法、安全和國防領域提供了許多機會。在安全領域，不僅是社會挑戰，也是公共任務執行的重要部分。目前荷蘭政府與科學機構合作，正在進行 AI 應用機會的研究，例如警務、司法和國防領域。

在警務領域，已在多個方面已經投入使用。例如，選擇相關影像資料進行調查研究及報警事件。為進一步推動 AI 在警務領域的研究和發展，警察機構成立了國家警察 AI 實驗室，探索 AI 如何提高警務的效率和效果。

在司法領域，開始 AI 應用的小範圍實驗，暫時不涉及現實中的法律案件或爭議。

在國防領域，國防部已經進行了 AI 及其對國防影響的研究，例如，使用 AI 進行反制水雷的水下無人機（Remus）和防禦系統（愛國者）的應用。

## (B)健康和醫療方面

AI 在預防、診斷和治療以及物流方面提供了新的機會，例如：AI 工具可以精確分析癌細胞，從而做出更好的診斷；閱讀 X 光片，更快地識別和分析疾病；可穿戴技術和健康應用程序幫助市民更好地自我管理健康。

## (C) 農業和食品方面

AI 可用於瞭解購買和消費行為，收集和分析社交媒體上有關水果和蔬菜的言論，以瞭解公民的消費行為。

## (D) 能源轉型及永續發展方面

AI 可應用於引擎、車輛和駕駛員、所有道路使用者、供應鏈、交通和運輸系統及環境中。AI 亦可以幫助發現市場趨勢、識別風險、減少交通擁堵、減少溫室氣體和空氣污染物排放、設計和管理運輸、分析旅行需求和行人行為。例如，荷蘭公共工程和水管理局（Rijkswaterstaat）使用 AI 預測交通擁堵、預防事故並優化基礎設施。

## B. AI 提供社會問題的解決方案

荷蘭政府希望充分利用 AI 執行公共任務，AI 應用可以改進不同政府機構的工作流程，並提供更好的社會問題解決方案。但這需要跨機關間良好的合作，例如；整合不同機關持有的數據。此外，試驗 AI 應用，並鼓勵市場參與者提出創新的解決方案，以改進公共任務的執行。

## C. 鼓勵創業

許多大型荷蘭企業已經在研究和應用 AI，提供更好的服務和提高生產力，但這些仍處於計畫或試點階段。中小企業的 AI 應用和 AI 新創公司和成長公司的數量尚待大幅增加，大型企業可以幫助中小企業和 AI 新創公司和成長公司開發和應用 AI 驅動的創新和新商業模式，這些企業面臨的問題往往是其他技術領域的共同問題，包括尚待汲取專業知識、缺少資本及人才、國際化。

### (2) 路徑二：創造適當的條件

為加速 AI 的發展，荷蘭政府希望在創造一個充滿活力的 AI 環境，行動方案包括：

- 發布 AI 研究議程。
- 建立一個領先歐洲的 AI 知識中心。

- 教育、文化和科學部（OCW）將投資超級計算機。
- 透過荷蘭創新網絡加強國際 AI 合作。
- 促進職業和高等教育的數字素養，包括中等職業教育（mbo）區域投資基金，以強化 mbo 與勞動市場的銜接。
- 設立國家數據科學實習計畫。
- 將數位素養納入小學和中學的修訂課程中，並透過數字化小學和中學教育議程，得到企業界的支持。

### （3）路徑三：加強基礎建設

為實現長久及負責任地發展 AI，荷蘭政府將加強基礎設施，包括法律和倫理框架，並保護公民的基本權利和價值觀，行動方案包括：

- 保護公民隱私、平等對待和自主權的基本權利和價值觀。
- 建構適當的法律和倫理框架，確保負責任使用 AI。
- 確保市場開放和競爭，並為消費者提供良好的保護。
- 保護國家的安全，確保公民、企業和政府的安全。

## 3. 荷蘭政府運用於 AI 於國防業務情形

### （1）政策規劃情形

荷蘭政府發展及運用 AI 於多面向行政業務，其中給人較敏感極具機密性之國防業務，荷蘭政府承諾將國防支出維持在至少佔國內生產總值（GDP）的 2% 以上，荷蘭國防部認為，AI 和數據科學對於保護荷蘭非常重要，爰提出 2035 年戰略願景（Strategic Vision 2035），並研擬「2021 - 2025 年戰略知識與創新議程（Strategic Knowledge and Innovation Agenda 2021-2025）」及「2023 - 2027 年數據科學與 AI 戰略（The Data Science and AI Strategy 2023 - 2027）」。  
謹將其國防業務發展及運用 AI 情形，分述如下：

#### A. 2035 年戰略願景

為確保國防部能夠透過 AI 和數據科學在未來獲得新技術並支援決策，確定未來 3 大主要目標。

### 目標一：提升技術

在人類控制下，使用 AI、數據科學、大數據和（半）自主系統，需要專注於創新且具備（資訊）技術方面具備教育背景。國防部的武器系統和部隊必須具有模組化並易於升級的特性，以便它們能夠與其他兼容的系統和部隊集結、整合。在創新過程中，特別關注倫理和法律方面。

### 目標二：資訊驅動的作戰

國防部的目標是在北約和歐盟內取得並保持顯著的資訊優勢，重點關注資訊為核心的作戰，關鍵概念之一是「資訊機動」，其核心在於運用軍事資訊能力，透過在作戰環境中創造有影響力的效果來塑造受眾的行為。此外，隨著收集大量數據，國防部利用數據科學和 AI 分析這些數據，有效管理複雜的情境和作戰，並擴大和促進其在日益複雜的部署能力。例如：AI 演算法可支持情景分析、視頻數據中的圖像識別，或支持數據收集的最佳路徑分配演算法。

### 目標三：成為可靠的合作夥伴

荷蘭國防部將更加公開透明，並加強向議會和公眾提供資訊的能力，使荷蘭能夠對威脅保持警覺並建立韌性。

## B. 2021 – 2025 年戰略知識與創新議程

戰略知識與創新議程（Strategic Knowledge and Innovation Agenda 2021-2025, SKIA）特別關注於短週期創新的能力，並擘劃知識建構、技術發展和創新，荷蘭國防部制定開發人工智能和機器人等關鍵技術的具體行動計畫，其重點關注 4 個領域：

- 加強國防專業知識。
- 將創新融入工作環境、文化和管理。
- 強化國防部的知識和創新合作夥伴間的合作。
- 加強國防知識與創新鏈的合作關係。

為實現這些目標，SKIA 著重於電子戰領域、傳感器系統及其在應對遙控飛行器系統中的應用。

## C. 2023 至 2027 年數據科學與 AI 戰略（The Data Science and AI Strategy 2023 – 2027）

荷蘭國防部諮詢包括產業領袖和跨大西洋合作夥伴在內的公共和私人合作夥伴後，制定「2023 - 2027 年數據科學與 AI 戰略」，作為未來 5 年內優化使用 AI 和數據科學的明確基礎，並每 5 年更新一次戰略。此外，荷蘭國防部正在制定框架和指南，記錄演算法的運作、所做的決策及其在工作場景中的應用。在軍事領域負責任地部署數據科學和 AI 也受到國際社會的高度關注，該部期望在國際標準和認證發展方面扮演領導角色。優化倫理法律框架以及在軍事領域中負責任地使用數據科學和 AI，並與北約和歐盟夥伴、荷蘭應用科學研究組織（TNO）、荷蘭皇家航空航天中心（NLR）及產業界密切合作，成為發展重點。數據科學與 AI 戰略選擇無人自主系統、軍事決策支持與情報、後勤和預警性維護、行政業務等 4 個 AI 的應用領域，並與北約和歐盟等戰略夥伴共同引導 AI 的發展。

### **（2）舉辦首屆全球負責任人工智能軍事領域峰會（REAIM 2023）**

荷蘭政府與韓國於 2023 年共同舉辦首屆全球負責任 AI 軍事領域峰會（REAIM 2023），這是全球第一個聚焦於負責任使用 AI 於軍事領域的會議，聚集來自世界 100 個國家、超過 2,000 名之政府、企業、學術界、新創公司和民間社會之與會者，就軍事衝突和戰爭中部署和使用 AI 達成共識。會議中政府代表表示，意識到 AI 在軍事領域的機遇與潛力，同時也瞭解其中的風險。因此，與會者共同呼籲在國際法的框架下，以不破壞國際安全、穩定和符合問責制度的方式，負責任地開發、部署和使用 AI 於軍事領域。荷蘭透過參與 REAIM 高峰會，承諾將負責任、符合國際法運用 AI 於軍事領域，同時也會尊重人權。荷蘭利用其在國際法領域的專業知識，透過強調法律原則的重要性並保持 AI 開發和部署過程中的問責性與透明性，為塑造全球 AI 治理框架做出貢獻。

### **（3）於無人作戰系統運用 AI 情形**




荷蘭國防部推行關於機器人和自主系統（RAS）的概念開發與實驗計畫，確定組織、作戰概念和無人系統的最佳組合，強化部隊的戰鬥力並提高士兵的保護。該單位隸屬於荷蘭陸軍，主要發展戰鬥無人地面系統、群體無人航空系統和自主性等關鍵主題。自動化功能（如導航）可以減少人類操作員的認知和身體負擔。無人系統的自主性是必要的，以確保在操作員與無人系統之間的數據鏈路被干擾或無法使用時，仍可以持續執行任務。自主性更可擴大於許多無人系統執行命令任務，並由少數人類監督。但荷蘭陸軍要求有自主性的無人系統仍可透過人類控制。

## 二、AI 監管機制及規範

### (一) 概述

全球對於 AI 術的興起和廣泛應用，正促使各國政府和國際組織加強監管框架的制定與執行。隨著 AI 技術在醫療、金融、交通、教育等多領域的應用日益深入，如何確保 AI 的使用安全、道德並且符合社會責任，成為全球關注的焦點。表 5 為近年來世界各國及組織針對 AI 監理規範之 12 項重要發展歷程。

表 5 AI 國際規範 12 項重要發展歷程

OECD AI 原則 (OECD AI Principles)	提倡使用創新、值得信賴、尊重人權和民主價值的 AI。各國利用該原則制定政策並建立其 AI 風險框架，為司法管轄區間之全球互通性奠定基礎。		2019 年 5 月
G20 AI 原則 (G20 AI Principles)			2019 年 6 月
GPAI (The Global Partnership on Artificial Intelligence)	成立 AI 全球夥伴聯盟(GPAI)，促進國家合作。		2020 年 6 月
聯合國教科文組織 AI 倫理建議書 (UNESCO AI Ethics Recommendation)	AI 道德框架由 193 個會員國通過。		2021 年 11 月
G20 新德里領袖宣言(G20 Leaders' Declaration)	G20 新德里高峰會通過利於創新 AI 法案。		2023 年 9 月
G7 廣島生成式 AI (AI) 進程 ( G7 Hiroshima AI Process)	舉辦論壇討論生成式 AI 衍生的風險。		2023 年 10 月
布萊切利宣言 (Bletchley Declaration)	於第 1 屆 AI 安全高峰會，由 28 個國家及歐盟簽署。		2023 年 11 月



聯合國大會 AI 決議 (UN General Assembly AI Resolution)	聯合國大會通過「抓住安全、可靠和值得信賴的 AI 促進永續發展的機會 (Seizing the opportunities of safe, secure, and trustworthy artificial intelligence systems for sustainable development)」，倡議於 AI 與 AI 之間架起橋樑。		2024 年 3 月
OECD AI 原則更正	針對於 2019 年 5 月發布 OECD AI 原則之 AI 定義等，予以更正。		2024 年 5 月
歐洲理事會 AI、人權、民主及法治框架公約 (Council of Europe Framework Convention on artificial intelligence and human rights, democracy, and the rule of law)	涵蓋 AI 相關生態系統的法律框架，讓 AI 可在進步發展的同時，確保不會對人權、民主及法治造成影響，並且使 AI 技術維持中立、平等的使用特性。		2024 年 5 月
AI 首爾高峰會 (AI Seoul Summit)	發布《首爾宣言》和《關於 AI 安全科學的國際合作首爾意向書》。		2024 年 5 月
歐盟 AI 法案 (EU AI Act)	歐盟 AI 法案生效，世界第一部全面性的 AI 法案。		2024 年 8 月

## (二) OECD

OECD 近年積極推動全球 AI 倫理和監管的標準化，確保 AI 技術的發展能夠符合人類利益，並建立共同的國際規範，採取發布 AI 倫理原則、建立 OECD AI 事件監測機制及成立 OECD AI 治理工作小組等。謹分述如下：

### 1. AI 倫理原則

- (1) 2019 年 5 月，OECD 發布「AI 倫理原則」，係全球首批針對 AI 的政府間原則，後來成為國際通行標準。該原則強調 AI 應該以人為中心，支持經濟增長並促

進公民福祉，計有 5 項原則，並對世界各國提出 5 項建議（詳表 6）。

表 6 OECD AI 5 項原則及其建議

5 項原則	5 項建議
A. 以人為本和公平：AI 應尊重人權和民主價值觀，並應促進福祉和公平。	A. 投資 AI 研發。
B. 可解釋性與透明性：AI 決策過程應該是透明的，並且系統應該能夠被解釋，使用者和受影響者得以理解其運作方式。	B. 發展包容性 AI 生態系統。
C. 穩健性、安全性與問責性：AI 應該具備穩健性，能夠應對各種風險，其運行應遵循安全和問責的標準。	C. 為 AI 塑造有利且可互通的治理和政策環境。
D. 數據治理與隱私保護：AI 的開發與應用應符合最高標準的數據隱私保護要求，保護個人的數據權利。	D. 提升人員能力並為轉型勞動市場轉型做好準備。
E. 跨國合作與標準化：鼓勵國際合作，以維持 AI 的全球治理和技術標準一致，避免監管碎片化。	E. 建立值得信賴的 AI 的國際合作。

(2) 2024 年 5 月，OECD 為了下列目的，將 AI 原則進行修正：

- A. 因應生成式 AI 錯誤訊息和虛假資訊，及認知維護資訊完整性日益重要。
- B. 解決預期目的以外的使用、故意濫用或無意濫用。
- C. 強調 AI 參與者應提供有關 AI 的信息，以確保透明度和負責任的揭露。
- D. 解決安全問題，以便如果 AI 有造成不當傷害或表現出不良行為的風險，則可以透過修復或退役；
- E. 強調在整個 AI 生命週期負責任的商業活動，
- F. 強調各司法管轄區需要共同努力，促進 AI 可互通治理和政策環境，以因應全球 AI 政策措施的增加。
- G. 強調日益重要的環境永續性。

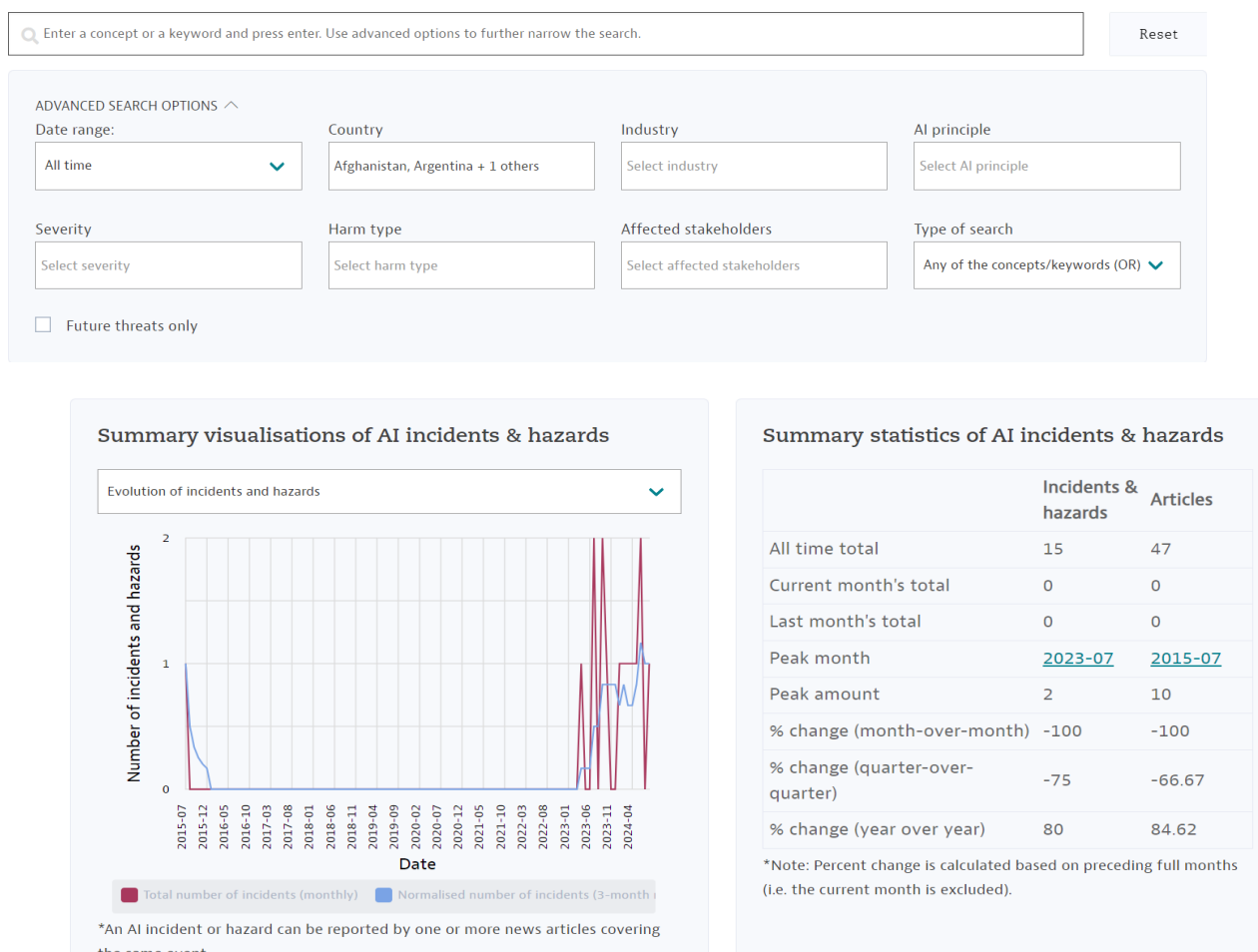
## 2. OECD AI 事件監測機制 (OECD AI Incidents Monitor, AIM)

與 AI 相關的立法正逐漸獲得重視，制訂有效政策需要證據、前瞻性和國際合作。OECD AI 事件監測器 (AIM) 記錄 AI 事件和危害，提供政策制定者、AI 從業者及全球所有利益相關者獲得有關 AI 系統風險和危害的寶貴證據。隨著時間的推移，AIM 將有助

於顯示風險模式，認知 AI 事件和危害及其多面性質。

使用機器學習模型識別全球在知名國際媒體上報導的 AI 事件和危害，並將事件和危害分類為 OECD AI 系統分類框架中的不同類別，包括其嚴重性、行業、相關的 AI 原則、傷害類型和受影響的利益相關者。AIM 基於每篇新聞文章的標題、摘要和前幾段進行分析。新聞文章來自事件註冊平台，這是一個監控全球新聞並能檢測新聞文章中報導的特定事件類型的新聞情報平台，每天處理超過 15 萬篇英文新聞文章。儘管認識到這些事件和危害可能僅代表全球所有 AI 事件和危害的一部分，但這些公開報導的事件和危害為提供證據基礎。事件和危害可以由一篇或多篇報導同一事件的新聞文章組成。為了減少與編輯偏見和虛假信息相關的擔憂，每個報告的註釋和元數據是從報導該事件和危害的最知名新聞媒體提取的，並根據 Alexa 流量排名進行評估。此外，事件和危害按報導該事件的文章數量和與特定查詢的相關性進行排序。最後，為了完整性，提供所有報導特定事件或危害的文章鏈接。該系統可透過設定篩選條件，得出結果（畫面截圖如圖 11）。

圖 11 OECD AI 事件監測機制



### 3.成立 OECD AI 治理工作小組(OECD Working Party on Artificial Intelligence Governance, AIGO)

OECD 數位政策委員會 (DPC) 設有 AI 治理工作小組 (AIGO)，負責監督 AI 政策制度情形。OECD 成員國提名負責各自國家 AI 政策的國家官員為該小組成員。

AIGO 負責監督 AI 政策和治理工作計畫，包括：

- 分析：審查成員國 AI 政策和行動計畫的設計、實施、監測和評估。
- 影響評估：評估 AI 技術的影響。
- 可信任 AI：發展負責任和可問責的 AI 方法。
- 測量與數據：監督 OECD AI 觀察站的趨勢與數據工作。
- 前瞻工作：進行 AI 及相關新興技術的前瞻性研究。

## (三) 歐盟

### 1.一般資料保護規則(General Data Protection Regulation, GDPR)

歐洲向來極為重視個人隱私保護，歐洲議會及歐盟理事會於 1995 年制定之「資料保護指令 (Data Protection Directive)」施行逾 20 年後，於 2016 年 4 月 27 日通過「一般資料保護規則 (General Data Protection Regulation, GDPR)」<sup>11</sup>，並自 2018 年 5 月 25 日起施行，整合整個歐洲的資料隱私法律規範。GDPR 規範重點如下：

- (1) 適用範圍：不論資料管理者或資料處理者於歐盟境內是否設立分支機構，只要其在跨境提供商品或服務的過程（例如：網路購物）中，有蒐集或處理歐盟居民之個人資料者，即適用 GDPR（第 3 條）。
- (2) 保護客體：明文規範個人資料包括位置資訊(location data)、網路識別碼

---

<sup>11</sup> 歐盟的法令包括以下 5 種類型：(1) 規則 (Regulation)：對所有成員國具有約束力，一旦採納，即具有直接適用性（無需成員國內部批准程序即成為國內法律體系的一部分）；(2) 指令 (Directive)：要求各成員國為達成目標而新制或修訂國內法，承擔法律義務；(3) 決定 (Decision)：具有法律約束力的法律形式之一，其適用對象不是一般性的，而是特定的，可能是特定的成員國、企業或個人；(4) 建議 (Recommendation)：歐洲委員會建議成員國政府、企業或個人採取某些行動或措施。雖然沒有法律約束力或強制力，但被視為促進成員國內法制化或修訂的手段；(5) 意見 (Opinion)：有時也被稱為「見解」，歐洲委員會對特定主題發表的意見，沒有法律約束力或強制力。

(online identifier)等，亦即納入 IP 位址、GPS 定位等(第 4 條)。

(3) 當事人之權利：除以往之資料查詢、複製、更正及刪除權外，更進一步賦予當事人得請求資料管理者及處理者刪除連結（第 17 條：被遺忘權），為強化網路環境之被遺忘權，將刪除權擴張至「公開個人資訊」之控管者有義務通知「刻正進行個人資料處理」之控管者刪除任何該個人資料之連結、副本或複製品)、要求以可共同操作之格式提供資料（第 20 條：資料可攜權）等權利。當事人同意必須具體、明確、受充分告知，如單純沉默、預設選項為同意，則不構成同意；個人資料之處理係以直接行銷為目的時，當事人有權在任何時間、且毋需任何費用拒絕該處理。

(4) 跨境傳輸原則禁止：保障歐盟公民個資只能在如同歐盟對隱私高度保障的地區進行利用：

A. 就歐盟境內之個人資料原則禁止跨境傳輸至歐盟以外之地區或國家。例外須符合下列情形之一者，方得為之：擬傳輸地區經評估具備「適當保護水平」（第 45 條）；資料管理者已提供適當保護措施（第 46 條）；當事人明確同意、履行契約或依當事人要求，為締約前之必要措施、基於重要公共利益之維護、為主張、行使或防禦法律上之請求權所必要、基於保護當事人之重要利益所必要、依法辦理之登記作業，而向公眾提供資訊等（第 49 條）。

B. 評估第三國個人資料保護程度充足與否時，應考量有否獨立監管機關之存在並有效運作、已否參與或簽署關於個人資料保護之國際協定或其他具法律拘束力之契約、或參與多邊或區域體系而生之義務等。第三國資料保護程度不足時，應禁止向該第三國為資料移轉（第 45 條）。此時控管者或處理者應採取適當之保護措施以彌補第三國對資料保護之欠缺，該等措施可包括利用有拘束力之企業守則（Binding Corporate Rules, BCR）、歐盟執委會採行或核准之標準資料保護條款、由監管機關授權控管者或處理者與第三國或國際組織之個人資料控管者、處理者或接收者之契約條款等（第 46 條）。

(5) 資料管理者部分：

新增資料保護影響評估(DPIA)、資料保護長(Data Protection Officer, DPO)等制度：

A. 於特別使用新科技之處理方式，考量該處理之本質、範圍、使用情形及目的後，認為可能導致自然人權利及自由高度風險時，控管者應於處理前，實行

該處理對於個人資料保護之影響評估（第 35 條）。於下列情形應指定具資料保護法律與實務之專業知識的資料保護官或資料保護長：除法院行使司法權外，由公務機關或機構執行個人資料處理處理時；控管者或處理者需要定期且系統性地大規模監控（regular and systematic monitoring）資料主體時；大規模處理特殊類型個人資料或與前科及犯罪相關之個人資料時；歐盟或會員國相關法律有明確要求時等（第 37 條至第 39 條）。

B. 明定資料控管者負相關舉證責任：須證明當事人知悉同意之事實及範圍及須證明其已遵守個人資料處理之一般原則（目的限制原則、資料蒐集最少原則、正確性原則、完整及保密原則等）。

**(6) 資料外洩通報義務與提高處罰額度：**資料控管者一旦發現侵害個資事件，應於發現後 72 小時內向監管機關通報（第 33 條），遲延通報造成資料當事人損害，應負損害賠償責任（第 82 條）。違反者，最高得處以 2 千萬歐元或該企業之前一會計年度總營收 4% 之罰鍰（第 83 條）。

## 2. 歐盟 AI 法案

歐盟為確保 AI 系統安全、尊重基本權利、促進人工智慧投資、改善治理，並鼓勵建立統一的歐盟 AI 市場，為相關技術之投資與創新創造支持性環境，遂於 2021 年 4 月提出歐盟 AI 法案，係全球首部針對 AI 之全面性法規。該法案已於 2024 年 5 月間通過，並於 8 月 1 日生效，適用歐盟全體會員國。該法案規範重點如次：

### (1) 以風險為基礎：

採用基於風險的方法監管 AI，將 AI 引發之潛在風險區分不同級別，相關定義如下：

A. 最小風險：多數 AI 皆屬最小風險，例如支援 AI 之推薦系統和垃圾郵件過濾器，由於該等系統對人們之權利與安全風險極小，無須承擔法案規定之義務。

B. 有限的風險：在使用 AI 系統過程中，使用者面臨的主要風險來自 AI 系統缺乏透明度，爰第 50 條制訂透明化相規定，例如：必須能夠讓使用者瞭解其正與 AI 系統互動；AI 產生之內容必須被標記為人工生成；須讓使用者知悉其正處於生物識別分類或情緒辨識系統之使用環境等。

C. 高風險：依據第 6 條規定，倘 AI 系統作為產品之安全零組件，或該 AI 系統本身即為獨立產品且適用於法案附錄 1 所列舉之法規清單，即可視為高風險

AI。此外，法案附錄 3 所敘明有關生物辨識、關鍵基礎設施、教育或職業培訓、評估使用公共或私人服務資格等 AI 系統與應用。倘對自然人之健康、安全及基本權利構成重大損害，亦將被視為高風險 AI，例如用於招募、評估貸款資格等 AI，依據第 8 至 15 條相關規定，高風險 AI 需要遵守相關要求，包括具備風險管理系統、高品質之資料集(data set)、活動記錄、明確的使用者資訊、人工監督和高水準的準確性及 網路安全性等條件。

- D. 不可接受之風險：依據第 5 條規定，對人們基本權利構成明顯威脅之 AI 將被禁止，如忽視使用者自由意志並操縱使用者行為之 AI 系統或應用程序，包括使用語音輔助鼓勵未成年人危險行為之玩具、政府或公司可進行「社會評分」之系統，此外，生物識別系統之即時用途將被禁止，例如在工作場所使用情緒識別系統、對自然人進行分類之 AI、及在公共場合將遠程生物識別系統用於執法等應用，惟倘生物識別系統用於尋找失蹤者、預防對自然人之生命威脅、識別犯罪嫌疑人等情況，則准予使用。

## (2) 適用對象之義務：

對於高風險 AI 及具系統風險之通用 AI 模型，法案針對相關人士規定不同之義務如下：

### A. 高風險 AI：

- (A) 提供者之義務：依據第 16 條規定，高風險 AI 提供者除須確保其符合第 8 至 15 條有關風險管理系統、數據治理 (data governance)、技術文件、透明化、確保網路安全等規定外，其他相關義務包括依第 17 條規定建立品質管理系統，以書面形式記載法規遵守，符合評鑑、設計驗證等相關資訊；依法案第 18 條規定保存技術文件、品質管理系統等文件；依第 19 條規定保存 AI 之日誌檔(logs)；確保 AI 符合第 43 條有關符合評鑑之規定等。
- (B) 歐盟境外第三國之提供者：第 22 條規定歐盟境外第三國之提供者在其高風險 AI 系統投入歐盟市場前，應書面任命 1 名位於歐盟境內之授權代表，執行授權書所委託之事宜，可委託事宜包括驗證第 11 條規定之技術文件、驗證第 47 條規定載明之符合性聲明、向主管機關提供必要文件與資訊、配合針對其高風險 AI 採取之任何行動等。
- (C) 進口商之義務：依據第 23 條規定，高風險 AI 進口者須確認提供者已落實相關義務，例如：AI 系統之符合性評鑑、起草技術文件等；另須確保 AI 系

統應貼有 CE 標籤並附有第 47 條規定載明之符合性聲明等；在 AI 系統之包裝或隨附文件上註記名稱、註冊商標、聯絡地址等資訊；向主管機關提供必要文件與資訊，配合針對其高風險 AI 系統所採取之任何行動等。

- (D) 經銷商之義務：法案第 4 條載明相關義務，如在高風險 AI 系統投入市場前，經銷商應驗證該 AI 系統是否帶有所需的 CE 標誌，並附有第 47 條中提到的歐盟符合性聲明副本和使用說明；倘經銷商認為其高風險 AI 恐不符第 8 至 15 條相關規定，不得將其 AI 系統投入市場等。
- (E) 部署者之義務：依據第 26 條規定，相關義務包括部署者應採取適當之技術和措施，依照高風險 AI 之使用說明據以使用該 AI；監控該系統之運作，並適時提供運作相關資訊予供應者；保存 AI 系統之日誌檔 (logs) 至少 6 個月等。另依據第 27 條規定，倘部署者係受公法管轄之機構或提供公共服務之私人實體，且其使用之高風險 AI 用於評估自然人之信用分數或生命風險，則該部署者須於使用前評估對基本權利之影響。

#### B. 具系統風險之通用 AI 模型：

依據第 3 條規定，通用目的 AI (GPAI) 模型係指一種 AI 模型，包括使用大規模數據進行自我監督 (self-supervision) 訓練之 AI 型，該模型顯示出顯著的通用性 (generality)，且無論以何種態樣進入市場，皆能夠執行各種不同的任務，並可整合到下游各種系統或應用程式中，但排除在進入市場前用於研究、開發、試驗活動 (prototyping activities) 之 AI 模型。相關提供者之義務如下：

- (A) GPAI 模型提供者案第 53 條規定 GPAI 模型提供者應履行之義務，涵蓋撰擬技術文件、揭露訓練模型之簡要資訊國家主管機關進行必要合作
- (B) 歐盟境外第三國之提供者：第 54 條則規定歐盟境外第三國之提供者在其 GPAI 模型投入歐盟市場前，應書面任命一名位於歐盟境內之授權代表，執行授權書所委託之事宜，並履行法案第 53 條及 55 條所載之相關義務。
- (C) 具系統風險 GPAI 模型之提供者：第 55 條另規定具系統風險 GPAI 模型之提供者應履行之義務，除應履行前述第 53 條及 54 條所載之相關義務外，應依據技術演進評估模型，以減輕潛在風險，並對模型及其基礎設施提供網路安全保護措施，倘意外事件之發生，需向歐盟 AI 辦公室及國家主管機關報告發生情形。



### (3) 歐盟相關執行單位：

- A. 歐盟 AI 辦公室：法案第 56 條規定，歐盟 AI 辦公室應推動歐盟層級之行為準則(code of practice)，以促進本法案之實施，並同時考量國際作法，此外，其業務範疇應確保前述第 53 條及 55 條所載之相關義務有效落實，該辦公室並可邀請 GPAI 模型提供者及相關國家主管機關參與產業規範之制定過程；法案第 64 條明定，執委會應透過 AI 辦公室建立歐盟在 AI 領域之專家能力。
- B. 歐洲 AI 委員會：依據法案第 65 條規定，將設立歐洲 AI 委員會(European Artificial Intelligence Board)，由各會員國指派一位代表組成該委員會，依據第 66 條規定行使職權，如該委員會將負責各會員國主管機關在落實本法案及相關行政管理之協調、就法案實施提供建議、協助 AI 辦公室支援相關主管機關建立和發展 AI 監管沙盒機制、促進與第三國或國際組織間之合作、向執委會就 AI 國際事務提供建言等。
- C. 諮詢論壇：依據第 67 條規定，將設立諮詢論壇 (Advisory Forum)，以向執委會與歐洲 AI 委員會提供技術知識及建言，該論壇成員將來自產業、學術、中小企業、新創等，另歐盟基本權利署(The Fundamental Rights Agency)、歐盟網路安全局(ENISA)、歐洲標準委員會(CEN)、歐洲電工標準化委員會(CENELEC)和歐洲電信標準協會(ETSI)應為諮詢論壇的永久成員。
- D. 獨立專家科學小組：依據第 68 條規定，執委會應透過施行細則，制定成立獨立專家科學小組之相關規定，該小組由具備 AI 領域專業知識者組成，負責向 AI 辦公室提供有關落實本法案之相關建言。
- E. 國家主管機關：第 70 條規定，每個會員國應至少設立或指派一個通知機構及一個市場監督機構作為國家主管機關，以履行本法案賦予主管機關之權責，會員國並須向執委會提交機關資訊，執委會應推動各會員國主管機關間之交流。

### (4) AI 監理沙盒：

第 57 條規定，會員國其主管機關應在國家層級建立至少一個 AI 監管沙盒 (Sandbox)，並於 2026 年 8 月 2 日前投入運作，該 AI 監管沙盒機制將提供一個可受控環境，以促進創新和競爭力，並協助 AI 系統在投入市場前，可於有限時間內完成訓練、測試、驗證等階段性任務，推動 AI 生態：系統之發展，特別是協助新創及中小企業開發之 AI 系統進入歐盟市場，執委會並可為 AI 監管沙盒的建立和運作提供技術支援、建

議和工具；第 58 條進一步規定執委會應採認施行細則(implementing act)，以明確規定 AI 監理沙盒之建立、執行、監督等細節；此外，第 59 條另明列，可在 AI 監管沙盒機制中利用個資開發 AI 系統之適用情形，例如：減緩氣候變遷、疾病檢測、改善環境品質等目的。

#### **(5) 對中小企業及新創之措施：**

第 62 條規定會員國應針對在歐盟擁有註冊辦事處或分支機構的中小企業（包括新創），在符合資格前提下，提供可優先進入 AI 監管沙盒、使用訓練計畫等資源，並協助回復對本法案之疑問，另需減免中小企業依據第 43 條進行符合性評鑑之相關費用。此外，AI 辦公室應開發單一資訊平台，方便歐盟境內所有營運商 (Operator) 使用，並透過宣傳活動，提高公眾對本法案相關義務之認識。

#### **(6) 罰款：**

依據第 99 條規定，違反禁用 AI 應用相關法規之企業，其罰款最高可達該企業全球年營業額之 7%，違反其他義務之罰款最高可達 3%，提供不正確資訊之罰款最高可達 1.5%。

#### **(7) 法案實施時間表：**

依據第 113 規定，目前法案已於 2024 年 8 月 1 日生效，並將於 2026 年 8 月 2 日起適用於歐盟全體會員國，其中有關被禁止的 AI 行為相關規定，將另於 2025 年 2 月 2 日適用；有關 GPAI、歐盟執行單位、罰款等相關規定，將另於 2025 年 8 月 2 日開始適用；與高風險 AI 系統相關之義務規定，將另於 2027 年 8 月 2 日起適用。

#### **(8) AI 法案與 GDPR 關聯**

AI 法案與 GDPR 共通之處，就是課以義務的對象不僅止於總部設於歐盟 27 個成員國的企業，而及於任何在歐盟成員國境內展開業務的人。AI 於歐盟 GDPR 合法原則下的討論：包含偏見的演算法、收集資料的目的性與最小化、揭露黑盒子秘密的透明原則，以及資料主體的刪除權。有研究顯示，GDPR 對於 AI 發展的影響，包含：偏見的演算法 v.s 公平原則；AI v.s 資料目的限制原則；AI v.s 資料最小化原則；黑盒子(black box) v.s 透明度原則；經訓練的 AI 模型 v.s 刪除權（被遺忘權）。違反者將最高處以 2 千萬歐元或全球營業總額 4%之行政罰鍰。

### 3. 歐盟 AI 協議(AI Pact)

歐盟 AI 法案於 2024 年 8 月 1 日生效，惟該法案對於高風險 AI 等相關規定義務僅在過渡期結束後開始適用，爰歐盟執委會刻正推動 AI 協議 (AI Pact)，盼企業簽署該協議，承諾在前述過渡期結束前即自願履行相關規定義務。歐盟執委會於 2024 年 9 月 25 日表示已有包括來自資通訊、電信、醫療保健、銀行、汽車和航空等產業之跨國公司和歐洲中小企業，超過 100 家企業簽署該協議。依據協議內容，簽署企業必須履行：(1) 採用 AI 治理策略，以推動組織內對 AI 之應用，及致力遵守 AI 法案；(2) 識別依據 AI 法案可能被歸類為高風險之 AI；(3) 提高使用人員的 AI 素養和意識，確保具道德和負責任的 AI 發展等承諾。超過一半的簽署企業除將落實上述 3 項核心承諾，另做出包括確保人類監督、降低風險以及標註特定類型的 AI 生成內容等額外承諾。在 AI 法案完全適用前，歐盟仍持續鼓勵產業界簽署上開協議。

### 4. 歐盟值得信賴 AI 道德指引 (ETHICS GUIDELINES FOR TRUSTWORTHY AI)

歐盟委員會 (European Commission) 於 2019 年發布「可信任 AI 倫理指引 (Ethics Guidelines for Trustworthy AI)」，說明值得信賴的 AI 應具備 3 項要素及 7 項關鍵要求，分別係：

#### (1) 3 項要素

值得信賴的 AI 應該是：**A. 合法**：遵守所有適用的法律和法規。歐洲、國家和地區的許多具有法律約束力的規則，包括但不限於：歐盟主要法律 (歐盟條約及其基本權利憲章)、歐盟二級法律 (例如：一般資料保護規則)、聯合國人權條約和歐洲委員會公約 (例如：歐洲人權公約) 以及眾多歐盟成員國法律。此外，還包括適用於特定 AI 應用的各種特定領域規則 (例如醫療保健領域)。**B. 道德**：尊重道德原則與價值觀。**C. 穩健**：AI 系統應以安全、可靠和可靠的方式運行，並且應採取保障措施，防止任何意外的不利影響。

#### (2) 7 項關鍵要求：

- A. 人類機構與監督**：AI 應該在人類的監督下運作。
- B. 技術穩健性和安全性**：AI 需要具有彈性和安全性，必須有備援計畫因應 AI 出現問題時。
- C. 隱私和資料治理**：除了確保充分尊重隱私和資料保護外，還必須建立充分的資料治理機制，考慮到資料的品質和完整性，並確保資料的合法存取。

- D. **透明：數據、系統和 AI 商業模式應該透明**：建立可追溯性機制，並可向利害關係人解釋。人類需要意識到他們正在與 AI 交互，並且必須瞭解系統的功能和限制。
- E. **多元化、非歧視和公平**：必須避免不公平的偏見，因為它可能產生多種負面影響，從弱勢群體的邊緣化到偏見和歧視的加劇。為促進多樣性，AI 應該向所有人開放，並讓相關利益估線人參與整個生命週期。
- F. **社會與環境福祉**：AI 應該造福全人類，必須確保它們是可持續的和環境友好的。此外，也應該考慮到環境，包括其他生物，並且應該仔細考慮其社會和社會影響。
- G. **問責制**：應建立機制來確保 AI 及其結果的責任和問責。可審計性能夠評估演算法、數據和設計流程。此外，應確保提供充分且易於獲得的補救措施。

#### （四）荷蘭

荷蘭在 AI 技術的發展和監管方面也表現積極，特別關注如何平衡技術創新與社會倫理需求。荷蘭政府致力於建設“負責任的 AI”，以推動社會的信任與技術的可持續發展。

1. **數據與隱私保護**：荷蘭對於數據隱私非常重視，並依據歐盟「一般資料保護規則（GDPR）」制定相關措施，要求 AI 在使用個人數據時必須滿足最高標準的隱私保護要求。這不僅有助於保障個人權利，也可提升公眾對 AI 技術的信任。
2. **AI 倫理委員會**：荷蘭成立 AI 倫理委員會，負責審查和建議政府在 AI 技術上的政策方向，確保 AI 發展遵循倫理和社會責任。
3. **推行演算法影響評估及演算法登記冊制度，強化演算法透明度**：

##### （1）基本權利和演算法影響評估（FRAIA）

荷蘭政府非常重視 AI 演算法是否侵犯人權，要求荷蘭公共機構內使用演算法前，必須先進行基本權利和演算法影響評估（FRAIA），使團隊能夠在早期階段以結構化方式討論與演算法部署相關的所有相關關注點。目標是防止演算法在其後果尚未被充分理解的情況下被部署。荷蘭政府希望透過 FRAIA，相關各方在考慮是否開發演算法應用時能進行平衡的討論，也幫助公務員識別使用演算法的風險，並採取適當的措施來應對這些風險。FRAIA 包含許多關於需要討論的主題的問題，並且在政府組織考慮開發、委託開發、

購買、調整和/或使用演算法的任何情況下都必須制定答案，而回答這些問題須與單位內部和外部利害關係人進行協商。這類問題的範例包括：促使使用演算法的公共價值是什麼？使用該演算法以及根據該演算法做出有針對性的決策的法律依據是什麼？什麼類型的資料將用作演算法的輸入以及資料從哪些來源取得？

FRAIA 的 4 個評估步驟：第一：評估使用該演算法的原因、根本動機和預期效果。第二：說明演算法本身。詢問有關所使用的數據的問題，然後詢問演算法以及負責任地使用數據的條件。第三：演算法的實作和使用以及如何處理輸出。第四：是評估演算法是否損害基本權利。

## (2) 公共演算法登記冊

荷蘭政府於 2022 年 12 月 21 日發布公共演算法登記冊<sup>12</sup>，希望對政府使用的演算法進行合法檢查，以確定其是否存在歧視和任意性。荷蘭政府希望透過提高政府使用的演算法的透明度、為其使用設定明確的要求並確保適當的監督，並應為使演算法的應用和結果更可解釋性，做出重要貢獻。該登記冊係對外開放，每個人都可以看到哪些有影響力的流程使用了演算法。該資訊允許監控演算法。當人們不同意演算法的使用時，演算法寄存器會指出可以在哪裡提出異議。荷蘭資料保護局(Dutch DAP)於 2023 年 1 月起擔任演算法監管機構。目前，政府機構被鼓勵但沒有義務將其使用的演算法添加到註冊表中，惟因應歐盟 AI 法案生效，荷蘭登記冊將從 2025 年起為強制性。依據荷蘭公共演算法登記冊揭露，截至 2024 年 10 月 4 日止，超過 350 個荷蘭公部門已於該登記冊登記超過 500 個演算法。

---

<sup>12</sup> 網址：<https://algoritmes.overheid.nl/>

## 第四章 歐盟審計院、荷蘭審計院運用 AI 於審計業務概況及查核報告

審計人員可以在審計生命週期的風險評估、審計規劃、實地工作和報告等各個階段運用自動化能力。重要的是要正確看待自動化、分析和 AI：它們是工具，就像電腦一樣，它

圖 12 AI 運用於審計業務之分析面向



們不會取代審計人員，而是會改變審計過程。AI 可用於審計過程中對大量交易進行各種分析（如圖 12），這些技術在提高審計效果方面發揮著至關重要的作用，幫助專業人士識別風險、確保合規並根據數據，提出改善決策：<sup>13</sup>

本章將說明歐盟審計院、荷蘭審計院查核歐盟、荷蘭政府發展及運用 AI 情形、歐盟審計院、荷蘭審計院發展及運用 AI 於審計業務情形，以及 AI 與公平性審計間之關聯。

### 第一節 歐盟審計院、荷蘭審計院查核歐盟、荷蘭政府發展及運用 AI 情形

#### 一、歐盟審計院（European Court of Auditors, ECA）

##### （一）查核 AI 應考慮的要素及準備工作<sup>14</sup>

ECA 認為查核審計 AI，是為確保其完整性和有效性。隨著 AI 系統成為受審核單位

<sup>13</sup> 資料來源：SAI20(2023)，負責任的 AI 彙編。

<sup>14</sup> 資料來源：歐盟審計院「2024-2025 年人工智慧初始策略與部署路線圖（Artificial Intelligence

業務流程的一部分，所以必須對其進行查核。需要考慮的關鍵問題是，受審核單位是否運用 AI 於業務作業、是否制定 AI 策略並建立相關治理機制及如何解決運用 AI 伴隨而來的問題。表 7 提供查核人員一些查核參考要素。

**表 7 查核 AI 需要考慮的要素**

要素	內容
數據完整性和減少偏見	AI 模型的效果取決於其訓練所使用的數據。AI 系統的操作人員應確保所使用的數據是準確的、具有代表性的，且不存在扭曲產出結果的偏見。
遵守法規和標準	AI 系統必須遵守相關的法律和倫理標準，從而確保這些系統被負責任使用，保護敏感數據，並維護公眾信任。
透明性和可解釋性	為維持信任和問責，AI 系統應該具有透明性，其決策過程應具可解釋。
性能監控和評估	定期評估 AI 系統是必要的，以確保其有效運作並持續滿足其目標。
對員工和流程的影響	查核 AI 對受審核單位的員工和流程的影響。運用 AI 不應降低工作質量或導致員工喪失在關鍵流程和程序的技能。

查核 AI 系統的準備工作包括：專門的培訓提高審計人員和管理層的 AI 素養；建立 AI 特定的檢查清單和程序來調整審計框架；與制定審計標準和框架的專業協會及國際工作組織進行合作；持續學習，以確保查核方法隨著 AI 技術、法規和政策發展而調整。

## （二）查核歐盟投資 AI 執行情形

ECA 非常重視歐盟投資 AI 的情形，以促使歐盟未落後於世界大國。ECA 於 2024 年 5 月發布歐盟投資 AI 執行情形查核報告—「歐盟人工智慧的雄心—未來需要更強的治理和更集中、更多的投資 (EU Artificial intelligence ambition – Stronger governance and increased, more focused investment essential going forward)」<sup>15</sup>。謹摘述如次：

### 1. 查核緣起

AI 涵蓋快速發展領域的新興技術，包括機器人、大數據和雲端運算、高效能運算、

initial strategy and deployment roadmap 2024-2025)」

([https://www.eca.europa.eu/ECAPublications/ECA-AI-Strategy-2024-2025/ECA-AI-Strategy-2024-2025\\_EN.pdf](https://www.eca.europa.eu/ECAPublications/ECA-AI-Strategy-2024-2025/ECA-AI-Strategy-2024-2025_EN.pdf))

<sup>15</sup> 資料來源：ECA 全球資訊網 ([https://www.eca.europa.eu/ECAPublications/SR-2024-08/SR-2024-08\\_EN.pdf](https://www.eca.europa.eu/ECAPublications/SR-2024-08/SR-2024-08_EN.pdf))。

光子學和神經科學。美國長期以來一直是 AI 領域的領先者，而中國大陸計畫在 2030 年成為全球 AI 領導者，該兩方都依賴科技巨頭的大量私人投資。歐盟公、私部門於 2018-2020 年間投資於 AI 為 200 億歐元，並在接下來的 10 年中每年達到 200 億歐元。歐盟內使用 AI 的企業比例在成員國之間有顯著差異，歐盟目標係於 2030 年時，境內 75% 的公司使用 AI。本次報告核係 ECA 第一次對歐盟發展 AI 生態系統的有效性進行查核。查核結果為歐盟的 AI 計畫績效提供洞見，以利於歐盟未來發展 AI。

## 2. 查核範圍及方法

(1) **查核範圍**：ECA 評估歐盟在促進歐洲 AI 生態系統發展中的現有角色及下列行動之有效性，但 ECA 並未查核歐盟計畫中提到的 AI 人才和技能的發展及 AI 法案：

- 歐盟在 2018 年和 2021 年協調歐盟 AI 計畫、進行監管改革以促進 2018-2023 年期間歐盟在數據和可信 AI 方面的投資的行動。
- 於 2018 年歐盟 AI 計畫通過後，歐盟資助的措施的執行情形。
- 2014-2022 年間的歐盟資助的 AI 研究與創新的執行情況。

(2) **查核方法**：

- 訪談：ECA 審查歐盟執委會的內部、公開文件及相關數據，並與官員進行訪談；與 OECD AI 觀測站和美國聯邦審計總署（科學、技術評估和分析團隊）的代表討論國際標準。
- 問卷調查：對 27 個國家負責協調 AI 政策的機關進行問卷調查（收到 20 份回覆），並訪談 3 個國家（比利時、芬蘭和西班牙）的相關機關。
- 隨機抽樣：10 個由「地平線 2020 (Horizon 2020)」計畫<sup>16</sup>資助的已完成 AI 研究項目，這些項目涉及環境、智能移動和工業機器人等歐盟 2021 計畫中的優先行業。目的在檢視歐盟執委會在宣導和運用方面的做法。ECA 對計畫受益者進行了現場訪問。為獲取對私部門參與地平線計畫的回饋，ECA 訪問相關公私合營夥伴關係（大數據、機器人和 AI、數據與機器人）的代表。

---

<sup>16</sup> 「地平線 2020 Horizon 2020」計畫是歐盟 2014 年至 2020 年的研究和創新資助計畫，預算近 800 億歐元。



### 3. 查核發現

#### (1) 歐盟協調和監管歐盟對 AI 投資的框架仍在進行中

- 歐盟 AI 計畫的投資目標過於模糊。
- 歐盟委員會與成員國的協調效果有限。
- 歐盟雖已採取行動實現單一數據市場，惟其仍處於初始階段。
- 自 2018 年以來，歐盟逐步採取措施，以建立 AI 的監管框架。

#### (2) 歐盟確定了 AI 創新的促進因素，但仍待推動

- AI 計畫對創新者提供資本的規模有限。
- 歐盟資助的中小企業 AI 基礎設施雖解決重要需求，但有延誤，且各項配套措施的協同效應尚未顯現。

#### (3) 歐盟執委會雖增加對 AI 研究與創新 (R&I) 的投資金額，但對成效缺乏全局瞭解

- AI 研究與創新投資 (R&I) 缺乏協調和評估框架。
- 歐盟執委會在 AI 研究與創新 (R&I) 結果的應用和推廣作為與計畫的相關性不足。

### 4. 查核建議

根據上開查核發現，ECA 建議歐盟執委會：

- (1) 重新評估歐盟的 AI 投資目標。
- (2) 評估是否持續支援中小企業資本，以協助中小企業專注於 AI 創新。
- (3) 確保歐盟資助的 AI 基礎設施能妥善運作。
- (4) 制定 AI 研究和創新支出的績效目標和指標，並定期監測執行情形。
- (5) 加強運用 AI 研究成果。

## 二、荷蘭審計院（Netherlands Court of Audit, NCA）

### （一）荷蘭中央政府 AI 運用情形之查核<sup>17</sup>

由於歐盟 AI 法案已於 2024 年 8 月間通過，NCA 為瞭解荷蘭中央政府機關 AI 是否按預期運作以及符合歐盟 AI 法案要求，針對該國 70 個機關之 AI 使用情形及目的地進行查核，並希望荷蘭政府自我檢視現行使用之 AI，是否存在歐盟 AI 法案於 2025 年 2 月開始禁止使用之類型，及早妥為因應。NCA 於 2024 年 10 月 16 日發布之查核結果。

#### 1. 查核目標

- 目前使用哪些 AI 系統？
- 正在進行或已經進行哪些 AI 實驗？
- 使用 AI 系統有哪些機會，是否已經產生？
- 如何評估和減輕運用 AI 衍生的風險？

#### 2. 查核技術與方法

- （1）問卷：請受審核機關完成一份關於其 AI 系統的問卷，包括目前使用的 AI 系統以及他們正在試驗或在過去 5 年中試驗過的 AI 系統。請受審核機關簡述所使用 AI 系統的狀態、結果和根據歐盟 AI 法案的預期風險分類。
- （2）訪談：對部分受審核機關進行深入訪談，瞭解他們預期使用 AI 帶來的機遇和風險。調查的重點是 AI 系統的部署。

#### 3. 查核發現

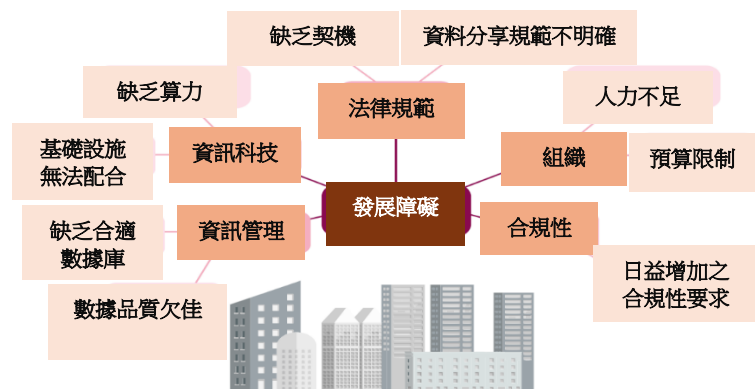
- （1）受審核之 70 個機關表示正在使用或曾經使用 433 個 AI 系統，且大多自行開發（in-house），目前中央機關使用 AI 尚不普遍。大多數系統（167 個）仍處於實驗階段，且大部分機構（88%）使用的系統不超過 3 個。僅有 5% 的系統被公布於演算法註冊表中。使用 AI 最多的機構是警察和員工保險局（UWV），分別使用 23 個和 10 個系統。

---

<sup>17</sup> 資料來源：荷蘭審計院全球資訊網。<https://english.rekenkamer.nl/publications/reports/2024/10/16/focus-on-ai-in-central-government>

- (2) 中央機關不確定其許多 AI 系統是否按預期運作，對於其一半以上的 AI 系統，並未權衡其機會與風險。
- (3) 約三分之二的 AI 系統主要用於改善內部流程，分析和處理大量資訊，舒緩人力短缺，而其運用並未對公民和企業產生直接影響。自動將語音轉換為文字或即時匿名化文件的 AI 技術可以為政府節省大量時間和資金。相比由公務員手動操作，自動化 AI 系統進行內部文件分析的效率更高，成本也更低；另外三分之一的 AI 系統對公民和企業有直接影響，最常見的是檢查和執法程式（82 個系統）。一些檢查和執法組織使用 AI 風險模型，根據歷史數據預測未來犯罪的風險，進而利用這些預測用於選擇公民和企業進行額外檢查。
- (4) 受審核機關經常不知道其 AI 系統是否運作正常，此外 35% 的受審核機關不清楚運用 AI 是否達到預期目標，肇因該機關並未設定使用 AI 可以達成的目標，或是根本無法確認系統是否有效運作。更弔詭的是，根據受審核機關表示，已停止使用之 141 個 AI 系統中，有 82 個（58%）達到預期目標或表現更好，惟因缺乏進一步開發的資源等原因，而被停止使用。
- (5) AI 的使用帶來機會，卻也衍生風險，而這些風險仍必須滿足今年發布的歐盟 AI 法案規範。根據歐盟 AI 法案，所有 AI 系統都必須進行風險評估，具有不可接受風險的 AI 系統將被禁止，未來高風險系統則需滿足額外的條件。雖然荷蘭行政機關大多將其使用之 AI 系統分類為“最低風險”，但這並不代表沒有風險，並未反應實際，仍可能存在侵犯隱私、資訊安全薄弱或對公民和企業造成不公平損害的風險。

圖 13 荷蘭中央機關發展 AI 遇到之障礙態樣



- (6) 阻礙中央機關發展 AI 潛力的障礙，例如：缺乏專業知識和能力、不確定的數據共享法律法規、不合適的基礎設施以及越來越重的合規負擔等（詳圖 13）。

## （二）荷蘭政府 AI 演算法之查核<sup>18</sup>

由於荷蘭使用 AI 之行政機關日益增加，AI 演算法管理成為政府服務營運管理中重要的一部分。預計這種發展將在未來幾年繼續下去。與私人公司一樣，政府正在越來越多地使用演算法來實現工作自動化、解決問題和做出預測。而政府機關使用演算法也帶來了一些挑戰，例如：

- 透明度不足：演算法在中央政府中的運作方式及其對政府行動的影響可能不夠清楚，或者無法清楚地向公眾解釋。這可能與所使用的技術（例如：神經網路），或者是由於其複雜性（例如：演算法可能涉及太多變數）有關。
- 偏見與歧視風險：演算法或其使用的數據集可能含有某些偏見，導致產生歧視的結果。雖然人類也有內在的偏見，但使用演算法時風險在於其可能主要依賴程式員或數據科學家的決策（例如，所使用的數據）。程式員或數據科學家可能缺乏關於特定情境的具體知識與經驗，例如對一個補助申請的決策。
- 學習過程的不可預測性：演算法從數據中學習時，往往難以預測它具體會學到什麼，或難以預見可能出現的不被接受的學習效果。數據中的某些相關性可能會導致演算法產生歧視的結果。
- 外部供應商的依賴性：中央政府使用的許多演算法來自外部供應商，包括內置演算法的 IT 系統。這些演算法所使用的確切數據和機制通常由供應商擁有，供應商可能會希望保護這些資訊。在涉及責任或個人數據處理等方面時，政府不能或不願完全依賴供應商提供的資訊，使得政府在分析和管理與演算法相關的風險時更加困難。

### 1. 查核目標

- （1）中央政府及其相關組織在哪些業務及流程中使用演算法，存在哪些類型或類別的演算法，以及使用演算法所帶來的風險和影響是什麼？
- （2）中央政府及其相關組織如何管理演算法的運行並控制其質量？

### 2. 查核技術與方法

---

<sup>18</sup> 資料來源：荷蘭審計院全球資訊網

(<https://english.rekenkamer.nl/publications/reports/2021/01/26/understanding-algorithms>)。

NCA 參考政府內、外部專家意見開發一個審計框架，以評估演算法在實踐中的品質和負責任的使用，並發現演算法的潛在弱點。NCA 擇選荷蘭政府 3 個演算法，並依該框架從下列 5 個面向評估：

- (1) **治理與問責 (Governance and Accountability)**：界定責任和專業知識、演算法生命週期的管理、使用演算法的風險因素，以及與外部利益相關者就責任等方面的協議。
- (2) **模型與數據 (Data and Model)**：模型與數據標準處理有關數據質量以及演算法的模型的開發、使用和維護的問題。這些標準包括關於數據中可能存在的偏見（從倫理角度出發）、數據最小化以及模型輸出是否經過測試的問題。
- (3) **隱私 (Privacy)**：部分演算法使用特殊類別的個人數據，必須遵守有關個人數據處理的法定規定。「一般資料保護規則 (GDPR)」是此審計框架的重要參考來源。
- (4) **IT 一般控制 (IT General Control, ITGC)**：ITGC 是組織為確保其 IT 系統可靠且符合道德標準而採取的控制措施。這些控制措施包括傳統的 IT 控制，例如：訪問權限管理、持續性和變更管理。此審計框架中納入的 IT 一般控制專注於與演算法相關的日誌數據、訪問權限和密碼管理。用於 IT 一般控制的主要標準是國際 ISO/IEC 27002 標準和政府資訊安全基準。
- (5) **倫理 (Ethics)**：倫理不僅是評估演算法的單獨方面，而是上述 4 個方面的組成部分，倫理與所有 4 個方面都息息相關。

### 3. 查核發現

#### (1) 在治理與問責面向

受審核的演算法在遵守治理與問責要求的程度上存在差異。此外，大部分演算法並未採取生命週期管理系統，在設計和實施階段花費大量的時間，但在其維護和持續運行方面卻不然。維護預算不足、維護不當或人力資源不足，最終可能導致演算法無法滿足新的倫理或法律標準。

#### (2) 在模型與數據面向

與模型和數據相關的問題包括演算法模型設計的方法和數據質量。在數據管理方面存在 2 個潛在風險：

●**歷史數據的使用**：歷史數據可能無法反映某些社會變化，這意味著過去的做法可能沒法被應用到當前的情境中。例如：什麼樣的能力是好的經理應具備的？這個問題的答案隨著社會趨勢的變化而變化。

●**數據偏見**：如果特定的族群在過去受到不同的對待，該演算法將會吸納這種偏見。

NCA 對 3 個演算法的分析顯示，並非所有相關專業領域的人士都參與了演算法的開發。雖然隱私權專家、程序員或數據專家通常會參與，但法律專家和政策顧問往往被排除在外。這可能導致演算法無法遵循所有法律和倫理標準，或未能推進相關的政策目標。

### (3) 在隱私面向

歐盟一般資料保護規則 (GDPR) 是隱私和數據保護的主要監管框架。根據審計框架對演算法進行了測試。隱私方面涉及的要點包括 GDPR 的個人數據處理登記、隱私影響評估、使用數據的法律基礎和數據最小化。NCA 查核的 3 個演算法，認為適用於演算法的隱私要求方面的合規性各不相同。在其中一個演算法的案例中，隱私政策、所使用的數據和演算法未公開提供足夠的詳細信息。這對於民眾來說非常重要，因為他們需要知道使用了哪些數據、演算法是如何運作的、如何影響他們。隨著數據使用量的增加和演算法的複雜性上升，這將成為未來更加重要的問題。在 NCA 評估的演算法案例中，發現民眾個人無法輕易獲得有關中央政府使用的演算法和數據的信息。民眾個人提交的個人數據和信息屬於他們，並且他們必須知道他們的數據是如何被使用的。數據處理登記並非在所有情況下都公開可用，演算法相關的隱私聲明並不非這麼清晰易懂。雖然在某些情況下，演算法的運作和所使用的變數已在立法中明確規定，但這些資訊卻不易閱讀或理解。

### (4) 在 IT 一般控制 (ITGC) 面向

ITGC 主要涉及訪問權限及其管理、資料備份。在 NCA 評估的 3 個演算法，其中 2 個演算法幾乎沒有資訊表明是否符合 ITGC 相關標準，而受審核單位亦無法提供足以證明其能夠充分控制相關風險之資料。NCA 認為這有 2 個原因：(1) 演算法由外部服務提供商管理。儘管相關官員假設這些外部服務提供商有適當的 IT 控制，但他們並不知道實際情況如何。當 NCA 要求提供證據時，相關部門的官員無法提供，或者無法在短時間內提供；(2) 雖然該機關已設置更高或不同的 ITGC 標準，但這些標準納入演算法的具體實務運作。

### (5) 在倫理 (Ethics) 面向

NCA 界定倫理的 4 個部分，分別係：尊重人類自主性、預防傷害、公平性、可解釋性和透明性。謹分述如次：

#### **A. 尊重人類自主性**

查核發現，這 3 個演算法作為輔助資源運作，該演算法並不（或尚未）做出任何自動化決策。

#### **B. 預防傷害**

為防止任何損害，演算法必須始終執行其預期的功能。此外，必須保護人們的隱私和相關數據。未經授權的訪問可能導致數據被更改、損壞或丟失。查核發現請詳前述 ITGC 說明。

#### **C. 公平性**

公平性意味著演算法考慮到人口的多樣性，不對某些個體、群體或其他實體進行歧視。如果不採取有效措施，演算法可能會在某些方面對特定個體或群體產生不良的系統性偏見。NCA 評估的 3 個演算法，其中 1 個案例是由外部供應商測試演算法是否存在不良結果。在另 1 個案例中，外部供應商會事先測試所有數據，以評估該數據是否絕對必要，從而使演算法達到其目的。

#### **D. 可解釋性和透明性**

演算法的擁有者有義務解釋其設計演算法的方式以及它的運作原理。所有 3 個演算法都具有可解釋性，且在這 3 個案例中，模型設計者努力在可解釋性和性能之間取得平衡。為了能夠解釋這些程序，管理當局必須清晰地記錄它們的運作。

### **4. 審核建議**

#### **(1) 發布清晰、一致的定義和質量要求**

NCA 建議內閣採用一套清晰、一致的術語和具體的演算法質量要求。清晰、一致的定義和質量要求將促進知識共享、簡化流程並防止誤解。

#### **(2) 告知社會大眾有關演算法的信息並解釋如何獲得進一步信息**

NCA 建議對於政府行為或針對特定案件、個人或企業的決策具有重大影響的演算法，內閣使社會大眾能夠獲悉哪些數據用於這些演算法、演算法如何運作及它們的結果會產生什麼影響，可能的做法係建置類似於提供有關大型 IT 項目的信息的儀表板。

### (3) 制定記錄有關使用演算法的契約條款及有效的合規監督流程

NCA 建議內閣確保充分記錄演算法的工作範疇、組織、監控和評估，因為這清楚地表明演算法是否適合其目的並能持續符合要求。尤其是在演算法係外包或從外部供應商購買取得情況下，確保所有有關責任的要求以契約形式律定清楚。

### (4) 將審計框架轉化為演算法的品質要求

NCA 建議內閣確保每個行政機關將審計框架轉化為一套實用的設計標準或演算法開發的品質要求。

### (5) 力求所有相關專家參與演算法的開發

NCA 認為演算法的開發涉及所有相關的學科和專業知識類型，爰建議行政機關除邀請技術專家之外，法律專家、倫理專家和政策顧問也應該參與其中。

### (6) 確保現在及未來提供清晰的信息，說明 IT 一般控制 (ITGC) 的運作

NCA 建議行政機關確保與演算法相關的官員獲得並保留有關 ITGC 質量的訊息，可以透過要求管理演算法的一方提供 IT 審計報告等正式聲明，證明 ITGC 符合相關規範。

## 第二節 歐盟審計院、荷蘭審計院發展及運用 AI 於審計業務情形

### 一、歐盟審計院 (European Court of Auditors, ECA)

ECA 認為 AI 應用協助審計工作，增強審計能力，使審計工作更加及時、高效且基於合理的依據，但 AI 不會取代審計人員的批判性思維和專業判斷。ECA 為使在部署和使用 AI 系統時堅持最高的倫理標準，防止偏見，優先考慮透明度，並保持謹慎和專業懷疑精神，忠實於獨立性、客觀性、誠信、透明度和專業精神的核心價值觀，並利用 AI 推動問責、透明度和信任的助推器，以最終提升歐盟的整體信譽，遂於 2023 年 10 月發布「2024-2025 年人工智慧初始策略與部署路線圖 (Artificial Intelligence initial strategy and deployment roadmap 2024-2025)」<sup>19</sup>。謹將內容摘陳如次：

---

<sup>19</sup> 資料來源：歐盟審計院全球資訊網([https://www.eca.europa.eu/ECAPublications/ECA-AI-Strategy-2024-2025/ECA-AI-Strategy-2024-2025\\_EN.pdf](https://www.eca.europa.eu/ECAPublications/ECA-AI-Strategy-2024-2025/ECA-AI-Strategy-2024-2025_EN.pdf))。

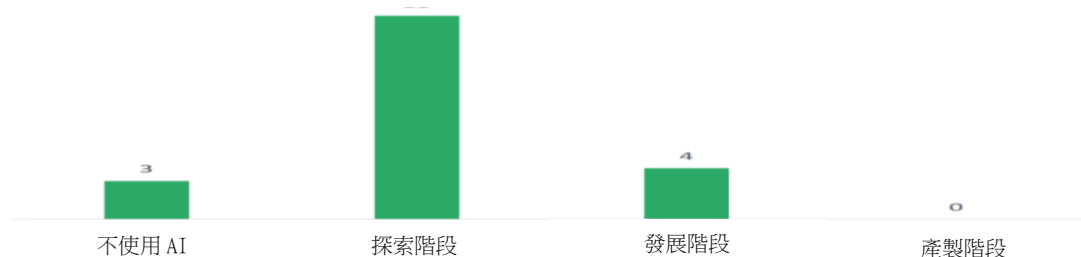


## (一) 目前歐洲國家最高審計機關發展/運用 AI 現況

### 1. 推行階段

ECA 於 2024 年 2 月向歐盟國家 27 個最高審計機關 (SAI) 調查，瞭解發展 AI 現處階段，共收到 23 份回覆，在受訪的 SAI 中，有 16 個 (70%) 目前處於探索階段，4 個 (18%) 正在開發一些 AI 工具 (詳圖 14)。

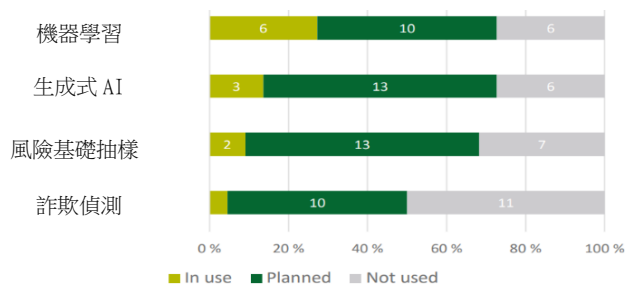
圖 14 歐盟 SAI AI 推行階段自我評估結果



### 2. 運用 AI 技術面向

有 8 個 SAI 表示已經在使用一些 AI 技術。大多數 SAI 規劃在未來使用這些技術，1 個 SAI 表示正使用生成式 AI 協助編寫電腦程式碼 (如圖 15)。

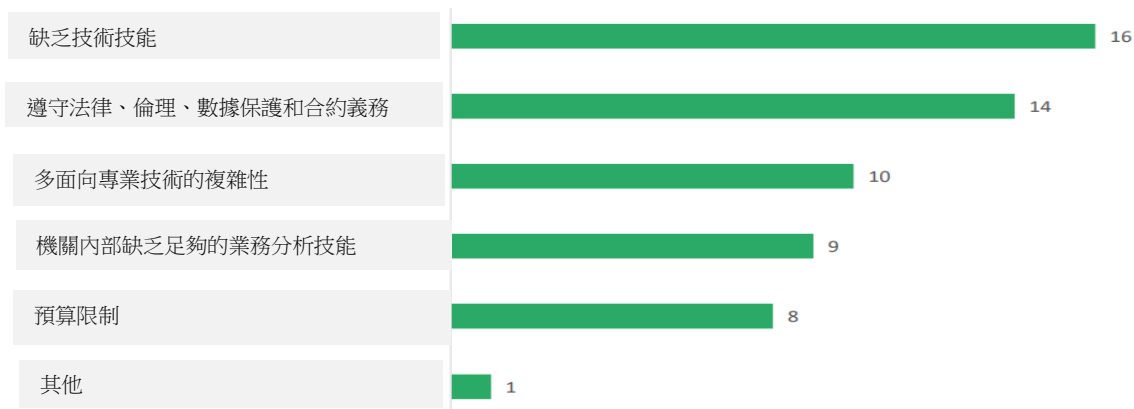
圖 15 歐盟 SAI 正使用或規劃使用特定 AI 技術情形



### 3. 在推行 AI 時遇到的挑戰類型

受訪 SAI 遇到最多的挑戰是缺少技術技能、其次依序為符合法規、倫理、資料保護等 (詳圖 16)。

圖 16 SAI 推行 AI 遭遇挑戰的類型



## （二）發展 AI 目標

### 1. 透過 AI 增加查核工作效率

透過有效利用 AI，逐步提升審計程序的效率，改善審計人員分析能力，並促進更加靈活的審計方法。

#### （1）提升 ECA 審計人員如何使用 AI 的技能

審計人員具備足夠的 AI 素養是成功部署 AI 工具的必要前提。為充分發揮 AI 在審計流程中的功能，必須提升審計人員的 AI 素養。ECA 開始實施全面的培訓計畫，提供對 AI 技術及其運作的基本理解、實用的指導和技術。

#### （2）確保 ECA 在技術上為 AI 做好準備

ECA 將對現有系統和流程進行全面的審查和改進，包括升級技術基礎設施以具備 AI 能力、修訂政策納入 AI 倫理和治理，以及調整程序優化 AI 的實施。最近，國際標準化組織（ISO）發布一項有關 AI 管理系統的新標準，ECA 考慮實施與其相關的部分。

#### （3）引入 AI 協助審計流程

ECA 將識別並實施最先進的 AI 工具，能夠處理和起草報告、高效分析大型數據並提供有別於傳統審計方法的建議，協助審計人員提高審計效率，使其能夠進行更深入的審計。

### 2. 建立基於 AI 的項目、系統和流程的審計能力

ECA 致力於提升其基於 AI 的項目、系統和流程的審計能力和量能，以便在變化多端且具有挑戰性的環境中提供強有力且可靠的證據。這一目標與 ECA 數位轉型策略相一致。隨著 AI 在歐盟的金融和政策倡議中開始應用，必須相應地發展 IT 審計方法和工具。

#### （1）提高 ECA 審計人員對 AI 工作原理及其相關風險的理解

第一步是確保審計人員深入瞭解 AI 技術及其使用方式，並認識到這些技術所帶來的風險。ECA 將啟動針對性的培訓和工作坊，以加深審計人員對 AI 演算法、數據依賴性和潛在偏見的知識。ECA 將訓練審計人員具備洞察力，使其能夠質疑和批判性評估 AI 系統。

## (2) 將 AI 方面納入審計方法論

ECA 將更新和擴展 IT 審計方法論以及 AWARE<sup>20</sup>中的其他指引，納入特定於 AI 的問題。這些將涵蓋數據質量、演算法透明度、道德使用以及遵守歐盟法規和標準等方面。這應該能夠更澈底和有根據地評估基於 AI 的過程和輸出。

## (3) 為歐盟及國際上有關 AI 的議題並做出貢獻

ECA 將積極參與有關 AI 的專業討論，尤其是在審計方面。這一目標有助於將 ECA 保持在公共審計專業的前端。

## 3. 將 ECA 在 AI 方面的知識提供給歐盟成員國最高審計機關

根據與同行合作的策略，承諾將在 AI 方面累積的知識和經驗隨時提供給其他尋求優化 AI 部署的歐盟成員國最高審計機關。將分享已被證明有效的最佳實踐、指導方針和工具，並將其放在 AWARE 的公共版本中可供使用。透過建立資源庫並促進工作坊和知識共享會議，幫助其他最高審計機關駕馭在審計中使用 AI 的複雜性。

## (三) 進行發展/運用 AI 風險評估

ECA 認為 AI 雖然有助於審計工作，但必須充分意識相關的風險，並確保有相關減緩風險的配套措施。ECA 列出針對部署 AI 的一般風險及運用於審計工作的風險，及其減緩風險配套措施（附錄 7、附錄 8）。謹摘述如次：

### 1. AI 的一般風險

#### (1) 有限的透明度和對 AI 邏輯的監管

許多 AI 系統，尤其是基於深度學習的系統，由於其複雜且不透明的內部運作過程，被視為「黑箱」。有時很難理解和解釋系統生成結果所經歷的過程。此種情況容易限制 AI 系統的透明性，並易招致對其審計成果的可靠性提出質疑。ECA 認為可以透過 A. 選擇運作方式更易於理解和跟蹤的 AI 模型；B. 確保模型可提供產生結果的過程或軟件代碼的逐步描述，從而允許監督和重新執行；C. 為審計人員提供培訓和方法指導，使其理解輸出結果及其局限性，確保模型提供用於生成特定輸出的代碼和/或過程步驟，記錄

---

<sup>20</sup> 查核方法及指引線上平台 (AWARE)：ECA 將審計資料整合至 AWARE，可依查核選案、規劃、查核方法及報告撰擬等步驟，並亦可以合規、財務及績效等面向，提供審計人員查詢參考。

AI 輸出在審計工作中的使用方式等減輕這類風險。

## (2) 偏見及歧視的產出

AI 系統輸出結果的質量取決於用來訓練系統的數據，系統可能會內化訓練數據中存在的潛在偏見，導致有偏見甚至歧視性的輸出，進而影響審計成果的可靠性。因此，制定有關 AI 輸出可接受使用的明確規定，並為審計人員提供培訓是非常重要的。此外，ECA 認為可透過定期重新評估使用的 AI 模型，並優先選擇訓練數據偏見較少的模型、使用基於 ECA 知識庫輸出的模型，而不是依賴於原始訓練數據中的知識等方式減輕此類風險。

## (3) 隱私及資料保護議題

AI 系統涉及大量數據的使用，可能包括敏感或個人資訊，如果可能，應優先選擇位於歐盟的供應商。

## 2. 部署 AI 的業務風險

### (1) 取得及維持成本太高

ECA 的 IT 預算有限，無法大規模採購或維持最佳的 AI 系統，且 ECA 內部可能缺乏將初步概念驗證轉化為所有審計人員都能使用的產品的技術能力。為克服此困境，ECA 將繼續探索跨機構合作的可能性，分享技能、基礎設施、解決方案等。ECA 規劃建立一個 AI 能力中心，整合來自 ECA 不同部門的專業知識和技能。同時，如果取得的 AI 雖非業界最先進，但也可以考量一些成本較低或資源需求較少的模型仍可能足夠滿足 ECA 業務需求的 AI 系統。

### (2) AI 模型的快速過時

由於過去兩年 AI 領域的發展非常迅速，存在技術發展速度快於審計機關向審計人員提供端到端服務能力的風險。當啟動發展專案時，往往會使用當時最有前景的技術，但當發展成品提供給審計人員時，可能會被認為已經過時而不被使用。影響部署速度的因素包括：透過試點項目驗證多個使用案例；確保遵守法規進行安全評估並制定安全計畫；以及調整作業流程基礎設施以適應這項新技術。可以透過保持靈活的方式和適當的基礎設施來減輕這些風險，同時並行探索不同選項，並為每個使用案例選擇最合適（且風險最小）的方案。對於市面上商業產品，這些風險的相關性較小，因為升級到更先進的模型可能只需購買新的許可。優先選擇歐盟供應商和開源模型，還可能有助於更快地獲得法律和數據保護方面的批准。

### (3) 使用 AI 造成聲譽損害

為避免私人未經授權使用 AI 的風險，應該為審計人員提供安全、性能良好的 AI 工具。此外，在審計工作中，對 AI 的誤用或缺乏應有的謹慎可能會對審計機關的聲譽造成損害。因此，為審計人員提供清晰的指導和充分的培訓是不可少的。

### (四) AI 發展路徑圖

路線圖基於兩個原則：1. 優先考慮可以在市場上獲得，或由其他機構或跨機構工作組開發；2. 當需要將 AI 應用於非常特定的審計業務，或當數據和信息的敏感性要求對工具進行嚴格控制時，則由 ECA 自行開發。

路徑圖 6 項行動方案：

#### 1. 制定詳細的溝通計畫據以執行。

這項策略的主要目標之一是建立一個共享的 AI 框架，因此，該策略和提出的路線圖必須在 ECA 內部及外部進行妥善的溝通。溝通計畫的目標是保持 ECA 對 AI 的知識與時俱。ECA 預計針對不同的溝通對象使用多種管道進行溝通，溝通計畫詳表 8。

表 8 ECA AI 溝通計畫

溝通對象	溝通訊息	溝通形式	溝通頻率
ECA 資深管理階層	進展狀態，高層次問題	質量保證/決策文件	每 6 個月
ECA 審計人員	有關 AI 應用、倫理以及在審計工作具體案例中的利用的基本知識。	主題簡介	2 次/年
	AI 的成就、新工具、能力中心等	ECA 新聞	4 次/年
	關於 AI 的一般訊息	AI 知識文件	每週
	AI 能力中心活動的協調	ECA AI 網絡：作為能力中心的一部分，在 Teams 上建立的一個對所有 ECA 審計人員開放的論壇	每天
歐盟最高審計機關	合作與知識交流	歐盟網絡活動	2 次/年
世界各國最高審計機關	ECA 的 AI 活動	特定媒體（例如 INTOSAI 期刊、EUROSAI 雜誌）和/或活動。	1 次/年

擴展的專業網絡	有關 AI 在審計中的一般資訊	個人的 LinkedIn	隨時
其他機構的 AI 決策者	比較 ECA 與其他機構的做法。	雙邊會談	隨時

## 2. 為 ECA 審計人員研擬 AI 培訓路徑

數據團隊與培訓團隊合作，專門針對 ECA 在 2024 年的具體需求，於 2023 年 12 月成立一個新的 AI 培訓方案，法律服務部門也提供有關 AI 法律風險和版權方面的課程。與藉由 EU-Learn 提供的培訓計畫相比，數據團隊專業人士組織的課程目的係精確滿足 ECA 審計人員（審計人員、幕僚單位等）的需求，並利用數據團隊在審計任務中工作的經驗。培訓後收到的一致正面回饋，顯示對數據團隊成員提供的培訓的認可，促使 ECA 在 2024 年和 2025 年進一步擴展這一培訓路徑。一旦 ECA 獲得生成式 AI，將為審計人員提供有關其使用的特定培訓。ECA 目前正在評估增加由內部專家主講 AI 倫理的特定課程。

## 3. 強化提供數據服務

AI 不一定是一項獨立的服務，而是一種可以提高審計機關各種現有業務生產力的促進因素。數據團隊是早期採用者之一，他們預計增強現有服務，以納入 AI、協助審計人員評估如何在其戰略和日常業務營活動中納入 AI（具體項目，詳表 9）。

表 9 增強提供的數據服務項目

服務項目	結合 AI 增加服務
資訊科學	<ul style="list-style-type: none"> <li>● 使用 AI 分析非結構化數據（基於風險的樣本提取）。</li> <li>● 使用 AI 對項目列表進行分類和/或從一般描述中識別相關問題。</li> <li>● 引入機器學習以在大型數據集中進行異常檢測。</li> <li>● 在數據科學工具集中整合生成性 AI。</li> <li>● 在支援審計任務時，利用生成式 AI 代碼助手提高數據團隊的生產力。</li> </ul>
IT 審計	<ul style="list-style-type: none"> <li>● 對 AI 系統進行審計，並提出未來的 AI 審計工作以獲得專業知識。</li> <li>● 透過添加與 AI 相關的風險和控制，提升當前的 IT 審計方法論，並起草 AI 系統的標準審計計畫。</li> <li>● 研擬內部審計 AI 系統的指導方針。</li> </ul>
自動化	啟動試驗，為 ECA 的機器人添加 AI 能力，

	例如分析發票、分類電子郵件等。
調查建議	測試基於 AI 的調查分析工具，並為審計人員提供使用指南。
培訓、輔導與主要利益相關者的聯繫	<ul style="list-style-type: none"> <li>● 積極參加歐洲最高審計機關組織和國際最高審計機關組織工作小組。</li> <li>● 探索與專業協會（例如 ISACA）合作的機會，以提高 AI 審計能力。</li> <li>● 數據團隊將針對可能出現在查核工作中的具體需求提供有關 AI 的專門輔導課程。</li> <li>● 數據團隊將提供 AI 專業知識，以支持與各自國家審計機關合作。</li> </ul>

#### 4. 建立新工具或強化工具的項目提案組合

準備一系列詳細的項目提案，每個提案目的在利用新技術的潛力或顯著強化現有的 IT 工具。這些提案將對前能力、資源以及來自審計人員和非審計人員的具體需求進行徹底分析後制定。

##### 4.1 建立“審計草稿助手”工具

該 AI 助手可以根據審計人員提供的想法或概念列表撰寫文本段落；校對文件以提高英文撰寫質量，並使其符合 ECA 的格式風格；概述文件並找出它們之間的聯繫。初步測試顯示，只有非常大型的語言模型（LLM）才能有效處理英語語言。相比之下，可以在 ECA 內部地端運行的 AI/LLM 由於其相對較小的規模，並不具備必要的複雜性。因此，截至 2024 年 2 月，唯一可行的選擇是購買 ChatGPT 的企業授權。主要供應商（如 OpenAI）目前正在採取必要的步驟，以適應歐盟市場需求（例如：在愛爾蘭建立子公司，並透過已與歐盟簽訂框架合同的公司提供產品，包括 Microsoft）。OpenAI/ChatGPT 可以透過其產品的企業授權獲取，或者作為 Microsoft Office 365 中的一個可選組件（稱為 Copilot）。

##### 4.2 發展並部署一個工具，透過基於問答的互動方式，從大型文檔集中檢索信息

審計團隊經常需要在他們收集的大量特定領域的文件中搜索信息，而這些文件有些不是公開資訊，因此為了確保數據的機密性且符合「歐盟數據保護條例（EUDPR）」，不建議使用像 ChatGPT 這樣的商業 AI。允許審計人員上傳所選文件並根據文件內容提問的工具稱為“檢索增強生成”（Retrieval Augmented Generation，簡稱 RAG）。這類工具由地端運行的 AI 組成，能夠在地端存儲的文件（如 PDF、Word 等）中搜索相關信息。使用者以聊天機器人的形式與 AI 交互，探索文件的內容。ECA 開發“PrivateGPT”工

具，初步測試顯示，在處理包含事實性和簡潔信息的文件時效果良好，但在需要高層次抽象推理的情況下（例如作出專業判斷或處理隱性或主觀信息時），AI 的效果往往不那麼理想。

#### 4.3 擴展 DORA 工具，為其添加 AI 後端，以進行摘要和其他自然語言處理任務

ECA 已開發一個名為 DORA 的文檔閱讀助手，審計人員可將審計文檔集上傳到 DORA，搜索主題、術語等，並建立僅包含與審計目標相關的頁面的“閱讀列表”。惟 DORA 尚未使用 AI，ECA 將透過添加基於 AI 的自動摘要功能以及基於上下文的搜索來增強 DORA。

#### 4.4 在 ECA 現有程式上新增語義搜索引擎

設計“智能搜索（Smart Search）”工具的工作原型，以針對 ECA 產出的特別報告進行全文語義搜尋。主要目的是探索語義搜索的潛力，以改善目前使用的基於關鍵字的方法。

#### 4.5 使用來自 AWARE 和其他受控質量來源的審計方法論內部知識來訓練聊天機器人

可以使用與行動 4.2 中所描述的類似工具來實施審計助手聊天機器人。與其他行動中描述的更通用工具的區別在於，ECA 將準備引入 AI/大型語言模型（LLM）。

#### 4.6 評估歐盟委員會的 DORIS 工具以進行調查分析

DORIS 是一個最初由歐盟委員會開發的工具，用於分析公共諮詢，使用者可以對回答進行分組、執行情感分析、從文本回覆中提取最重要的關鍵詞等。DORIS 與歐盟調查系統（EU Survey）整合良好，對於處理大量多語言的審計調查回覆非常有用。ECA 與歐盟委員會接洽，以評估 ECA 是否可以使用該工具。

#### 4.7 在機器人流程自動化中添加 AI 組件（為重複性任務的自動化增添智能）

ECA 已經設計機器人流程自動化（RPA），並開發數個執行自動化大規模但簡單和重複的任務的機器人。這些機器人可以潛在地擴展 AI 功能，以進一步自動化執行審計作業。具體來說，AI 有潛力提高那些至今只能透過人工完成任務的效率，可以免除因人工輸入文件所帶來的錯誤。

在 RPA 機器人中使用 AI 有兩個選擇：A. 評估供應商提出的 AI 擴充功能；B. 在 ECA 購買商業 AI 工具的許可後，或者當 ECA 能夠在內部運行強大的大型語言模型（LLM）時，可以將 AI 功能納入機器人中，為 LLM 添加查詢步驟並產出結果。例如，在下載了



非結構化發票後，機器人可以向 AI 查詢「提取到期日」之類的特定信息。

## 5. 建立新工具或強化工具的項目提案組合

為實現發展及運用 AI 的目標，需要來自多個領域的專家協同合作。ECA 已經有幾位分散在不同部門的專業人員從事 AI 工作。ECA 創造環境，使他們得以最大限度發揮所長，而 ECA 能充分利用這些知識。

## 6. 促進審計機關間及與國際合作

審計機關間合作有助於優化資源的附加價值，尤其是在如盧森堡等生活成本高昂的領域，這些資源既稀缺又昂貴。傳統上，參與審計機關間工作小組的主要目標是交換想法和資源，並汲取同行的審計經驗。審計機關與國際機構合作帶來的預期益處，詳表 10。

表 10 與國際機構合作帶來的預期益處

交流活動	預期益處
促進思想交流和定期同行檢視	機構可以避免閉門造車。定期的同行檢視能確保 AI 策略及其執行過程能受益於多元的觀點和經驗。
識別方法論和技術合作領域	可以整合專業知識、資源和努力，更有效地應對複雜的挑戰。這種方法不僅加速進展，還能確保解決方案穩健、經過測試，並適應不同的情境。
善用對方既有成就	包括策略、路線圖、項目文件、業務和技術分析文件、使用者反饋、代碼以及架構設計。透過共享這些資源，可以大幅減少重複工作和成本，確保投資用於創新而非重新發明輪子。
分享訓練課程	確保寶貴的見解和經驗不局限於單一機關，而是對更廣泛的受眾開放，促進一個具備良好能力的社群，共同為 AI 領域做出貢獻。
資源共享	創新實驗室之間的協同作用對 AI 項目的有效開發和實施至關重要。透過共享資源，包括人員、知識和基礎設施，機構可以優化投資並降低招聘成本。
倡導更廣泛地重用歐盟委員會開發的工具	可以確保使用歐盟資金開發的 AI 產品和解決方案具備最大影響力、可及性和持久性。最有效的方法是準備詳細的業務需求，並仔細定義使用 AI 所帶來的審計成果。

## (五) 自行研發 AI 或購買現貨—雲端商業產品 (Cloud-based commercial

## products) 與地端工具 (on-premises tool)<sup>21</sup>的比較?

機關決定如何在組織內部署基於 AI 的工具時，有兩種可能的策略。第一種是購買一個或多個商業產品的許可證；第二種是利用內部 IT 資源來安裝和維護 AI 工具的所有必要組件（即語言模型及相關使用者工具）。表 11 為雲端與地端比較。

表 11 雲端與地端比較

面向	雲端	地端
契約性	通常涉及持續的訂閱模式。合約通常是標準化的，且談判空間有限。	涉及前期投資。合約更具客製化。
資料保護	數據存儲在機關之外，通常位於多個位置。這引發了對數據所有權、對個人數據的控制以及遵歐盟數據保護條例 (EUDPR) 的擔憂	提供對數據存儲的更多控制，與 EUDPR 及其他數據保護法規緊密對齊，這對於敏感政府數據至關重要。
合法性	受托國法律管轄的影響，這可能會造成複雜性，特別是在跨境數據傳輸方面。	法律合規性則更為直接，由歐盟法律規範。
資料安全	需要不斷監控以確保遵守不斷變化的法規，例如 EUDPR。	相對而言，法規合規性更具可預測性，但需要內部專業知識來維持標準。
技術性	大多數與部署和維護相關的技術方面都是外包的，這使得成本更可預測，但限制了組織對產品發展的控制。	地端 AI 需要新的 IT 架構組件；內部 IT 專業人員需要學習新技術並獲得新技能。組織對產品的發展保持完全控制。
AI 未來安全法規	遵守任在於供應商。	機關可能需要負責評估地端方案的安全性。
績效及可靠度	提供可擴展的性能和高可靠性，但直接控制較少，透過共享基礎設施和資源提供成本效率和管理便利性。	性能和可靠性取決於前期投資規模。擴展或縮減 IT 基礎設施則更為困難。

<sup>21</sup> 地端工具 (on-premises tool): 與雲端解決方案相對，這些工具安裝並運行在企業 IT 伺服器或個人筆記本電腦上。

## 二、荷蘭審計院

NCA 目前尚未有 AI 發展整體策略，考量風險及資安，所以尚未運用 ChatGpt 於審計業務，惟刻正探索運用 AI 於審計工作的可行性，規劃於 2025 年啟動生成式 AI（例如：ChatGpt）的先導運用，於網路安全、大數據處理、數據挖掘、撰寫摘要、將之前紙本資料轉換為數位化後進行全面搜尋（holistic search）等，未來將視成效再據以決定整體策略內容。荷蘭持續與國際間保持互動交流，以掌握 AI 運用於審計業務之可行性。

## 第三節 公平性審計（Leave No One Behind）

### 一、AI 的發展及運用對特定族群不利影響

AI 的發展及運用對特定族群可能產生不利影響，已引起國際上學界、各國政府及社會廣泛關注。這些影響通常與數據偏見、技術不平等及倫理問題等因素密切相關，包括：

**（一）數據偏見：**AI 的訓練依賴於大量的數據，如果這些數據本身存在偏見，則系統的決策也可能會反映這些偏見。例如：

1. **種族和性別歧視：**在人員招聘過程中使用的 AI，若基於歷史數據進行訓練，因為過去的數據反映了不平等的招聘慣例可能會將某些種族或性別的應徵人員排除在外。
2. **社會經濟地位：**在信貸評估中，若數據未充分考慮某些經濟弱勢群體的情況，則可能導致這些群體無法獲得貸款。

**（二）技術不平等：**AI 技術的發展可能加劇數位鴻溝，導致某些族群面臨技術不平等的問題：

1. **教育與技能差距：**教育資源不足的特定族群或地區可能無法獲得必要的數位技能訓練，無法有效參與 AI 驅動的經濟活動，進而影響其就業機會和經濟狀況。
2. **可接觸性：**偏遠地區或經濟弱勢群體可能無法獲得良好的網路連接和設備，導致他們無法充分利用 AI 技術帶來的便利。

**（三）隱私與安全問題：**AI 的運用可能導致個人隱私的侵害，特別是在監視技術和數據收集方面：

1. **監視和監控**：某些族群或地區（如少數族裔社區）可能成為過度監視的對象，導致社會的不安與不信任。

2. **數據濫用**：如果個人數據被不當使用或外洩，則可能對特定族群的權益造成損害。

(四) **決策不透明**：缺乏解釋性，許多 AI 的運作方式是「黑箱」，即其決策過程不透明，受影響的群體無法瞭解為何會遭遇不利的結果，無法提出異議或尋求補救。

(五) **社會心理影響**：

AI 技術的使用可能對特定族群在心理層面，產生下列影響：

1. **信任問題**：在使用 AI 的機構（如政府、企業等）未能充分解釋或改進其技術的情況下，促使特定族群更加不信任這些機構。

2. **自我實現預言**：如果某些族群被標籤為高風險，AI 的決策可能強化這一標籤，造成自我實現的預言，進一步加深社會不平等。

## 二、國外對於 AI 的發展及運用對特定族群不利影響之研究

歐洲理事會（Council of Europe）報告認為，成員國於發展 AI 同時，必須防止和減輕歧視風險，特別關注那些權利受到不成比例影響的群體，這些群體包括女性、兒童、老年人、經濟弱勢群體、LGBTI<sup>22</sup>社群成員、殘障人士以及種族、民族或宗教群體。成員國必須避免使用可能導致歧視或產生歧視性結果的 AI，並保護個人免受第三方使用此類 AI 的影響。在 AI 生命週期的各個階段，來自這些群體的有效代表積極參與並進行有意義的諮詢，是防止和減輕對人權產生不利影響的重要組成部分。此外，還需要特別關注 AI 開發中所使用的訓練數據的透明性。人權影響評估（HRIA）和其他形式的人權盡職調查應定期重複進行，並提供適當且易於使用的問責和補救渠道。成員國在執法背景下使用 AI 時應採用最高水準的審查。此類系統在部署之前需要進行獨立審核，以檢查是否存在任何可能導致特定群體事實上受到定性分析的歧視性影響。如果檢測到此類影響，則該系統不得使用。<sup>23</sup>

---

<sup>22</sup> LGBTI (Lesbian, Gay, Bisexual, Trans, Intersex)，係指女同性戀、男同性戀、雙性戀、跨性別及雙性人。

<sup>23</sup> 資料來源：歐洲理事會（2019），Unboxing Artificial Intelligence: 10 steps to protect Human Rights。

歐盟委員會 (European Commission) AI 高階專家小組於 2019 年 4 月提出「可信任 AI 道德準則 (Ethics Guidelines for Trustworthy Artificial Intelligence)」，AI 應滿足的 7 項關鍵要求<sup>24</sup>，才能被認為是值得信賴的，其中之一係多元化、非歧視和公平，必須避免不公平的偏見，因為它可能產生多種負面影響，從弱勢群體的邊緣化到偏見和歧視的加劇，為促進多樣性，AI 應該向所有人開放，並讓相關利益相關者參與整個生命週期，透過培訓和教育，使所有人利害關係人瞭解可信賴的 AI 並接受過相關培訓。此外，應建立機制來確保 AI 及其結果的責任和問責，可審計性可評估演算法、數據和設計流程。該指引提到應特別關注涉及較弱勢族群的情況，例如兒童、身心障礙者、殘疾人和其他歷史上處於不利地位或面臨被排斥風險的人以及情況，其特徵是權力或資訊不對稱，例如雇主和工人之間，或企業與消費者之間，並且以清晰、主動的方式向利害關係人傳達有關 AI 的訊息能力和局限性，實現現實的期望設定，以及關於要求得到落實。對他們正在處理 AI 的事實保持透明。

學者研究認為，AI 將影響廣泛的工作，但可能不會平等地為所有就業部門帶來機會，尤其對低薪專業人士。而 AI 在勞動力中的使用有可能加劇現有的種族和性別差異，除非當前 AI 產品的開發能夠解釋並防止 AI 在未來工作場所產生的有害影響。目前大約三分之二的工作都受到一定程度的 AI 自動化的影響，四分之一的當前工作任務可以被 AI 取代，AI 的部署預計將提高生產力、產生效率並創造更多高薪就業；然而，並不能保證因 AI 而失業的低薪工人能夠獲得新創造的機會。報告顯示，雇主對 AI 的使用將導致對 STEM (Science, Technology, Engineering, Mathematics, 科學、科技、工程、數學) 學生的需求增加。女性、黑人等在 STEM 領域職業的代表性不足，除非 STEM 領域雇主、政府和學術機構努力解決這一差異，否則 AI 將進一步加劇這一趨勢。AI 轉型的挑戰和機會之一是透過技能提升和培訓計畫來保護弱勢群體，使技能提升和培訓得確保就業障礙不會更進一步惡化。雇主和政府需要共同努力，讓勞動階層因應 AI 轉型而做好準備，並幫助勞動階層提升爭取新的高薪工作的競爭力。雖然 AI 仍在發展，但政府可利用立法技術、公私夥伴關係以及民間部門的成功案例，防止 AI 的不公平運用。學者認為除了應遵守的特定 AI 原則之外，若要將 AI 應用於弱勢群體，則提出建議：1. 在設計 AI 和進行數據收集時，應有明確的目的。收集數據目的是什麼？即使技術環境發生變化，數據可能以不同的方式被運用，最初的目的是否仍能滿足？數據收集、分析以及使用目的不應隨時間改變，因為一旦政府改變使用目的，目標群體之前所同意使用

---

<sup>24</sup> 7 項關鍵要求分別係：(1) 人類機構與監督、(2) 技術穩健性和安全性、(3) 隱私和資料治理、(4) 透明、(5) 多元化、非歧視和公平、(6) 社會與環境福祉、(7) 問責制。

其資料將不再適用，因為他們並未同意政府新的使用目的；2. 設定嚴格的記錄保存時間表，以尊重個人對其數據收集同意可能隨時間而改變的情況；3. 數據收集的利益應大於風險，數據收集不應對個人帶來更大的風險，亦即數據收集的利益應遠遠超過敏感信息被盜取或被惡意使用的潛在風險。<sup>25</sup>

歐洲輿論報導，AI 可能無法正確識別缺乏四肢、臉部差異、不對稱、言語障礙、溝通方式或手勢不同的人，或使用輔助設備的人，無論從技術或政策角度來看，以殘疾人為中心的 AI 研發仍然是一項複雜的任務，而 AI 演算法扮演很重要的角色。減輕特定群體的演算法風險是一個相當複雜的過程，包括建立針對這些群體的風險類別和影響評估，考慮演算法背後的社會和歷史因素，確保數據的存取以及多利益相關者的監督和參與。<sup>26</sup>

### 三、執行公平性審計之必要性及關鍵要素

公平性審計 (Fairness Audit) 運用於 AI，係指對 AI 或自動化系統進行審查，以確保這些系統在設計和運行過程中不會引發偏見或歧視，特別是對特定群體造成不公平的結果。公平性審計的最終目標是確保 AI 的應用對所有群體都公平，並能夠識別和減少任何潛在的偏見或歧視，以促進社會公平與包容。這一審計過程旨在評估 AI 是否公平，並確保其決策透明、負責任且符合道德標準。其關鍵要素包含：(一) **數據偏見檢測**：AI 通常依賴大量數據進行訓練，如果這些數據本身存在偏見，系統就可能產生偏向特定群體的結果。審計需要檢查數據源是否包含性別、種族、年齡或社會經濟地位等方面的偏見；(二) **演算法透明性**：公平性審計應檢查 AI 所使用的演算法，確保其運作方式透明且可解釋。這樣可以讓系統的決策過程更具透明性，避免算法“黑箱”問題，從而能夠及時發現潛在的不公平決策；(三) **影響評估**：審計需要評估 AI 對不同群體的實際影響。這包括分析系統的結果是否對某些弱勢群體產生了負面影響，並檢測系統是否無意中強化了社會中的既有不平等；(四) **監管與合規**：應審核 AI 是否符合相關的法律和道德規範，特別是在涉及人權或數據隱私的領域。例如，一些國家已經制定了專門的法律來防止 AI 中的偏見或歧視性決策；(五) **持續監測與改進**：公平性審計不應是一次性的，而應該持續進行。隨著數據和技術的變化，AI 可能需要不斷調整和改進，以確保

---

<sup>25</sup> William Lacy Clay Jr., Yvette Puckett Cravis, Amaris Trozzo, The Impact of Artificial Intelligence on Vulnerable Populations in the Workforce.

<sup>26</sup> EURONEWS (2023), Can emerging AI strategies protect people with disabilities and other vulnerable groups?

長期保持公平性。

AI 技術的發展和運用對特定族群的影響是多方面的，涉及社會、經濟和倫理等多個層面。為減少這些不利影響，審計機關允宜進行公平性審計，促使政府研擬 AI 數據透明性及技術包容性等措施，確保 AI 技術的發展能夠真正服務於全社會，避免造成 AI 帶來不平等現象。

經訪談荷蘭萊登大學 Sarah Giest 教授表示，政府發展 AI 同時應兼顧弱勢團體，可以考慮就學齡兒童從小施予 AI 教育，從基礎教育著手，使大家有平等機會接觸，減少未來競爭力的差距；對於目前低技術的勞工階層，則可透過在職訓練，輔導成為 AI 基本操作員，不必塑造每個人都成為 AI 的編碼員；此外，審計機關在執行公平性審計時，允宜注意政府 AI 系統使用數據資料時，該資料內涵之偏差造成決策失誤，以及由於 AI 是依據投入數據資料，經過演算法產生通常是 0 與 1 兩個絕對的決策，但是福利發放等行政行為，有時不是如此絕對，尚須考量個案特殊因素及其配套措施，所以政府發展 AI 時，是否將人性因素（Human-in-the-hoop）納入考量。

## 第五章 結論及建議意見

### 第一節 結論

AI 發展近年來方興未艾，成為世界顯學之一。本次赴荷蘭研習，經研讀相關文獻，並與荷蘭審計院審計人員進行訪談，瞭解該院審計實務，參加萊登大學公共行政管理系研討會暨訪談該校教授 Sarah Giest 博士，探究發展 AI 同時應如何兼顧弱勢團體及公平性之審計作法，歸納歐盟、荷蘭 AI 及其課責制度的發展重點及趨勢，結論個人心得如下：

#### 一、Risk management：以風險評估管理 AI

將 AI 帶來風險及影響評估區別等級，透過系統性地識別和分析潛在的風險點，確保合規性與法規遵循，減少偏見與歧視風險，提升系統的可靠性和穩定性，促進持續監控與改善。與荷蘭審計院訪談發現，於未能有效確保掌控風險前，寧可不運用 AI。

#### 二、Privacy：重視資料隱私

AI 系統通常需要大量數據來訓練和運行，而這些數據往往包含個人資訊。如果缺乏隱私保護措施，使用者的敏感數據可能會被洩露或濫用，進而引發安全和倫理問題。即

便歐盟 AI 法案通過，但是未能免除應該遵守一般資料保護規則（General Data Protection Regulation, GDPR）的責任，要求企業和政府機構確保個人數據在收集、儲存、處理和分享的過程中都得到適當的保護。歐洲人非常重視個人資料隱私之保護，不會接受因為發展 AI，而犧牲個人隱私，所以歐盟及荷蘭在發展 AI 同時，都再再強調保護隱私權。

### 三、Algorithms：強調演算法之重要性及透明性

演算法是 AI 系統的核心，決定 AI 如何處理數據、做出決策以及與使用者互動。提高演算法的透明性有助於提升 AI 的可靠性和公信力，尤其是在公共和商業領域廣泛應用 AI 技術的當下。荷蘭政府雖然推行政府機關部門使用 AI 演算法登記冊制度已數年，力求 AI 演算法之透明性，惟經荷蘭審計院查核發現，目前曾使用、已使用 AI 之演算法僅 5% 顯示於該登記冊，顯示離預期仍有大段距離，荷蘭審計院亦促請荷蘭政府加強演算法透明性。

### 四、Human- in- the loop<sup>27</sup>：以人為本之中心思想

在發展 AI 的過程中，以人為本的理念具有關鍵重要性。AI 技術的終極目標應是改善人類生活、增進福祉、促進社會進步，因此，人應始終被置於 AI 系統設計和應用的核心位置，確保 AI 產品的可用性、易用性和可理解性，以使人們能夠信任和順利使用這些技術。其次，以人為本能更好地保護個人權利，當 AI 涉及到個人隱私和決策時，重視人類需求和倫理原則有助於在系統設計時減少潛在的偏見，並減少對人類權利的侵犯，例如，考慮到透明性、數據隱私、以及公平性原則，可以防止 AI 因偏見或誤用而造成不公平的影響。此外，以人為本的 AI 還能促進社會可持續發展，當 AI 技術服務於整體社會利益，例如醫療、教育、環境保護等公共領域，其價值可以被更廣泛地體現出來，從而造福更多人群。這種以人為中心的理念能激發更多人接受和使用 AI 技術，進一步推動其發展與創新，同時兼顧弱勢團體。蒐集歐盟資料及訪談荷蘭審計院、萊登大學 Sarah Giest 博士，均一再強調 AI 應以人為本的觀念。

---

<sup>27</sup> 同註 4。



## 第二節 建議意見

### 一、行政機關部分

#### (一) 及早制定我國 AI 法案，健全 AI 發展及運用法制架構

在 AI 技術日益普及的現代社會，制定符合臺灣國情的 AI 法案顯得尤為重要。AI 技術的發展和運用需要有明確的法律框架來規範，這不僅可以保障技術的負責任應用，同時也能提升國際間的競爭力。根據「全球負責任 AI 指數 (Global Index Responsible AI, GIRAI)」2024 年的調查結果，臺灣在 6 項指標分數<sup>28</sup>，以 AI 監理制度方面的評分最低，這顯示出臺灣在法制建構上仍有很大的改進空間，隨著歐盟 AI 法案的通過，臺灣企業必須遵守歐盟「一般資料保護規則 (General Data Protection Regulation, GDPR)」，以確保個人數據在收集、儲存、處理和分享的過程中都得到適當的保護。我國國家科學及技術委員會雖已於 2024 年 7 月 15 日預告制定「人工智慧基本法」草案，惟仍尚待完成立法程序。為研究、開發與整合人工智慧之技術資源，藉以發揮我國之智慧創新能力，同時兼顧國人生活之社會與自然環境均衡發展，並促成人工智慧發展倫理規範之制定，建議主管機關參考歐盟 AI 法案，將 AI 引發之潛在風險區分不同級別，並採用基於風險的方法監管 AI，完善健全 AI 發展及運用法制架構，在制定 AI 法案時，可參考歐盟的經驗，強調數據隱私保護、演算法透明性和以人為本的設計原則，不僅有助於建立公眾的信任，還能促進技術的可持續發展。

#### (二) 加強投資 AI 及完備基礎建設並鼓勵民間積極參與，促進公私部門合作

加強投資 AI 及完備基礎建設並鼓勵民間積極參與，順應世界潮流，是推動 AI 發展的關鍵。歐盟已啟動總預算超過 79 億歐元之資助計畫，塑造歐洲社會和經濟的數位轉型，致力於將數位技術帶給企業、公民和公共管理部門，提供策略性資金來應對這些挑戰，支援超級運算、AI、網路安全、先進數位技能等關鍵能力領域的項目，及 AI 創新計畫，支持新創企業和中小企業開發符合歐盟價值觀和規則的值得信賴的 AI，致力於開發歐洲工業生態系統以及公共部門的新穎用例和新興應用程式。此外，荷蘭於 2019 年 10 月即已頒布 AI 策略行動計畫，希望善用世界級網絡、數據中心及主機提供商等良好

---

<sup>28</sup> 6 項指標分別係：政府整體發展框架 (Government Frameworks)、政府行動 (Government Actions)、非政府組織行動 (Non-state Actions)、人權與 AI (Human Rights and AI)、建構負責任 AI 的能力 (Responsible AI Capability) 及負責任 AI 的治理制度 (Responsible AI Governance)。

基礎條件，發展及運用 AI。臺灣期望能在 AI 技術的競爭中立於不敗之地，並因應全球化的趨勢，發揮其在國際舞台上的影響力，建議主管機關投入技術創新和國際合作增加對 AI 技術的投資，建立現代化的基礎設施，包括高速互聯網、雲計算平台和大數據分析中心等 AI 應用的基石，並鼓勵私人企業積極參與 AI 技術的研發和應用，提供資金和政策支持，創造一個有利於創新的環境。透過公私合作，能夠加速技術進步，並確保技術成果真正應用於改善社會福利。

### （三）研擬歐盟 AI 法案生效後，對我國企業衝擊之因應措施

隨著歐盟 AI 法案生效，臺灣企業面臨著新的挑戰與機遇。該法案對 AI 技術的使用設有嚴格的規範，包括數據隱私保護、演算法透明性和以人為本的設計原則等，這些要求對於臺灣企業的國際接軌和市場競爭力提出了更高的要求。首先，臺灣企業需加強對數據隱私的重視，遵守歐盟「一般資料保護規則 (GDPR)」，確保個人數據在收集、儲存、處理和分享過程中的安全性。這不僅能提升企業的國際信譽，還能避免因違規而面臨的高額罰款和法律責任。其次，企業應提升演算法的透明度，確保 AI 系統的決策過程公開、可追溯。這不僅有助於建立公眾信任，還能提高系統的公平性和可靠性。企業需要投入資源研究和開發透明度更高的 AI 技術，並建立相關的監管和審計機制。再者，以人為本的設計原則要求企業在開發和應用 AI 技術時，考慮對用戶的影響，避免技術濫用和偏見。企業應積極參與國際 AI 倫理標準的制定，確保技術應用符合人權和道德規範。在資安方面，政府應增加對 AI 技術的投資，完善基礎設施，並鼓勵私人企業積極參與 AI 技術的研發和應用。透過公私合作，加速技術進步，創造有利於創新的環境。此外，需加強 AI 領域的教育和培訓，培育具有全球視野和創新精神的 AI 人才，確保臺灣在國際 AI 競爭中立於不敗之地。臺灣企業需從法律合規、技術創新、透明性提升和人力資源培育等方面全面應對，確保在全球 AI 發展潮流中保持競爭力。建議主管機關研擬歐盟 AI 法案生效後有助我國企業適應衝擊之因應措施，帶領我國企業及早適應歐盟 AI 法案，俾利開拓歐盟市場。

### （四）秉持不遺漏任何人 (Leave no one behind) 的精神，深根教育國民及企業，因應 AI 時代來臨

AI 技術的發展是一把雙刃劍，它既能帶來無限的創新與機遇，也可能加劇社會的不平等。政府應該秉持「不遺漏任何人」的理念，從教育、訓練、審計、法規等多方面入手，確保 AI 技術的發展不僅有助於經濟增長，也能促進社會公平，並保障弱勢群體的基本權益。在這樣的基礎上，AI 將能夠真正成為推動社會進步的力量，而非加劇社會分

裂的工具。正如荷蘭萊登大學 Sarah Giest 教授於訪談內容所提及，教育是縮小社會差距的關鍵，特別是在 AI 這樣的高科技領域。政府應該從基礎教育層級開始，將 AI 教育納入學校課程，為學齡兒童提供平等接觸 AI 技術的機會，無論他們來自哪一個社會背景或經濟層級。AI 教育不僅僅是學習編程或數據分析，更重要的是要讓學生理解 AI 的基本概念、倫理和社會影響，這將有助於培養一代具備批判性思維的創新人才，為未來的競爭力奠定基礎。此外，對於目前處於低技術層級的勞工群體，政府可透過在職訓練計劃，提供專業技能提升課程，幫助這些群體掌握基本的 AI 操作技能。這些訓練課程不需要每個人都成為 AI 的編程專家，而是應該讓他們能夠掌握 AI 技術的基本應用，如數據輸入、操作簡單的 AI 工具等。這樣的培訓不僅有助於提升他們的就業競爭力，也能在職場中減少 AI 技術帶來的排斥效應。**建議行政機關在發展 AI 時，應該制定全面的法規和政策，確保 AI 技術的發展不會帶來對社會公正和公平的負面影響。這些法規不僅需要涵蓋 AI 的技術標準和倫理準則，還應包括如何保護弱勢群體的措施。對於企業開發 AI 技術，政府可以鼓勵企業在設計 AI 系統時，強化透明度和可解釋性，並設立相關機構進行監管和審查，以確保 AI 技術的使用不會對特定群體造成不公平或不當的影響。**

#### **（五）允宜爭取加入 AI 相關國際專業組織，以掌握 AI 國際最新發展**

在全球 AI 技術快速發展的背景下，無法單打獨鬥僅憑靠一己之力就能發展 AI。加入 AI 相關的國際專業組織對於掌握最新的技術趨勢和發展動向至關重要。這不僅有助於提升我國在 AI 領域的技術實力，還能促進國內外技術與經驗的交流與合作。透過參與這些組織，可以瞭解國際間最新的研究成果、技術標準及法規動向，從而制定出更具前瞻性和合規性的發展路徑。同時，加入國際組織也有助於提升我國在全球 AI 社群中的影響力，為國內 AI 企業和科研機構提供更多的國際合作機會。在這個數位轉型的關鍵時刻，積極參與國際 AI 專業組織，確保我們能夠及時掌握和應用最新技術，實現技術創新和產業升級。OECD 成立 OECD AI 工作小組和專家網絡，下分健康專家組、AI、數據和隱私專家組、OECD AI 指數專家組、AI 風險與責任專家組、AI 未來專家組、AI 事件專家組、計算專家組等 7 個組。我國雖現尚非 OECD 會員國，惟仍受邀以其他身分參與 OECD 活動，例如：以「參與方」身分參與 OECD 鋼鐵委員會、競爭委員會、漁業委員會，以「受邀方」身分參與造船工作小組、優良實驗室操作(GLP)工作小組、資源生產力與廢棄物工作小組、資深預算官員委員會等會議。**為使我國發展 AI 能持續與國際接軌，爰建議行政機關積極爭取參與 OECD AI 工作小組和專家網絡等 AI 國際相關專業組織，以掌握 AI 國際最新發展，交流與學習專業知識。**

## 二、審計機關部分

### （一）關注演算法查核技術方法之國際發展趨勢，以發展演算法審計模式

隨著 AI 技術在各國政府和公共服務領域的應用日益廣泛，對 AI 演算法的監管和審計愈來愈重要。AI 技術的運行通常依賴於資料 (Data)、算力 (Computing Power) 和演算法 (Algorithm) 這三大要素，而其中的演算法，作為決策過程中的核心要素，對政府公共政策的執行、社會公平性、數據隱私以及決策透明度等方面產生了深遠影響。AI 演算法在公共政策執行中的廣泛應用，不僅提升了政府運作的效率與精確度，還帶來了對透明度、公平性、數據隱私等敏感問題的挑戰。因此，如何有效地查核和審計政府使用的 AI 演算法，已經成為現代數位治理中的一個關鍵課題。政府運用 AI 進行決策時，公眾有權知曉這些決策背後的演算法邏輯。AI 系統如果運行不透明，可能會引發對政府決策的不信任。演算法審計的目的之一，就是檢查政府使用的 AI 演算法是否能夠以可解釋的方式展示其運行過程，並確保其對決策結果負有責任。這樣的審計能夠促使政府在使用 AI 技術時保持透明，讓公民理解其背後的運作原理。此外，AI 演算法在處理大量數據時，可能會無意中引入偏見，導致對特定群體的不公平對待。例如，某些群體的需求可能在訓練數據中未被充分考慮，從而影響 AI 系統的決策結果。演算法審計能夠通過檢查演算法的設計與運行過程，發現是否存在偏向性，並確保 AI 技術的公平性。這樣可以避免 AI 技術在實施公共政策時對弱勢群體產生歧視性後果。再者，政府使用的 AI 系統通常需要處理大量涉及個人隱私的數據。這些數據的處理、存儲和使用必須符合隱私保護法規和安全標準。演算法審計有助於確保這些數據不會被濫用或泄露，並且其處理方式遵循法律規定。例如，在歐盟的「一般資料保護規則 (GDPR)」規範下，演算法審計對於保障個人數據隱私至關重要。目前，儘管 AI 演算法審計領域仍處於初步發展階段，特別是歐盟等地的最高審計機關尚未形成統一的查核技術和方法，但國際社會對此問題的關注與探索正在逐步加深，隨著歐盟及荷蘭最高審計機關逐漸重視演算法審計，演算法審計的發展有賴技術創新和國際合作的共同推進。**爰建議審計機關允宜關注國際上有關演算法查核技術方法之發展趨勢，並發展演算法審計模式。**

### （二）廣續執行公平性審計，減緩 AI 落差惡性循環，促使社會永續發展

雖然 AI 提供了許多機會，但它也引發了與透明度和公平性相關的擔憂。這些問題包括 AI 對隱私的影響、AI 系統中的偏見和歧視，以及公眾對 AI 演算法的理解不足。隨著 AI 在政府部門的應用日益廣泛，審計機關在執行公平性審計時應該更加重視數據偏差所可能引發的決策失誤。在許多行政決策中，數據是 AI 系統作出決策的基礎，然而，

數據本身可能包含了歷史上的偏見或系統性問題。例如，社會上某些群體在過去的數據中可能未受到足夠的關注或代表，這會使得 AI 系統對這些群體作出不公平的決策。因此，審計機關應該檢查 AI 系統所依賴的數據來源，並確保這些數據能夠充分反映社會各界的多樣性。此外，AI 系統的決策通常是基於數據經過演算法計算後所得到的結果，這些結果往往是 0 或 1 這樣的絕對決策。然而，在行政行為如福利發放、健康照護等領域，決策往往需要考慮更多的個案特殊性與人性因素。**建議審計機關允宜要求政府在設計 AI 系統時，將人性因素納入考量，以確保 AI 不僅僅依賴數據和演算法，而是能夠處理複雜的、具有多變性的現實情境，並做出更符合人類福祉的決策，並廣續執行公平性審計，降低 AI 落差帶來社會惡性循環影響，力求社會永續發展。**

### （三）研酌配合審計人員任用條例修正，加強進用具備 AI 相關專業知識人才

荷蘭審計院受益於其彈性之人才晉用制度，近年為發展數位審計及 AI，招募具備 AI 專業知識背景人才，為該院發展 AI 演算法審計框架及審核該國政府發展 AI 情形，提供其專業知識，發揮功能。我國公務人員晉用制度雖尚未如荷蘭等國彈性，惟審計人員任用條例於 112 年 11 月 29 日經總統修正公布，擴大審計人才延攬之範圍，審計機關得以任用資訊處理職系考試及格者擔任審計人員。鑑於審計機關審計人員目前多以財務、工程稽察背景居多，亟待具備 AI 相關專業知識人才到部，協助規劃 AI 運用於審計業務；此外，我國政府亦推行 AI 行動計畫，尚待我國審計機關持續查核其計畫執行良窳，及世界大國審計機關已逐漸關注演算法審計發展，我國審計機關允宜研議演算法審計制度。**爰建議審計機關允宜研酌配合審計人員任用條例修正，加強進用 AI 相關專業知識人才。**

### （四）持續與國際專業機關（組織）交流，汲取國際最新 AI 發展及運用新知

荷蘭審計院透過參加國際研討會、會議或是視訊會議與其他國家最高審計機關互動，研究有關 AI 審計發展議題。歐盟審計院亦倡議會員國最高審計機關應經常交流與分享審計經驗，並於「AI 發展路徑圖」規劃辦理每年 2 次歐盟網絡活動，認為透過促進思想交流和定期同行檢視、分享訓練課程及資源共享，避免閉門造車，確保 AI 策略及其執行過程能受益於多元的觀點和經驗，優化投資並降低人員招聘成本等。隨著歐盟等地對 AI 技術監管力度的加大，其他國家和地區也將參與到建立統一審計標準的過程中。這將有助於推動全球範圍內的 AI 技術合規性，並提升公民對政府使用 AI 技術的信任，審計機關亦應與時俱進，掌握國際最新 AI 發展及新知。我國審計機關近年透過薦送審計人員參加美國專業機構訓練課程、參加 OECD 審計人員聯盟論壇、國際內部稽核協會及亞洲區內部稽核協會國際研討會、參訪歐美先進國家審計機關或與其進行線上會議，

瞭解世界 AI 發展及運用，惟 AI 發展瞬息萬變，有待不斷學習，方能不與世界潮流脫節。爰建議審計機關持續與國際專業機關（組織）交流，藉由同行交流、掌握私部門最新技術與方法，反思運用於政府審計面向，期能促進政府良善治理。

### （五）廣續加強審計人員 AI 相關訓練，增進審計人員專業培力

最高審計機構的任務是提供對公共服務的獨立和客觀的檢查，並作為政策制定者所需的反饋機制，以識別改進和機會的領域。因此，最高審計機構需要採用雙管齊下的方法，將 AI 整合到其運作中，一方面作為 AI 的審計者，另一方面作為使用者。最高審計機構必須調整其技能和技術，以便能進行演算法審計，這將需要開發新工具和技術，以有效評估財務報表及其他關鍵信息的完整性。最高審計機構還需要探索利用 AI 構建審計工具的方法，以提高審計的質量，提高效率，並幫助審計機關保持相關性。本部審計人員訓練委員會於 2023 至 2024 年共計開辦 AI 治理與倫理、AI 法律問題、數位永續-數位永續雙軸轉型之現況與發展、生成式 AI—LLM 大型語言模型介紹與應用，及政府機關應用 AI 之案例探討等 21 項 AI 領域之理論、實務及案例相關訓練課程，師資遍及產、官、學、研等界，為審計人員介紹 AI 知識基礎及操作技能，提供本部審計人員 AI 轉型所需專業知識。歐盟審計院於 2023 年 10 月發布「2024-2025 年人工智慧初始策略與部署路線圖 (Artificial Intelligence initial strategy and deployment roadmap 2024-2025)」將「為 ECA 審計人員研擬 AI 培訓路徑」列為 6 項行動方案之一，著重於提供 AI 相關法律課程，亦與數據團隊結合，提供 AI 專業及倫理課程，顯見歐盟審計院於發展 AI 過程，極為重視培訓審計人員。荷蘭審計院亦非常重視審計人員在職訓練，除了透過實體及線上課程之外，鼓勵所屬取得文憑，以增進專業知識與技能。歐盟審計院於 2023 年 12 月成立 AI 培訓方案，除了有關 AI 法律風險和版權方面的法律課程之外，由數據團隊專業人士組織的課程，精確滿足審計人員需求，將於 2024 年和 2025 年進一步擴展此一培訓路徑，及增加由內部專家主講 AI 倫理的特定課程。爰建議審計機關廣續加強訓練審計人員專業知識，增進審計人員培力，以因應 AI 發展瞬息萬變。

## 參考文獻

1. European Court Of Auditors (2023), Artificial Intelligence initial strategy and development roadmap 2024-2025.
2. Netherlands Court of Audit (2024), Focus on AI in the Dutch central government.
3. Leif Z Knutsen, Jo E Hannay, Michael A Riegler (2024), Artificial Intelligence in the Public Sector - An Agenda for Responsible Innovation through Learning.
4. Global-Index (2024), The Global Index on Responsible AI.
5. UNESCO & OECD (2024), G7 TOOLKIT FOR ARTIFICIAL INTELLIGENCE IN THE PUBLIC SECTOR.
6. Netherlands Court of Audit (2021), Understanding algorithms.
7. Netherlands Court of Audit (2023), AI & Algorithmic Risks Report Netherlands.
8. European Commission (2019), HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE.
9. Springer (2024), The Very Long Game- 5 Case Studies on the Global State of Defense AI.
10. Netherlands Court of Audit (2024), Highlights AI & Algorithmic Risk Report Netherlands.
11. OECD (2024), EXPLANATORY MEMORANDUM ON THE UPDATED OECD DEFINITION OF AN AI SYSTEM.
12. OECD (2024), THE POTENTIAL IMPACT OF ARTIFICIAL INTELLIGENCE ON EQUITY AND INCLUSION IN EDUCATION.
13. European Court Of Auditors (2024), EU Artificial intelligence ambition.
14. The Netherlands (2024), Strategic Action Plan for Artificial

Intelligence.

15.OECD (2024), THE OECD TRUTH QUEST SURVEY- METHODOLOGY AND FINDINGS.

16.黃心怡、曾冠球、廖洲棚、陳敦源（2021），當人工智慧進入政府：公共行政理論對 AI 運用的反思。



1. Does the NCA have an overall strategic plan for implementing AI? Looking ahead, would you consider developing AI solutions in-house or acquiring them from external providers?

荷蘭審計院（NCA）是否擬具發展 AI 的整體策略規劃？NCA 未來傾向自己發展 AI 還是外購？

2. In which areas of audit operations is the NCA currently utilizing AI? What has been the most challenging issue encountered so far?

NCA 目前在哪些業務有運用 AI？所遭遇的問題是甚麼？

3. Since the launch of the Audit Framework for Algorithms published by NCA in 2021, could you share your experience in applying this framework to audit AI projects developed and used by the Dutch governments? Have you encountered any notable insights or obstacles in the auditing process? Based on your experiences as senior auditors, what areas do you believe require adjustments.

NCA 於 2021 年推出 AI 演算法審核框架，可以分享運用這個框架於審核荷蘭政府 AI 的經驗嗎？過程有什麼啟發或是遭遇甚麼障礙嗎？基於這些經驗，該框架有哪些需要修改之處？

4. When auditors face uncertain while reviewing AI algorithms, perhaps due to a lack of expertise, what resources or mechanisms does the NCA provide to assist them? What is your view on using third-party certification as an alternative solution? What could be its potential advantages and disadvantages?

倘若審計人員缺乏審核 AI 演算法的專業能力？NCA 提供何種機制或資源，予以協助？如何評價第三方認證機制為替代方案及其優點、缺點？

5. In your opinion, which administrative tasks and government services should be a) encouraged to use AI, and b) prohibited from using AI?

您認為政府那些業務適合使用及應禁止使用 AI？

6. In response to the advent of the AI era, how does the NCA plan to recruit and train auditors to enhance their professional skills in relation to AI implementations and AI auditing? Is there an overall training plan in place?

因應 AI 時代來臨，NCA 如何招募人員，及如何訓練審計人員以增強發展 AI 及審核 AI 的專業技能？NCA 有專責訓練單位嗎？擬具整體訓練計畫嗎？

7. Could you share any insights gained from your interactions with international professional bodies (organizations) (if any) or other supreme audit institutions from different countries over the past three years?

可以分享過去 3 年和國際專業組織或國外最高審計機關交流有關 AI 議題的收獲嗎？未來國際交流 AI 議題的重點有哪些？是否與臺灣保持交流審計經驗？

## 荷蘭中央政府會計帳戶法案 2016 (Government Account Act 2016) <sup>1</sup> (摘錄<sup>2</sup>)

### 第 1 章 總則

#### 第 1.1 節 定義

- 中央政府審計服務局：負責在中央政府內執行審計職能的財政部單位。
- 理事會和辦公室：國會、國家理事會、審計院、國家監察人、荷蘭騎士榮譽團的秘書處、國王辦公室、阿魯巴總督辦公室、庫拉索總督辦公室、聖馬丁總督辦公室、選舉委員會以及情報和安全服務審查委員會。

### 第 2 章 中央政府總預算及課責制度

#### 第 7 條 遞交年度報告及最終財務報告

##### 第 2 節 遞交年度報告

- 每年，財政部長必須在預算年度的次年 3 月 31 日之前，將年度報告轉交給審計院進行審計。同時，中央政府審計服務局編製的有關年度報告的審計報告也必須一併轉交。
- 財政部長必須在審查年度的次年 5 月的第三個星期三向國會眾議院提交由審計院審計的年度報告，除非該日期恰逢眾議院休會期間或與公共假日重疊。在這種情況下，政部長必須在與眾議院院長和審計院院長協商後，最遲於同年 6 月 1 日提交年度報告。
- 當需要時，財政部長必須在眾議院討論年度報告的日期之前，儘快向眾議院提交對審計院所提出異議的看法聲明。

---

<sup>1</sup> 資料來源：荷蘭審計院全球資訊網  
(<https://english.rekenkamer.nl/publications/publications/2018/01/01/government-accounts-act-2016>)

<sup>2</sup> 除第 7 章外，其餘章節謹摘錄與審計業務相關之條文。

### 第 2.38 節 遞交中央政府年度財務報告

- 財政部長必須在預算年度的次年 4 月 21 日之前，將中央政府年度財務報告提交給審計院進行審計。同時，中央政府審計服務編製的有關年度報告的報告也必須一併提交。
- 財政部長必須在審查年度的次年 5 月的第三個星期三，向國會提交由審計院審計的中央政府年度財務報告，除非該日期恰逢眾議院休會期間或與公共假日重疊。在這種情況下，財政部長必須在與眾議院院長和審計院院長協商後，最遲於同年 6 月 1 日提交中央政府年度財務報告。

### 第 2.39 節 遞交有關預算最終執行數的法案 (Submission of Bills concerning the final variances)

- 必須在與預算報表相關的年度的次年 5 月的第三個星期三向國會眾議院提交通過預算最終執行數的法案，除非該日期恰逢眾議院休會期間或與公共假日重疊。在這種情況下，必須在與眾議院院長和審計院院長協商後，最遲於同年 6 月 1 日提交法案。

## 第 8 條 免責

### 第 2.40 節 免責

- 在收到審計院出的核准通知之前，不得授予免責。在此之後，還必須通過第 2.39 節提到的通過預算最終執行數的法案。

## 第 3 章 預算管理及財務管理標準

### 第 1 條 一般標準

#### 第 3.2 節 預算管理

- 部長及各理事會和辦公室負責確保預算管理的有序性和可審計性。

#### 第 3.2 節 財務管理

- 部長及各理事會和辦公室負責財務管理的效率、合規性、有序性和

可審計性。

### 第 3.4 節 業務管理

- 部長及各理事會和辦公室負責實體資源的獲取、管理和處置的效率、合規性、有序性和可審計性。

## 第 2 條 財務紀錄及會計資訊標準

### 第 3.5 節 財務紀錄

- 財務記錄必須以可靠和可審計的方式進行結構化和保存。

## 第 4 章 預算管理及財務管理責任

### 第 1 條 部長之政策及業務管理

#### 第 4.1 節 預算管理及業務管理：總體

- 每位部長對其預算所依據的政策負有進行政策有效性和效率的定期審計的責任。
- 每位負責部門的部長對其負責的預算負有對該運營管理的有效性和效率進行定期審計的責任。

### 第 2 條 理事會與辦公室部門

#### 第 4.4 節 理事會及辦公室管理

- 內政與王國事務部長負責管理國會、國家理事會、審計院、國家監察人、荷蘭騎士榮譽團秘書處、阿魯巴總督辦公室、庫拉索總督辦公室、聖馬丁總督辦公室和選舉委員會的預算。

### 第 3 條 私法下的法律行為

#### 第 4.7 節 私法下法律行為的初步議會審查程序

- 若涉及的法律行為的財務利益低於財政部長就第 1 項(a)至(d)項

所指定的金額，則第 1 項所提及的程序不適用於進行私法上法律行為的意圖。

- 相關的部長必須就進行第 1 項(a)所提及的私法上法律行為的意圖與審計院進行諮詢。相關的部長必須給予審計院合理的諮詢時間。在與審計院協商後，相關的部長必須將擬議決策提交內閣。

## 第 6 條 資產負債表管理及財政部長的其他任務§ 6.

### 第 4.20 節 規定

財政部長可以為中央政府制定以下規則：

- 有關審計和運營管理事務的部門諮詢機構的組成、組織和任務及信息提供。
- 中央政府審計服務局的任務、組織和質量控制。

## 第 6 章 中央政府以外公共資金管理的監督

### 第 1 條 中央政府以外公共資金管理的監督

#### 第 6.1 節 部長的監督

- 在不損害其他法定規定或歐盟法規規定的情況下，相關的部長必須監督法人、有限合夥、普通合夥以及從事職業或經營業務的自然人，如果他們直接、間接或有條件地從歐盟預算中獲得了補助、貸款或擔保，只要該歐盟成員國負責監督和審計該補助、貸款或擔保及其管理；

#### 第 6.3 節 部長的監督權力

- 相關部長可以根據會計記錄，對第 6.1 節所提到的法人、有限合夥、普通合夥及自然人的帳目進行審計。
- 相關部長有權要求檢查審計過第 1 項(a)和(b)項文件的會計師的審計程序和檔案，以確定該會計師所進行的審計是否可以在第 6.1 節提及的監督中依賴。會計師不得以依據國會法案對審計檔案中包含的機密數據所施加的保密義務為由拒絕允許檢查審計檔案。

- 在不損害國會法案其他地方規定的條款的情況下，相關部長不得對省、市政府、水利當局、博奈爾、聖尤斯特修斯和薩巴公共機構、職業和貿易的公共機構，以及根據《聯合安排法》成立的公共機構和聯合機構進行審計，除非是參與中央政府的公共機構和聯合機構，前提是它們已獲得第 6.1 節開頭及(a)項所提到的貢獻。

#### 第 6.4 節 部長在外包情況下的監督權力

- 如果會計職能或相關任務被外包給第三方，相關部長有權對該第三方或代表該第三方保留賬目的個人所維護的賬目進行第 6.3 節所提及的審計。

#### 第 6.5 節 會計師的責任

- 根據第 6.3 節允許檢查審計檔案並提供這些檔案中文件副本的會計師，對第三方因此遭受的任何損失不承擔責任，除非可以證明考慮到所有事實和情況，他不應該合理地允許檢查。

### 第 2 條 國庫銀行的監督

#### 第 6.7 節 部長對國庫銀行的監督

- 根據第 5.1 節所指定的法人，必須在相關部長的要求下進行審計，以確定其是否遵守第 5.2 至 5.4、5.7 和 5.8 節所提及的義務。該審計必須由負責審計該法人年度賬目的會計師進行。這些法人必須將審計報告轉交給相關部長。

### 第 3 條 歐盟支出聲明

#### 第 6.9 節 歐盟支出聲明

- 財政部長可在相關部長的同意下，為歐洲委員會編制一份有關荷蘭作為成員國對共同管理的歐洲基金支出的年度聲明。
- 財政部長必須將該聲明轉交給歐洲委員會。
- 財政部長必須將該聲明轉交給審計院進行審計。
- 在該聲明經審計院審計後，財政部長必須將其轉交給國會眾議院。

## 第 7 章 荷蘭審計院

### 第 1 條 組成與組織

#### 第 7.1 節 組成

1. 審計院由三名普通成員組成，這三名成員共同組成審計院的董事會，最多還可有三名特別成員。
2. 院長由內政與王國事務部長推薦，經皇家法令從普通成員中任命。
3. 審計院必須儘快通知國會眾議院其成員的任何空缺。
4. 審計院的董事會必須在收到第 3 項所提及的空缺通知後，儘快擬定至少四名推薦候選人的名單，並將其轉交給國會眾議院。在提出建議時，國會眾議院可以根據其認為合適的方式考慮推薦候選人名單。

#### 第 7.2 節 特別成員

1. 特別成員可以由院長邀請參加某些活動，在這種情況下，他們在該活動中擁有與普通成員相同的權力。特別成員也因此成為審計院董事會的一部分。
2. 如果普通成員缺席或無法參加，其職位將由特別成員代替。

#### 第 7.3 節 秘書長

1. 審計院設有一名秘書長。
2. 秘書長由審計院推薦，經皇家法令任命和解雇。審計院董事會可以對秘書長進行暫時停職。

#### 第 7.4 節 任命要求

1. 審計院的成員僅限於具有荷蘭國籍的人士。
2. 成員和秘書長不得擔任任何固定報酬或津貼的其他公職，或屬於任何依法律規定經選舉產生的公共機構的成員。前提是該不相容情況不會因其他法律規定而產生，在聽取審計院意見後，可



以經皇家法令豁免此條款。

3. 除第 2 項的情況外，成員和秘書長不得擔任其他足以影響可能與他們在審計院的職責的正當執行或他們的公正性和獨立性的維持或對此的信任的職位。
4. 審計院的成員和秘書長在擔任該職務外所擔任的任何職位，必須由院長每年公開報告。

### 第 7.5 節 終止與停職

1. 成員可隨時提出辭職，並且在達到 70 歲時必定終止其成員資格。成員資格的終止自下個月的第一天起生效。
2. 荷蘭最高法院可解雇或停職成員。「司法官法」（第 6A 章（第 46b、46c（1）（b）、（2）和（3）、46d、46i（1）（c）、46k 和 46q 條除外）可適用，前提是：
  - a. 審計院的院長對成員施以書面警告的紀律處分；
  - b. 第 46e 條中的“審計院的院長”一詞取代“審判官，即上訴法院或地方法院的院長、最高法院的院長或最高法院的檢察總長”；
  - c. 審計院的院長被視為上級；
  - d. “內政與王國事務部長”一詞取代“部長”；
  - e. 第 46i（4）條和第 46l（2）條所提到的建議由審計院提出；
  - f. 第 46p（5）條中的“審計院”一詞取代“相關法院或最高法院的檢察總長辦公室”。
3. 除第 2 項提及的解雇理由外，成員行為違反第 7.4（3）條的情況亦可作為解雇理由。
4. 可根據或依據內閣命令制定有關遣散費和疾病及失能條款的規定。

### 第 7.6 節 就任

1. 院長、其他普通成員、特別成員和秘書長在就任前，必須在國會面前宣誓（聲明和確認）以下誓言：
  - “我宣誓（聲明）我沒有以任何名義或任何藉口，直接或間接地向任何人提供或承諾任何禮物，以獲得我的任命。
  - 我宣誓（聲明並確認）我未曾接受，也不會接受任何人的禮物或承諾，無論是直接還是間接，作為引誘我在職務上做或不做任何事情的報酬。
  - 我宣誓（確認）我將忠於國王，始終尊重憲法，誠實、認真且公正地履行我的職責。
  - 願全能的上帝保佑我。（我如此宣誓和確認。）”
2. 在授權下，院長可以在審計院的董事會會議上向其他普通成員、特別成員和秘書長宣誓或進行確認。

#### 第 7.7 節 內部規則

- 審計院必須制定其工作的內部規則。這些內部規則必須在政府公報上公布。

#### 第 7.8 節 院長的任務和權力

1. 院長必須監督審計院的工作，確保本章所列條款的正確適用。
2. 院長必須確保所有寄送至審計院或寄送至他作為院長的文件，都由董事會在會議中進行審議，除非根據內部規則被排除在外。
3. 在院長缺席的情況下，最年長的普通成員必須行使他的任務和權力。

#### 第 7.9 節 決策

1. 審計院的董事會以多數票通過決策。
2. 在投票平局的情況下，院長擁有決勝票。
3. 審計院的董事會在會議中必須至少有多數成員出席，方可作出決策。

### 第 7.10 節 迴避義務

1. 成員和秘書長不得參與涉及他們、配偶、登記伴侶、同居者或與他們有三度內血緣或婚姻關係的任何人的討論或決策。
2. 成員和秘書長不得參與審計或對他們所編制的賬目或財務報表的決策。

### 第 7.11 節 法律地位

1. 中央和地方政府人事法中關於中央政府人員法律地位的規定應相應適用於審計院的官員。審計院的董事會可以任命、暫停和解雇審計院的官員，但如果任命或解雇需由皇家法令進行，則必須由審計院提出相應的建議。
2. 審計院可以授權秘書長負責官員的任命、暫停和解雇。
3. 官員必須在院長面前宣誓或作出聲明和確認。這一要求也可以適用於根據民法聘用合同在審計院工作的員工。

## 第 2 條 年度報告審核及績效審計

### 第 7.12 節 年度報告的審計

1. 審計院每年必須對中央政府的以下事項進行審計：
  - 年度報告中的財務報告信息；
  - 年度報告中非財務報告信息的編制；
  - 中央政府年度財務報告中的財務報告信息。
2. 審計院必須對中央政府的以下事項進行審計：
  - 預算管理、財務管理、物資運營管理以及為此目的保留的中央政府會計記錄；
  - 財政部國庫的中央會計記錄。

### 第 7.13 節 年度報告審計範圍

1. 第 7.12 節 (1) 所提及的審計目的在於確定第 3.8 至 3.10 節所述的標準是否已達成。
2. 第 7.12 節 (2) 所提及的審計目的在於確定第 3.2 至 3.5 節所述的標準是否已達成。

#### **第 7.14 節 年度報告審計報告**

1. 每年，審計院必須在第 7.12 節 (1) 和 (2) 所提及的審計中記錄其發現和結論於報告中。
2. 審計院必須對中央政府年財務報告中包含的中央政府賬目和第 2.35 節 (2) 和 (4) 所述的中央政府試算表發出批准聲明。
3. 如有必要，批准聲明可在最終變更於第 2.36 節所述的變更和必要時於第 7.22 節 (3) 所述的賠償法案通過的條件下發出。
4. 在採納如第 1 款所述的報告之前，審計院必須給予相關部長在合理期間內對其發現和暫定結論進行意見反饋的機會。
5. 具有保密性的資訊和發現不得包含在審計院的報告中，如此類資訊或發現可以保密方式發送給國會以供參考。

#### **第 7.15 節 年度報告審計報告的提交**

1. 審計院必須在預算年度的次年最遲於 6 月 1 日之前向國會提交第 7.14 節 (1) 和 (2) 所述的報告及批准聲明。
2. 如果審計院未能在預算年度的次年 6 月 1 日之前完成第 7.12 節 (1) 和 (2) 所提及的審計，則必須在該日提交一份有關審計進度的臨時報告。在這種情況下，審計院必須在隨後盡快提交最終報告和第 7.14 節 (1) 和 (2) 所述的批准聲明。

#### **第 7.16 節 績效審計**

- 審計院必須檢查中央政府所實施政策的有效性和效率。

#### **第 7.17 節 績效審計報告的提交**

1. 審計院必須通知國會有關其根據第 7.16 節所進行的審計所採

- 納的報告。
2. 如有必要，審計院必須向第 7.34 節 (8) 所述的機構通知第 1 款所述的報告。
  3. 在採納如第 1 款所述的報告之前，審計院必須給予相關部長在合理期間內對其發現和暫定結論進行意見反饋的機會。
  4. 具有保密性的資訊和發現不得包含在審計院的報告中，如此類資訊或發現可以保密方式發送給國會以供參考。

### 第 7.18 節 年度報告和績效審計的權限

1. 審計院可根據其認為執行任務所需的情況，檢查所有中央政府部門的所有財物、記錄、文件和其他信息載體，以其認為適當的方式進行。
2. 審計院有權要求檢查已根據其中一位部長或委員會或辦公室的指示進行審計或審查的會計師的審計計畫和檔案。審計院有權對審計檔案中的文件進行複印。會計師不得因依據法律或根據法律法規所施加的保密義務而拒絕允許審計院檢查其審計檔案。第 6.5 節相應適用。
3. 在執行其任務時，審計院可在不妨礙其進行自身審計的情況下，利用其他人進行的審計結果。
4. 審計院可在中央政府對其有興趣的情況下，行使本節所述的權力。
5. 部長及委員會和辦公室必須根據要求，向審計院提供其認為必要的信息，以便執行其任務。
6. 根據一位部長或委員會或辦公室的指示進行審計的會計師必須根據要求向審計院提供與其審計相關的審計計畫和檔案。

### 第 7.19 節 在外包情況下對年度報告和績效審計的權限

- 如果會計功能或相關任務外包給第三方，則審計院有權對相關第三方或代表該第三方維護帳目的個人進行第 7.12 節和第

7.16 節所述的帳目審計。

#### **第 7.20 節 對於秘密支出和收入的年度報告和績效審計的權限**

1. 審計院的董事會必須對秘密支出和收入進行審計。
2. 在執行審計時，董事會可安排其工作人員提供協助。

#### **第 7.21 節：異議程序之一般規定**

1. 根據第 7.12 節的審計，審計院可對財務管理、物質運營管理或與其相關的報告提出異議。
2. 審計院必須通知相關部長有關異議。
3. 在收到第 2 款所述的通知後的一個月內，部長必須告知審計院可能解決其異議的措施。
4. 在該期間屆滿後，審計院必須作出最終決定並相應通知部長。

#### **第 7.22 節：如果異議被維持的程序**

1. 審計院可以維持其第 7.21 節 (1) 所述的異議。
2. 如果審計院維持其異議，則必須通知相關部長和財政部長。
3. 如果異議涉及規範問題，則必須在收到第 2 款通知後兩個月內向國會提交一項通過賠償法案的法案。
4. 如果在該期間屆滿時尚未提交此類法案，則審計院必須相應通知國會。
5. 如果其異議與第 3 款所述的異議不同，審計院必須在該報告中通知其異議，如第 7.14 節所述。
6. 審計院還可在其有關中央政府年度財務報告的報告中附加相關備註。

### **第 3 條 其他任務和權限**

#### **第 7.23 節 應要求進行審計**

- 審計院可根據國會兩院或任何部長的要求進行審計。

#### 第 7.24 節 對中央政府以外的公共資金的審計

在不妨礙其他法律條文的情況下，審計院可對下列事項進行審計：

1. 法律人、有限合夥、普通合夥和從事專業或業務的自然人，若他們：
  - 直接、間接或附帶條件地從中央政府預算中獲得了贈款、貸款、擔保或具有贈款、貸款或擔保特徵的實物貢獻；
  - 直接、間接或附帶條件地獲得了相關部長簽發的有關研究與發展工作、能源投資、環保投資或租住住房投資的稅收減免聲明或根據部長令發出的指令；
2. 具有法定任務的法律人；
3. 直接或間接為了履行公共任務而獲得資金的人；
4. 直接或間接從中央政府獲得資金的公司或經濟實體。

#### 第 7.25 節 對中央政府以外公共資金的審計範圍

1. 根據第 7.24 節開頭及 (a) 至 (c) 款的規定，對法定人、有限合夥、普通合夥及自然人的審計目的是形成對：
  - 相關部長對第 7.24 節 (a) 至 (c) 款所提及的法定人、有限合夥、普通合夥及自然人所採取政策的看法
  - 相關部長對第 7.24 節 (a) 至 (c) 款所提及的法定人、有限合夥、普通合夥及自然人的監督情況的看法。
2. 在不影響第 1 款的前提下，對於第 7.24 節開頭及 (b) 款所提及的具有法定任務的法定人的審計目的之一是形成對：
  - 具有法定任務的法定人對公共資金的管理的看法；
  - 法定任務的履行情況。
3. 對於荷蘭中央銀行 (De Nederlandsche Bank NV) 的審計不涉

及執行歐洲聯盟運作條約的任務。

### 第 7.26 節 公共機構和聯合機構的審計範圍

根據第 7.24 節開頭 (d) 款所提及的公共機構和聯合機構的審計目的為：

- 形成對中央政府參與第 7.24 節 (d) 款所提及的公共機構和聯合機構的方式的看法。
- 評估相關部長對第 7.24 節 (d) 款所提及的公共機構和聯合機構所採取政策的看法。

### 第 7.27 節 國有企業的審計範圍

根據第 7.24 節開頭 (e) 款所提及的公共和私營有限公司的審計目的為：

- 形成對國家作為股東在這些公司的權利行使的看法；
- 評估相關部長對這些公司的政策的看法。

### 第 7.28 節 對歐盟預算資助的審計

- 在不影響其他法案或歐盟法規中規定的條款的情況下，審計院可以對法律人、有限合夥、普通合夥和從事職業或經營業務的自然人進行審計，如果他們直接、間接或有條件地從歐盟預算中獲得資助、貸款或擔保，只要該歐盟成員國負責監督和審計該資助、貸款或擔保及其管理。

### 第 7.29 節 對歐盟預算資助的審計範圍

- 根據第 7.28 節所提及的法律人、有限合夥、普通合夥和自然人的審計目的為形成對相關部長所進行的監督的看法，以履行根據歐洲聯盟運作條約對於所獲得的資助、貸款或擔保的財務管理及其審計或監督的義務。

### 第 7.30 節 其他審計報告的呈報

1. 審計院必須向國會通報其已採納的報告，該報告係根據第



- 7.23、7.24 和 7.28 節進行的審計。
2. 在必要情況下，審計院還必須向第 7.34 節 (8) 款所提及的機構通報第 1 款所述的報告。
  3. 在採納第 1 款所述的報告之前，審計院必須給予相關部長在合理期限內對其調查結果和暫定結論進行評論的機會。
  4. 具有保密性質的信息和調查結果不得由審計院納入第 1 款所提及的報告中。含有此類信息或調查結果的通訊可以作為保密資料發送給國會以供其參考。

### **第 7.31 節 對歐盟支出的聲明的審計**

- 審計院必須審計第 6.9 節所述的有關荷蘭對歐洲資金的支出的聲明，該資金由共享管理進行分配。

### **第 7.32 節 對歐盟支出聲明的審計報告的呈報**

- 審計院必須向國會和財政部長通報其已採納的報告，該報告係根據第 7.31 節進行的審計。
- 第 7.30 節 (2) 至 (4) 款的條款適用於此。

### **第 7.33 節 活動報告**

- 審計院必須在每年 4 月 1 日之前向國會呈報其前一年活動的報告。

### **第 7.34 節 對中央政府以外公共資金及公共機構和聯合機構的審計權限**

1. 在進行第 7.24 (a) 至 (d) 款所述的審計時，審計院必須盡可能多地利用其他機構進行的審計的調查結果。
2. 通過使用相關部長或第 8 款所提及的機構持有的文件，審計院可以獲知有關第 7.24 (1) (a) 至 (d) 款所述的法律人、有限合夥、普通合夥、自然人、公共機構和聯合機構的信息。
3. 如果審計院認為相關部長或第 8 款所提及的機構持有的信息

- 有必要這樣做，則審計院有權向相關的法律人、有限合夥、普通合夥、自然人、公共機構和聯合機構索取進一步的信息或要求其提供文件。
4. 通過使用檔案，審計院可以對第 7.24 (1) (a) 至 (d) 款所述的法律人、有限合夥、普通合夥、自然人、公共機構和聯合機構進行審計。第 7.18 (1) 和第 7.19 節適用於此。
  5. 審計院有權要求檢查已審計第 6.3 (1) (a) 和 (b) 款所提及文件的會計師的審計計畫和檔案。會計師不得以法律規定的保密義務拒絕讓審計院檢查其審計檔案。第 6.5 節適用於此。
  6. 如果審計院行使第 3 至 5 款所述的權限，則《一般行政法》中的第 5:12、5:13、5:15 和 5:17 (2) 和 (3) 款的條款亦適用於此。
  7. 審計院可在公共利益需要的期間和年份內行使本節所述的權限。
  8. 根據法律授權或根據法律設立的機構有責任告知審計院有關其對具有法定任務的法律人的監督結果，並必須根據審計院的規定提供其審計計畫。如果審計院要求，該機構必須提供其審計計畫。
  9. 審計院必須將其根據本節進行的審計情況告知相關部長。
  10. 本節不適用於省、市、排水機構、博奈爾、聖尤斯特歐斯和薩巴公共機構、職業及行業的公共機構、根據《金融監督法》和《BES 金融市場法》所述的金融機構和電子貨幣機構，以及根據《聯合安排法》建立的公共機構和聯合機構，

### 第 7.35 節 國有企業的審計權限

1. 第 7.34 節 (1) 至 (3) 條適用於國家持有至少 5% 已發行股本的公私有限責任公司，如第 7.24 節 (e) 所述，前提是任何進一步的信息必須通過相關的部長獲得，且僅限於年度賬目及審計年度賬目的會計師相關報告。

2. 第 7.34 節 (1) 至 (7) 和 (9) 條適用於國家持有超過 50% 已發行股本的公私有限責任公司，以及國有和私有有限責任公司持有（直接或間接）超過 50% 已發行股本的情況。

### 第 7.36 節 對歐盟預算貢獻的審計權限

- 第 7.34 節 (1) 至 (9) 條適用於第 7.28 節所述的歐盟預算貢獻。

### 第 7.37 節 王國內部的合作

1. 審計院可以與阿魯巴、庫拉索和聖馬丁的審計院及博奈爾、聖尤斯特歐斯和薩巴公共機構的聯合審計院合作。
2. 機密性的數據、發現和結論不得披露給第 1 款所述的審計院。

### 第 7.38 節 國際活動

1. 審計院可以根據其法定任務進行國際活動。
2. 在進行第 1 款所述的活動時，審計院可以與其他國家類似的研究機構合作。
3. 第 7.37 節 (2) 條同樣適用於第 1 款和第 2 款所述的活動。

### 第 7.39 節 溝通

- 審計院可以向財政部長、相關部長和國會發送其認為符合公共利益的任何溝通。

## 第 4 條 結論

### 第 7.40 節 與審計院的諮詢

1. 相關部長必須與財政部長和審計院諮詢有關：
  - a. 由或根據國會法案制定的有關審計院任務和權限的規則；
  - b. 第 4.7 節 (3) (a) 所提到的法律規定，僅限於涉及國家成立或共同成立私法法律人的行為。
2. 相關部長須與外交部長及財政部長協商有關草擬的歐盟立法，該立法涉及國家審計辦公室的地位、任務或權限。

3. 財政部長必須與審計院就根據本法制定的規則進行諮詢，但不包括有關：
  - a. 第 4.20 節 (1) 開頭部分 (a) 所提到的中央政府預算結構；
  - b. 第 4.20 節 (1) 開頭部分 (b) 所提到的預算程序；
  - c. 第 4.20 節 (1) 開頭部分 (d) 所提到的預算管理；
  - d. 第 4.20 節 (1) 開頭部分 (e) 所提到的有關預算的財務記錄。
4. 相關部長必須允許與審計院的諮詢有合理的時間。

## 訪談 Sarah Giest 教授題綱

- 一、 聯合國一直倡議不遺漏任何人 (Leave No One Beyond)，請問發展 AI 同時，政府應如何保護弱勢團體、應該如何降低不公平？例如：低收入家庭沒有多餘資源讓兒女學習有關 AI 的知識，基層勞工的工作將來可能會被 AI 所取代，都可能產生惡性循環或是雪球效應。

The concept "Leave No One Behind" was introduced by the United Nations and as a part of the 2030 Agenda for SDGs. The United Nations has advocated that concept for years. The trend of AI spreads all over the world. It seems to be unstoppable. With development and utilization of AI, the vulnerable groups don't have many resources to keep up with the trend. For example, the low-income family doesn't have money to learn knowledge about AI or the labors with low-level skills could lose their jobs because their employers use AI. It could be vicious circle or snowball effect. How should governments protect the vulnerable groups? What should governments do in the AI era to reduce inequalities?

- 二、 我國審計部近年來強調公平性審計，測使我國政府施政時應顧及弱勢團體。身為最高審計機關，審計部有責任監督我國政府 AI 政策執行良窳，審核其有效性及效率性。請問審計機關在執行公平性審計時應注意哪些面向（關鍵要素）？

NAO recently emphasizes the fairness audit to urge Taiwan government to take care of the vulnerable groups. In terms of a supreme audit institution, NAO has

responsibilities to monitor how Taiwan government develops and utilizes AI, and also audit the effectiveness and efficiency of government's AI policy. Could you tell me what aspects or key elements audit institutions should take into account when they conduct fairness audit?

# The Development of Artificial Intelligence And Its Accountability In Taiwan

CHUNG-HUANG (Eric) CHIU  
National Audit Office (Taiwan)





**WHY AM I  
HERE?**



**WHAT AM I  
GOING TO DO?**



# Table of Contents

## **PART I**

- **The Introduction of the National Audit Office (NAO), R.O.C. (Taiwan)**

## **PART II**

- **The Development of Artificial Intelligence In Taiwan**
- **The Accountability System of Artificial Intelligence In Taiwan**



# **PART I**

## **The Introduction of the National Audit Office, R.O.C. (Taiwan)**

# The Organizational Structure of R.O.C.(Taiwan) Government



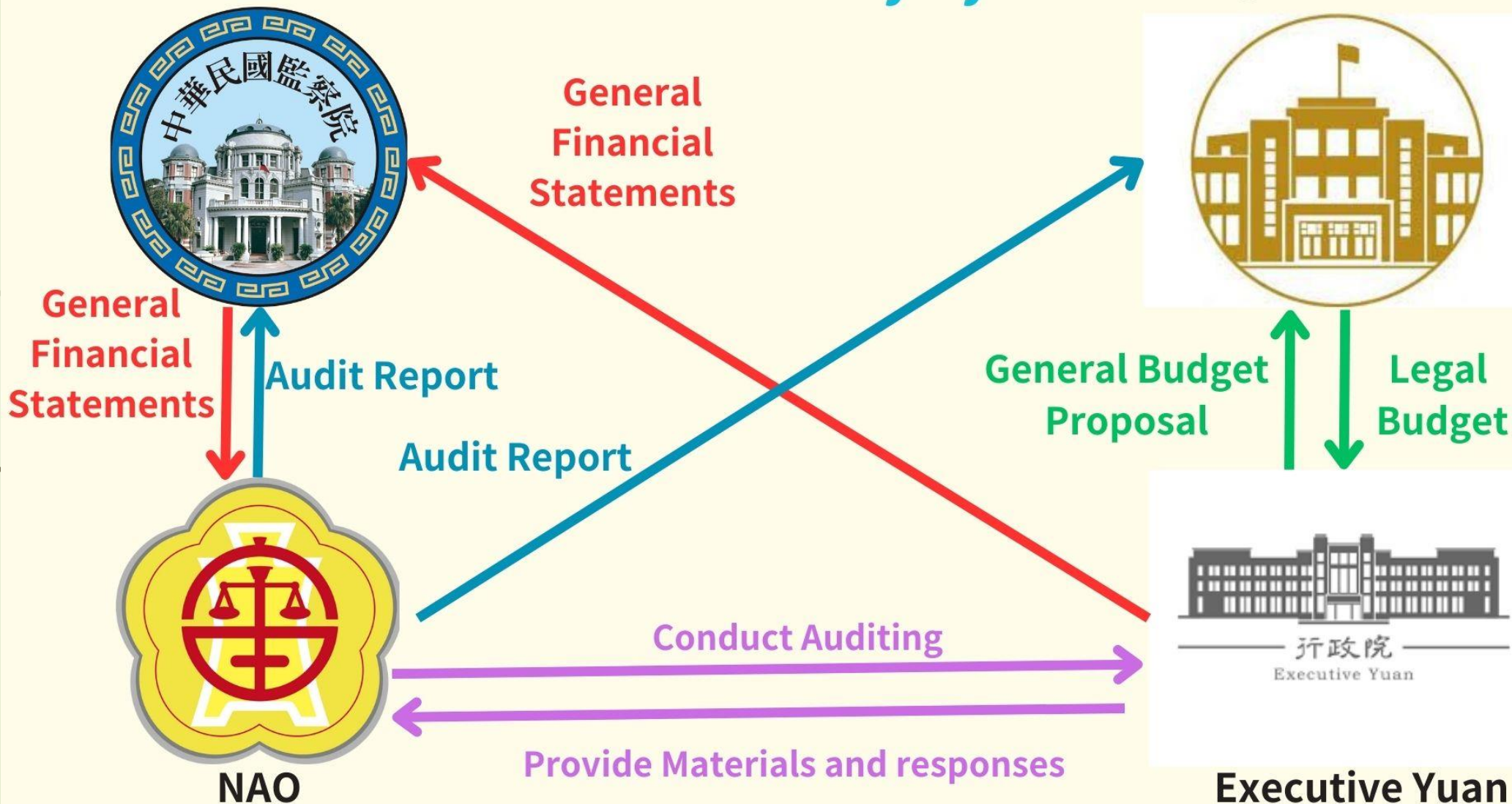
# NAO's Role in the Government Budgeting Procedure



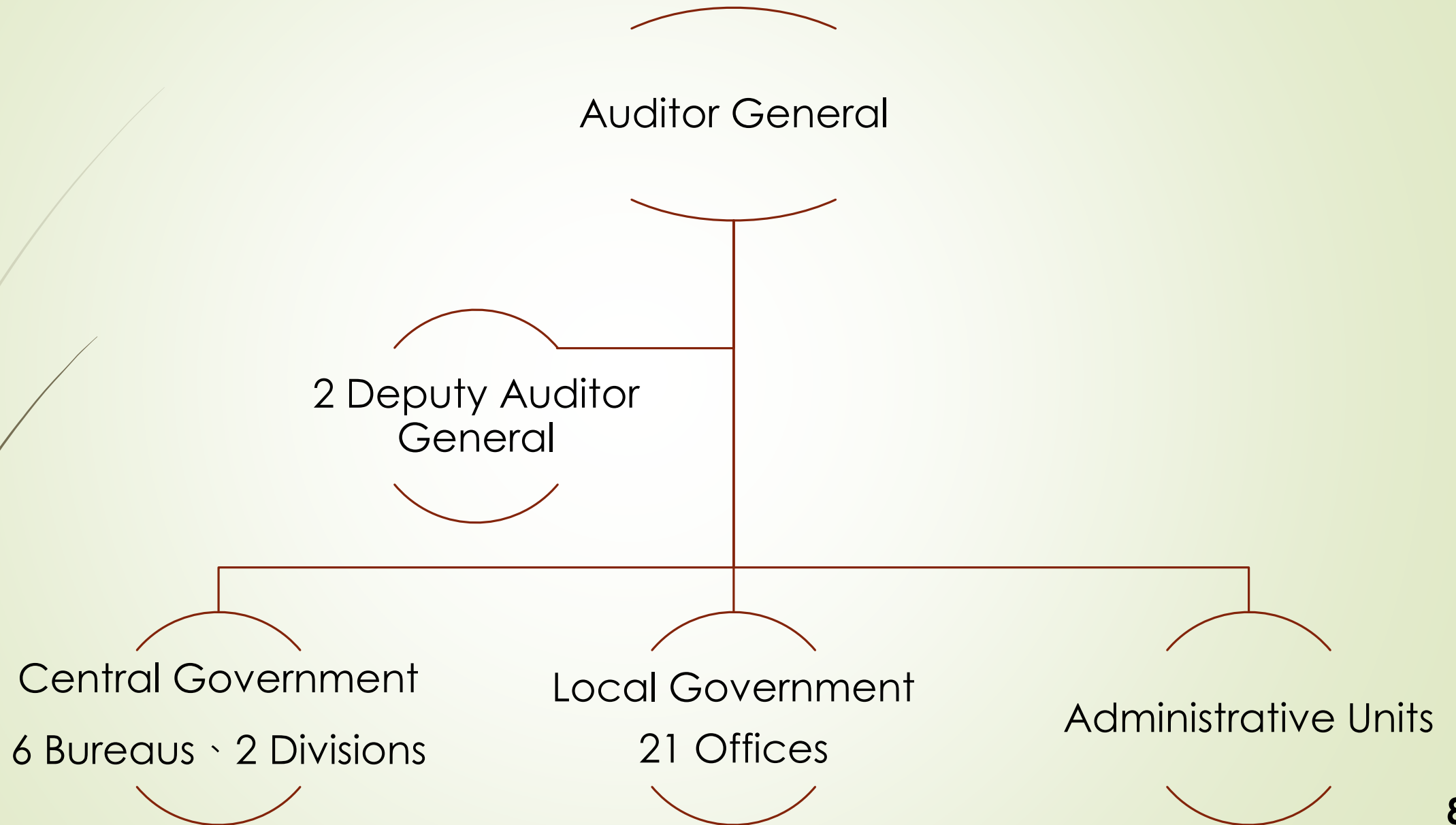
Control Yuan

# The Accountability System

Legislative Yuan




# The Organizational Chart of the NAO



# 8 Areas add NAO's Value

We focus on audits of national priorities, SDGs, digital advancements, provide insight through engaging experts and citizens in our work, and explore administrative agencies' good practices in order to maximize the value and benefits of the NAO, which is making a difference to the lives of citizens.




### Audit on Cross-cutting issues

Launch the mechanism for identifying and auditing key issues, learning from practice of GAO. **2021**

Establish Commission on Key Audit Issues. **2023**

**Specific chapters in annual audit reports**

- 8** **2020**
- 12** **2021**
- 12** **2022**
- 10** **2023**



**7 KEY AUDIT AREAS**




### Digital Auditing

Establish Sixth Bureau, starting new era of digital audit. **2022**

IT applications, e.g. ARBUTUS    

Conduct audits on cyber security & IT investment projects of the governments.




### SDGs Auditing

Establish Task Force on Auditing Implementation of Sustainable Development. **2005**

Disclose specific chapter of Audit on Sustainable Development. **Since 2015**

Release audit report of Government's Preparedness to Implement the SDGs. **2022**

Publish VDR of SDGs. **2019 & 2021**



### Administrative Agencies' Good Practices

- Discover Administrative Agencies' Good Practices during audit process.
- Launch platform for sharing good practices.
- Number of cases: 24 in 2021, 14 in 2022, 29 in 2023.



### Expert & Citizen

Consulting experts  
Committee engagement  
Participatory audit platform



### Innovation

Innovation case Review  
Creative ideas proposal  
Collaborative Innovation Camp  
AI Hackathon



### Professional Development

Training courses  
Audit Newsletter  
Digital Learning Platform



### International Affairs

International meeting  
Foreign counterparts visits  
Foreign guests greeting



# **PART II**

# **The Development of Artificial Intelligence and Its Accountability In Taiwan.**



# The Development of Artificial Intelligence In Taiwan



# National Digital Development Policies



# AI Taiwan 1.0

Innovation, Collaboration,  
Inspiration

AI for Industrial Innovation

AI International Hub

AI Pilot Project

AI Talent Program

Test Fields and Regulatory Co-creation

# AI Taiwan 2.0

Scientific & Technology Research,  
Talents, Governance

Semiconductor Manufacture

Medical and Health

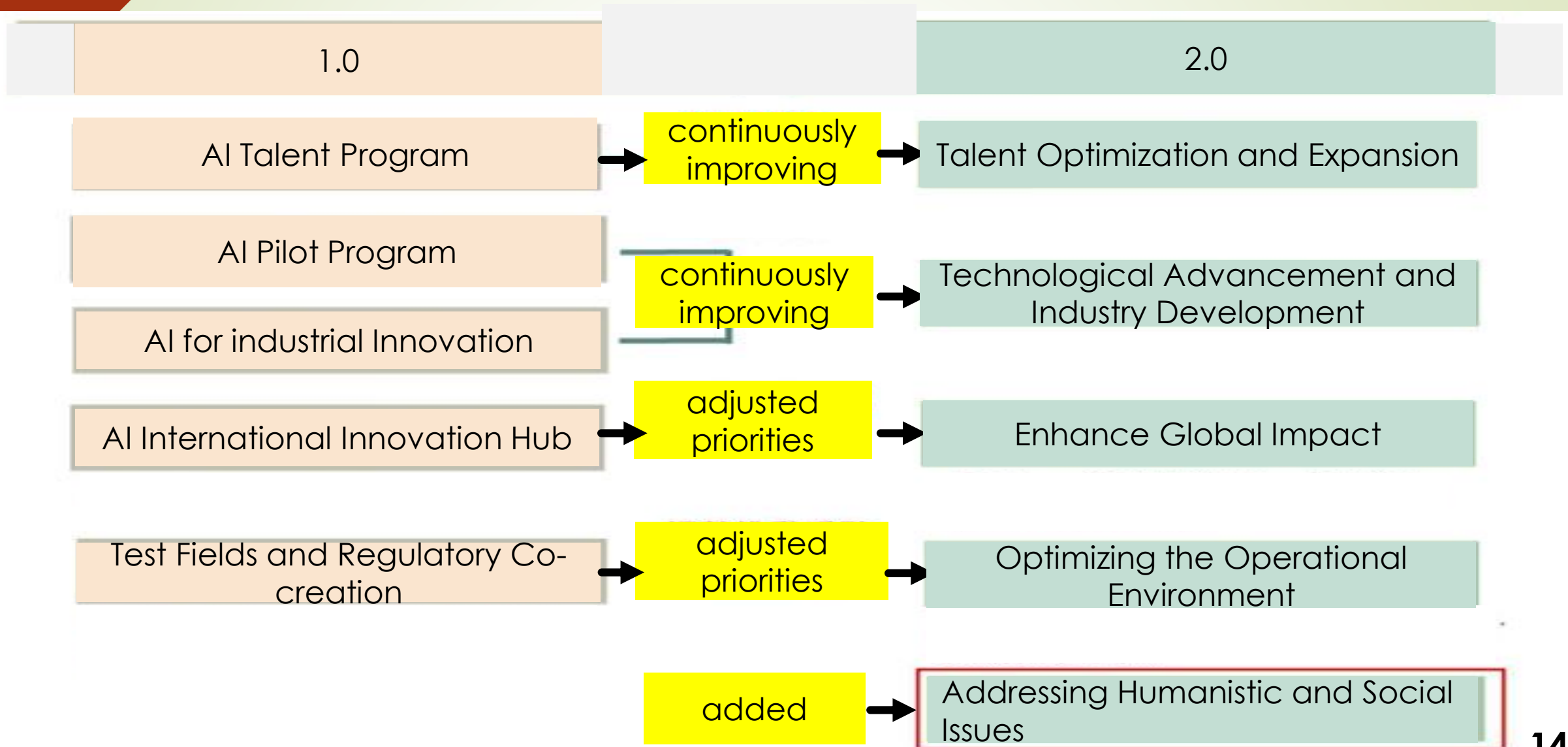
Environment

Smart City

Service

Core Tech

# The differences Between AI Taiwan 1.0 and 2.0

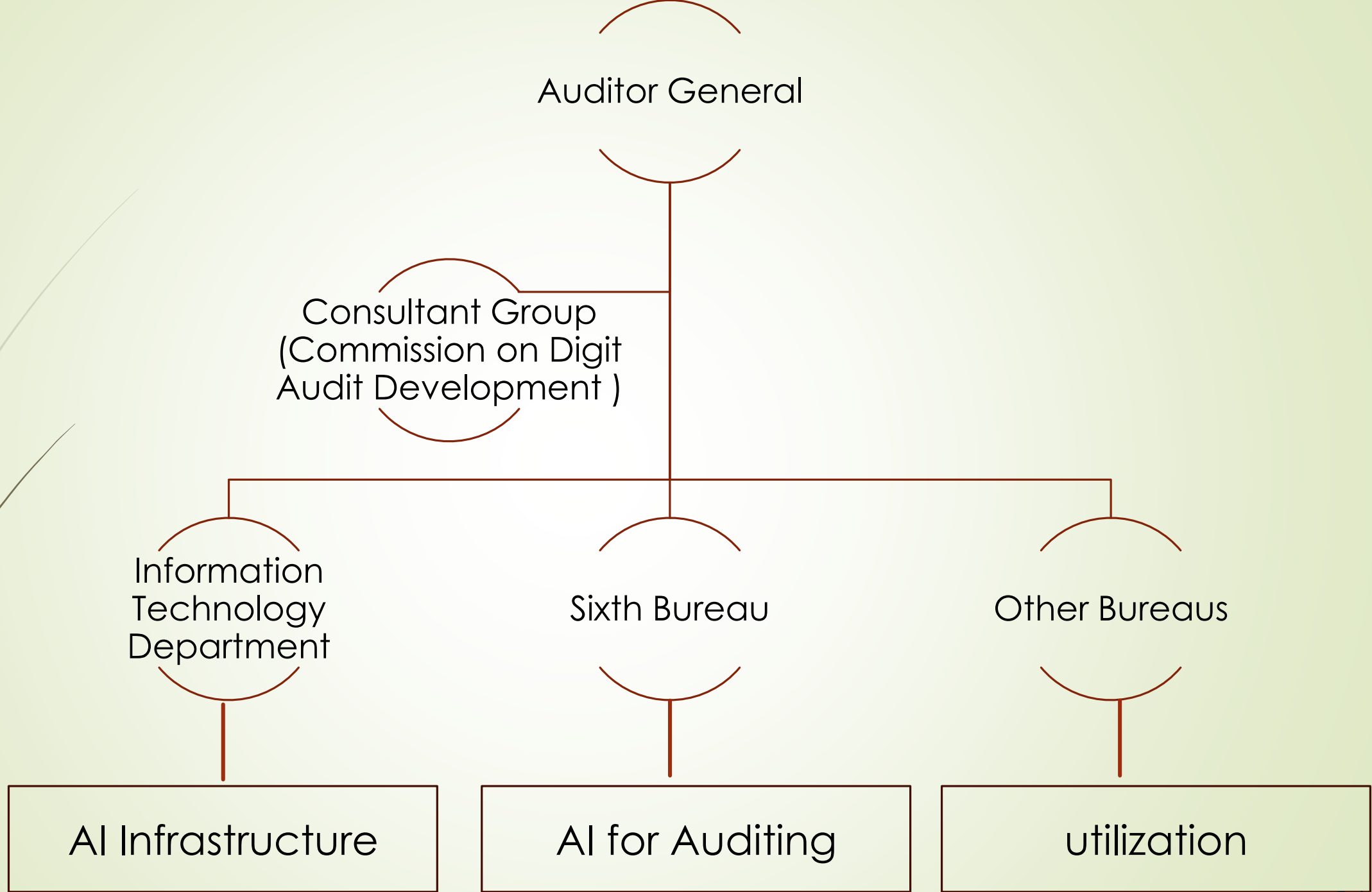




**Making a difference  
to the lives of citizens.**

---

# **The Accountability of Artificial Intelligence In Taiwan**



# The development of AI in the NAO



**Establishing the Sixth Bureau in 2022**

charging with audit of public digital service



**Applying ChatGPT to Audit**

Potential Audit Issues  
On-site Auditing



**Establishing Program of Applying AI Technology to Audit**

AuditGPT



**Introducing Training Projects**

Training Courses  
Hackathon Event

# High- Risk Audit Issues



Food Safety



Traffic Safety



Energy Transformation



Climate Change



Government Procurement





# A Real Case of Applying ChatGPT to Audit in the NAO

## Government Procurement Act

Lawbreakers' names will be published on the Government Procurement Gazette.

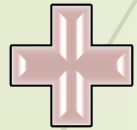
Lawbreakers will be prohibited from participating in tendering for up to 3 years.

# Big Data

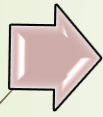
# A Real Case of Applying ChatGPT to Audit in the NAO



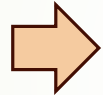
Audit Experience



ChatGPT



Program Code



In 10yrs  
2 million cases

Database of Judgement

Python

Government e-Procurement System



High-Risk Contractors v.s.

Contractors of Procurements  
(in 3 years · 7000cases)



18 Contractors



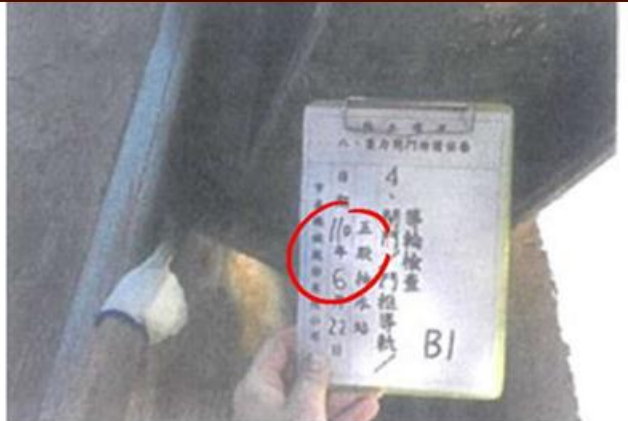
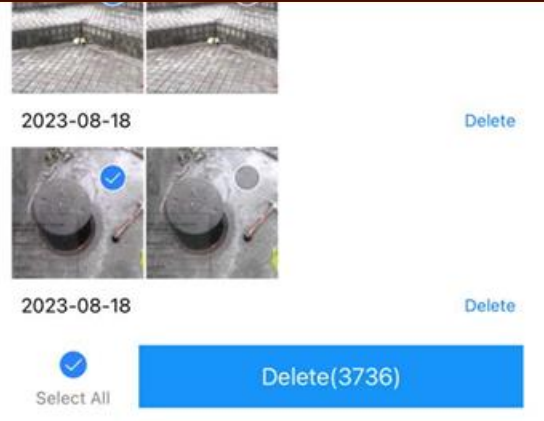
20

16 Contractors  
> € 4 Million





The same photo was used for invoicing by the same contractors for procurement cases in different years.



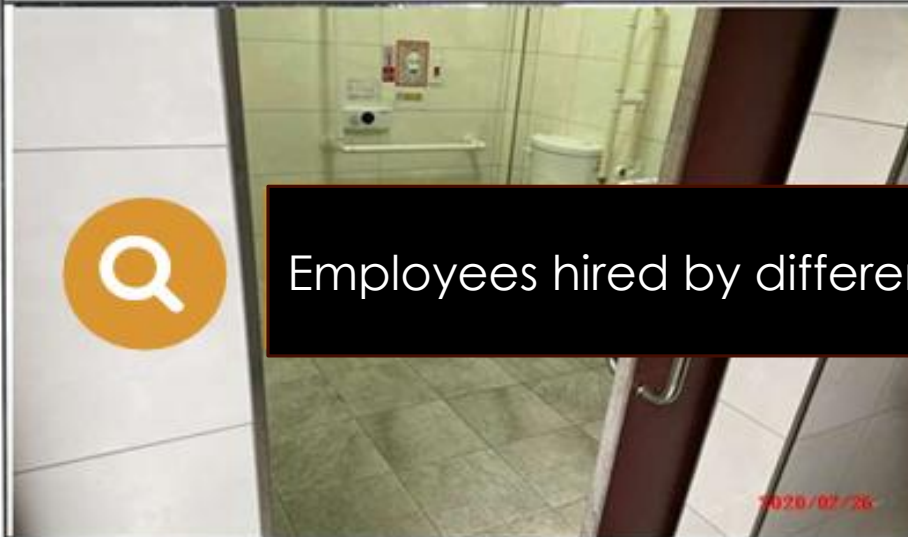
區公園廣場景觀維護(南區)

公園

施工前



施工後



區公園廣場景觀維護(北區)

公園

施工前



施工後



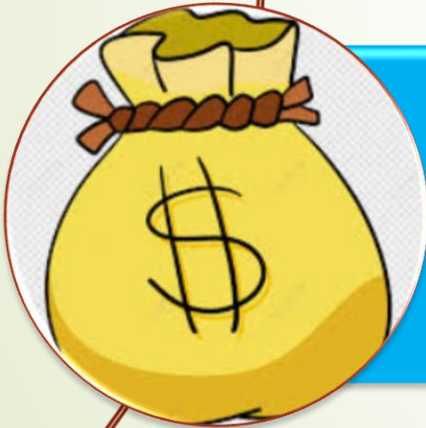
Employees hired by different contractors were found to be shared.



Without using the programs, auditors wouldn't be able to visually discern the differences in these photos.



**These cases were referred to the Ministry of Justice Investigation Bureau.**



**Part of the contract payments have been recovered.**



# The Audit on the Implementation of the Administrative Agencies' AI Projects



**Some ministries have yet to fully establish relevant legal frameworks and supporting measures.**



**It is recommended to establish a dedicated unit responsible for assessing whether AI products on the market comply with regulations.**



# In The Future

The trends in emerging technology development

The practices of international organizations and advanced countries





**The challenges to auditors examining the algorithms of government AI systems are significant. How should they be audited?**



**What are the pros and cons of third-party certification mechanism?**





*Thank you  
for Listening!*

聯合國教育、科學及文化組織 (UNESCO) 「準備狀態評估法：AI 倫理建議的工具 (Readiness Assessment Methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence)」

1.	一般面向
1.1	政府目前已經決定透過建立國家 AI 倫理委員會或其他措施，來實踐 UNESCO 建議書內容嗎？
1.2	政府是否已對從公權力介入（例如規範、策略或指導方針等）而受益於 AI 的特定群體排定優先順序？
1.3	政府是否在民眾使用涉及 AI 公共服務時，告知民眾被這些系統進行了檔案分析或決策？
1.4	哪些部門負責 AI 管理？
1.5	總體而言，在該國制定 AI 法規和政策時，最重要的挑戰是什麼？
2	法規面向
2.1	背景：法律/監管層面（包括實施和執行監管框架的能力）是應對成員國在實施建議時所需的機構和人力資源的關鍵面向，更廣泛地面對因各行各業日益採用 AI 而引起的重大社會變革。監管框架應該包括有效保護、執行、救濟和監測與 AI 系統的部署和使用相關的潛在危害的各個方面。這包括評估成員國是否採取適當的監管框架，以確保 AI 的道德開發和部署，以及監測和評估其實施和執行的機制。這一面向應有助於監控法律及具體條款的存在和有效性，以實施該建議。在隱私的情況下，這可能包括確保當前的數據隱私和數據保護在部署 AI 系統時不會受到損害、評估確保性別平等的監管，或防止濫用市場主導地位。
2.2	法律面向指標
2.2.1	AI 政策與規範
2.2.1.1 (質化問題)	是否擁有國家 AI 策略？如果沒有，是否有任何法律或策略對 AI 監管產生間接影響（例如，數據隱私或反歧視法律，或數字策略）？
評估指標	
2.2.1.1.1	是否已經評估國家 AI 策略的有效性？
2.2.1.1.2	國家 AI 策略或同等文件是否包含道德要素？
2.2.1.1.3	國家 AI 策略或同等文件是否由多元化團隊創建（包括男性和女性、少數族裔等）？
2.2.1.1.4	國家 AI 策略或同等文件是否於諮詢不同利害關係人（學術界、產業界、民間社會等）後訂定？
2.2.1.1.5	國家 AI 策略或同等文件是否提及 AI 對人權的影響？
2.2.1.1.6	國家 AI 策略或同等文件是否包括詳細的實施計畫？
2.2.1.1.7	是否有專責的機關負責推動國家 AI 策略或同等文件的實施？
2.2.1.1.8	國家 AI 策略或同等文件是否包括建議措施的預算分配？
2.2.1.1.9	國家 AI 策略或同等文件是否要求在某些領域的部署之前進行 AI

	影響評估？
2.2.1.1.10	AI 策略或同等文件是否包括監測和評估的條款？
2.2.1.2	國家是否已頒布任何具有約束力的 AI 法規（例如，採購指導方針）？如果尚未頒布任何 AI 法規，是否正在制定這類法規？
評估指標	
2.2.1.2.1	這些具有約束力的 AI 法規的有效性是否已經評估？
2.2.2	資料及隱私保護法
2.2.2.1 (量化指標)	該國在網絡安全指數上的得分。
2.2.2.2 (質化問題)	是否擁有數據保護法？如果沒有，是否正在制定這類法規？
評估指標	
2.2.2.2.1	數據保護法的有效性是否已經評估？
2.2.2.2.2	數據保護法是否賦予用戶對其數據的控制權並允許他們刪除數據？
2.2.2.2.3	數據保護法是否提到通知和同意框架，並具體說明在什麼情況下適用？
2.2.2.2.4	數據保護法是否包括數據使用的透明度要求？
2.2.2.2.5	數據保護法是否包括數據最小化的要求？
2.2.2.2.6	數據保護法是否強調需要進行數據保護或隱私影響評估的情況？
2.2.2.2.7	數據保護法是否包括針對敏感信息（例如健康數據）的具體規定？
2.2.2.2.8	數據保護法是否包括違規時的執行機制和賠償方案？
2.2.2.2.9	對於公營與私營實體所收集的數據，是否適用不同的數據保護標準？
2.2.2.2.10	數據保護法或其他法律是否保護隱私和/或尊重私人 and 家庭生活？
2.2.2.2.11	政府是否已設置資料保護專責機關或官員？
2.2.3	資料分享及可取得性
2.2.3.1 (量化指標)	在開放清單之得分
質化問題	
2.2.3.2	是否已簽屬國際資料開放憲章（Open Data Charter）？
2.2.3.3	是否已有國家數據共享框架？如果沒有，是否正在研擬中？
評估指標	
2.2.3.3.1	是否已經評估國家數據共享框架的有效性？
2.2.3.3.2	數據共享框架如何處理公共部門與私營部門之間、不同地理區域等的數據共享？
2.2.3.4	是否有開放政府數據政策？如果沒有，是否正在採納其他類似政策？
評估指標	

2.2.3.4.1	是否已經評估開放政府數據政策的有效性？
2.2.3.4.2	開放政府數據政策是否提到將數據集提供和可供研究使用？
2.2.4	採購法規與政策
2.2.4.1 (量化指標)	是否制定有關於採購 AI 或包含 AI 組件的產品/服務的法律或政策？如果沒有，這些法律或政策是否正在被採納？
評估指標	
2.2.4.1.1	已經被評估這些法律或政策的有效性是否？
2.2.4.1.2	在購買 AI 之前是否有特別的批准程序？
2.2.4.1.3	是否有包含人工智能相關條款的認證供應商名單？
2.2.4.1.4	認證僅涵蓋技術層面，還是同時涵蓋技術和道德層面？
2.2.4.1.5	認證是否僅包括人工智能系統，還是同時包括人工智能技術進入公共系統的代理？
2.2.4.1.6	該認證是自願的還是強制的？
2.2.5	資訊自由法/知識獲取法
2.2.5.1 (量化指標)	是否有資訊自由法？如果沒有，該法案是否正在被採納？如果沒有，原因是什麼？
評估指標	
2.2.5.1.1	此法案的有效性是否已經被評估？
2.2.5.1.2	根據此法案，個人是否可以請求有關 AI 在公共部門中使用方式的資訊？
2.2.5.1.3	使用和/或共享數據的各方是否有義務通知那些他們正在使用和/或共享數據的人？
2.2.6	正當程序及課責制度
2.2.6.1 (質化指標)	正當程序權利的主要法律是什麼？如果沒有，這樣的法律或政策是否正在通過？
評估指標	
2.2.6.1.1	該法律的效力是否經過評估？
2.2.6.1.2	在某些情況下，必須告知個人他們正在與 AI 進行互動嗎？
2.2.6.2	是否有法律或政策強調針對 AI 造成的危害的監控、糾正和補救機制？如果有，是哪些機制？如果沒有，這樣的法律或政策是否正在通過？
評估指標	
2.2.6.2.1	該法律的效力是否經過評估？
2.2.6.2.2	監管機構或法院是否可以根據本法律/政策索取有關 AI 及其內部運作的資訊？
2.2.7	網路安全和言論誠信
2.2.7.1 (質化指標)	是否有一個框架來通知和刪除違規內容（例如網路仇恨言論、錯誤訊息和虛假訊息）的政策？如果沒有，這樣的框架是否正在被採用？
評估指標	
2.2.7.1.1	有評估過此法律或政策之有效性嗎？
2.2.7.1.2	這些架構適用於那些案件

2.2.7.1.3	該法是否明確規範網路中間商的責任？
2.2.7.1.4	根據該法，網路中間商應承擔哪些責任？
2.2.7.2	是否頒布了有關 AI 對社群媒體影響的任何法律或政策，包括透明度、錯誤訊息、虛假訊息和仇恨言論？ 如果沒有，這樣的框架是否正在被採用？
評估指標	
2.2.7.2.1	有評估過此法律或政策之有效性嗎？
2.2.8	公部門能力
2.2.8.1 (質化指標)	是否有政府策略/計畫來提高公共部門的數位技能？
評估指標	
2.2.8.1.1	有數位技能提升計畫嗎？
2.2.8.1.2	如果有，新進人員是否必須參加數位技能提升計畫？
<b>3</b>	<b>社會/ 文化面向</b>
3.1	背景：考慮與 AI 的道德發展和部署相關的因素，包括包容性、社會和文化多樣性。如果開發和部署 AI 的團隊非常同質，可能會導致 AI 無法充分反映社會組成的複雜性和多樣性，這意味著 AI 產生的結果可能會放大結構性偏見。 其次，它將解決人們對 AI 技術的態度，包括大眾的接受程度。它還應有助於揭示成員國的社會價值觀和偏好，從而導致對技術的某些態度並直接影響相關的社會選擇。 其旨在解決的一些問題如下：首先，它將解決尊重包容性和社會文化多樣性的問題，包括 AI 生命週期各個階段的性別代表性水準，以及受該技術影響的不同社區和少數群體。這一方面旨在解決目前 AI 領域中女性和少數族裔的差距和代表性不足，包括缺乏女性開發人員、研究人員、教授等。
3.2	社會/ 文化面向指標
3.2.1	多元化、包容性及平等
評估指標 (量化指標)	
3.2.1.1	使用網際網路的性別差距
3.2.1.2	使用網際網路的城鄉差距
3.2.1.3	STEM 課程中男性/女性高等教育畢業生的百分比
3.2.1.4	在科學或數學領域希望成為 STEM 專業人員的頂尖女孩/男孩的比例
3.2.1.5	男、女在科學表現的差距
3.2.1.6 (量化指標)	是否已制定任何法律或政策來縮小性別之數位差距？如果沒有，這樣的法律或政策是否正在制定中？
評估指標	
3.2.1.6.1	是否已評估此法律或政策之有效性？
3.2.1.7 (質化指標)	是否已制定任何法律或政策來縮小城鄉之數位差距？如果沒有，這樣的法律或政策是否正在制定中？
評估指標	
3.2.1.7.1	是否已評估此法律或政策之有效性？



3.2.1.8 (質化指標)	是否已頒布任何與增強 AI 勞動力多元化相關的法律或政策？如果沒有，這樣的法律或政策是否正在制定中？
評估指標	
3.2.1.8.1	是否已評估此法律或政策之有效性？
3.2.1.8.2	科技公司是否需要發布多元化統計數據？
3.2.1.8.3	是否採用平權行動標準來提高整個 AI 生命週期的多樣性？
3.2.1.8.4	是否有加強 STEM 多樣性的計畫？
3.2.1.8.5	大學是否需要發布多元化統計資料？
3.2.1.8.6	政府承包商是否需要遵守多元化標準？
3.2.1.9	是否有線上內容和資料可用於所在國家/地區的所有官方語言訓練 AI 系統？
3.2.1.10	是否有線上內容和資料可用於用原住民語言訓練 AI？
3.2.2 (量化指標)	公眾參與與信任
3.2.2.1	該國線上服務指數得分
3.2.2.2	電子參與指數得分
3.2.2.3	信任政府網站和應用程式情形
3.2.2.4	對 AI 信任情形
3.2.3.1 (質化指標)	是否已頒布政策以解決 AI 對環境和永續性的影響？
評估指標	
3.2.3.1.1	是否已評估該項政策的有效性？
3.2.3.1.2	是否明確提及 SDG 和/或 ESG？
3.2.3.1.3	是否有具體考慮 AI 對土地和水資源利用的影響？
3.2.3.1.4	在某些情況下使用 AI 之前是否必須進行環境影響評估？
3.2.3.1.5	是否有具體考慮 AI 需求對能源及其相關碳足跡的環境影響？
3.2.3.1.6	是否已具體考慮 AI 對環境的影響？(例如，自動駕駛個人車輛對交通相關溫室氣體排放的影響，或利用 AI 增加化石燃料勘探)
3.2.4	健康及社會福利
3.2.4.1 (質化指標)	是否已採用數位醫療政策？
評估指標	
3.2.4.1.1	數位醫療政策(或同等政策)的有效性是否經過評估？
3.2.4.1.2	數位健康政策(或同等政策)是否提及 AI 技術？
3.2.4.1.3	數位健康政策(或同等政策)是否涵蓋身心健康？
3.2.4.1.4	數位健康政策(或同等政策)是否考慮 AI 對兒童的影響？
3.2.5.1 (質化指標)	是否實施了有關利用 AI 保護文化遺產的政策？
評估指標	
3.2.5.1.1	該政策的效果是否經過評估？
3.2.5.1.2	該政策是否允許社區成員參與 AI 的開發或監管以保護文化遺產？如果沒有，是否有保護少數族裔和原住民語言的政策提到 AI

	和數位科技的影響？
3.2.5.2 (質化指標)	是否已實施有關使用 AI 保護文化遺產的政策？如果沒有，是否有文化遺產保護政策提到 AI 和數位科技的影響？
評估指標	
3.2.5.2.1	該政策的效果是否經過評估？
3.2.5.2.2	這項政策是否允許社區成員參與 AI 的開發或監管以保護少數民族和原住民語言？
<b>4</b>	<b>科學/教育面向</b>
4.1	背景：此面向旨在評估一個國家 AI 相關的研發水準，包括 AI 相關出版物和專利的數量，以及參與研發的 AI 研究人員和工程師的數量。它還將調查 AI 倫理研究，例如反映關注 AI 倫理的出版物的數量。教育面向可以包括為學生提供教育機會，例如 AI 相關學位課程、AI 開發人員的終身教育計劃以及公眾的教育機會。例如，這方面將檢查是否有專門的計劃來熟悉公眾並為他們提供與技術相關的技能，這些技能可能有助於跟上新的工作要求。除了教育機會之外，該面向還將檢視 AI 或資訊通信技術相關學科的學生（畢業生）、專業人士和公眾的數量。它還可以包括 STEM 畢業生的數量，作為 AI 開發和部署的重要前提。此面向有可能解決人口中 AI 和電子技能的水準。最後，該面向將包括為 AI 相關和非 AI 相關學位的學生提供 AI 道德教育，以及為經理、開發人員或產品設計師等專業人士提供 AI 道德課程。
4.2.1	研究與創新
4.2.1.1 (量化指標)	科學研究經費支出
評估指標	
4.2.1.1.1	研究與開發總支出 (GERD) 佔國內生產毛額 (GDP)
4.2.1.1.2	自然科學和工程研究與開發總支出 (GERD) 佔國內生產毛額 (GDP) 的比例
4.2.1.1.3	政府研發預算撥款 (GBARD)
4.2.1.1.4	政府對 AI 研究和開發的政府資助有估算嗎？
4.2.1.2 (量化指標)	研究產出
評估指標	
4.2.1.2.1	人均 AI 數量及人均 AI 相關出版數量
4.2.1.2.2	AI 及 AI 相關出版的人均引用次數
4.2.1.2.3	人均 FAcT 出版數量
4.2.1.3 (量化指標)	道德 AI 的研究
評估指標	
4.2.1.3.1	人均 AI 倫理出版數量
4.2.1.3.2	國內每年人均舉辦的 AI 倫理會議數量
4.2.1.3.3	均致力於 AI 倫理的研究中心和/或部門數量

4.2.1.3.4	人均涵蓋 AI 倫理的 AI 研究中心和/或部門數量
4.2.1.4 (量化指標)	AI 人才
評估指標	
4.2.1.4.1	大學 AI 研究人員(電腦科學家、資料科學家、機器人專家、AI 倫理研究人員)人均數量
4.2.1.4.2	人均 Kaggle 大師數量
4.2.1.5 (量化指標)	創新產出
評估指標	
4.2.1.5.1	人均 AI 專利授權量
4.2.1.5.2	GitHub 上的人均 AI 儲存庫提交數量
4.2.2	教育
4.2.2.1 (質化指標)	教育策略
4.2.2.1.1	是否有任何法律或政策將 AI 或其他數位工具整合到教育系統中?
評估指標	
4.2.2.1.1.1	是否對這項法律或政策的功效進行過評估?
4.2.2.1.1.2	是否有關於如何培訓教育工作者講授 AI/技術道德的法律或政策?
評估指標	
4.2.2.1.2.1	是否對這項法律或政策的功效進行過評估?
4.2.2.2 (量化指標)	教育基礎設施
4.2.2.2.1	小學、國中和中學為教學目的連接網路的比例。
4.2.2.2.2	小學、國中和中學擁有教學用電腦的比例。
4.2.2.3 (量化指標)	課程內容
4.2.2.3.1	人均致力於 AI、機器學習或資料科學的高等教育課程數量。
4.2.2.3.2	人均提供 AI、機器學習或數據科學一項或多項模組的高等教育課程數量。
4.2.2.3.3	人均提供數位人類學、技術哲學、AI 倫理學或相關/類似學科的一個或多個模組的高等教育課程數量
4.2.2.3.4 (量化指標)	是否有既包含 AI 技術又包含倫理方面的教育計畫(技術面可能包括編碼、機器學習、統計、資料科學等;倫理面向可能包括資訊倫理、科技哲學、隱私問題、科技的社會影響等)?
評估指標	
4.2.2.3.4.1	是否在小學教育中,可以包括讓學生熟悉程式設計或數位彈性的課程(例如,線上安全、螢幕時間、數位素養)有既包含 AI 技術又包含倫理方面的教育計畫(技術面可能包括編碼、機器學習、統計、資料科學等;倫理面向可能包括資訊倫理、科技哲學、隱

	私問題、科技的社會影響等)？
4.2.2.3.4.2	是否在中學教育中，可以包括讓學生熟悉程式設計或數位彈性的課程（例如，線上安全、螢幕時間、數位素養）有既包含 AI 技術又包含倫理方面的教育計畫（技術面可能包括編碼、機器學習、統計、資料科學等；倫理面向可能包括資訊倫理、科技哲學、隱私問題、科技的社會影響等）？
4.2.2.3.4.3	是否在大學教育中，可以包括讓學生熟悉程式設計或數位彈性的課程（例如，線上安全、螢幕時間、數位素養）有既包含 AI 技術又包含倫理方面的教育計畫（技術面可能包括編碼、機器學習、統計、資料科學等；倫理面向可能包括資訊倫理、科技哲學、隱私問題、科技的社會影響等）？
4.2.2.3.4.4	在持續教育、職業學校和技術/專業培訓機構中，可以包括讓學生熟悉程式設計或數位彈性的課程（例如，線上安全、螢幕時間、數位素養）有既包含 AI 技術又包含倫理方面的教育計畫（技術面可能包括編碼、機器學習、統計、資料科學等；倫理面向可能包括資訊倫理、科技哲學、隱私問題、科技的社會影響等）？
4.2.2.4 (量化指標)	教育程度
4.2.2.4.1	STEM 畢業生接受高等教育的比例
4.2.2.4.2	高等教育中資訊通信技術畢業生的比例
4.2.2.4.3	人均數據科學、機器學習或機器人課程畢業生人數
4.2.2.4.4	人均 AI 相關博士人數
4.2.2.4.5	人均 AI 相關博士後人數
4.2.2.4.6	Coursera 全球技能報告數據科學排名
4.2.2.5 (質化指標)	AI 教育普及率
4.2.2.5.1	是否有針對一般人群的 AI 技術課程？
評估指標	
4.2.2.5.1.1	如果是，它們是否免費且有多種語言版本？
4.2.2.5.2	是否有針對一般大眾的 AI 倫理課程或模組？
評估指標	
4.2.2.5.2.1	如果是，它們是否免費且有多種語言版本？
<b>5</b>	<b>經濟面向</b>
5.1	背景：旨在解決該國 AI 生態系統供應方的規模和實力，這對於開發反映特定國家及其人口的特殊需求和條件的 AI 解決方案的能力非常重要。它將著眼於科技產業的規模，包括開發或部署 AI 系統及其員工的公司數量。它還將解決 AI 領域公共和私人投資的數量問題。它將有助於追蹤 AI 產業的成長。增強這方面的能力將關係到支持該國 AI 生態系統發展的能力，包括對 AI 技術和人才的投資吸引力。
5.2	經濟面向指標
5.2.1	勞動市場

(量化指標)	
5.2.1.1	需要 AI 相關技能的職缺比例 (最好是線上職缺)。
5.2.1.2	目前擔任資料科學家的員工比例。
5.2.1.3	相對 AI 技能滲透率。
5.2.1.4	AI 人才集中。
5.2.1.5 (質化指標)	是否有應對 AI 對勞動市場影響的策略? 這包括受自動化影響的工人的重新技能、提高工人的技能以利用 AI 提供的機會以及考慮人類技能相對於 AI 系統的軟技能優勢和互補性等問題。
評估指標	
5.2.1.5.1	該策略的有效性是否經過評估?
5.2.2 (量化指標)	中間消費
5.2.2.1	企業在 AI 服務 (包括軟體即服務) 上的支出佔中間消耗的比例是多少?
評估指標	
5.2.2.1.1	這些 AI 服務傾向於國產還是進口?
5.2.3 (量化指標)	投資及產出
評估指標	
5.2.3.1	企業電腦程式設計、諮詢及相關活動部門的人均研發支出。
5.2.3.2	SIC 代碼 62.0 (電腦程式設計、諮詢和相關活動) 的人均 GDP。
5.2.3.3	高科技出口占貿易的比重。
<b>6</b>	<b>技術及基礎建設面向</b>
6.1	背景: 如果沒有相關的基礎設施, AI 的開發和基於 AI 的解決方案的實施就無法在全國範圍內推廣。因此, 此面向旨在評估 ICT 和相關技術基礎設施的水平。除此之外, 該面向將評估網路連接和存取、資料中心的可用性、雲端運算能力和超級電腦。鑑於數據對於 AI 技術的至關重要, 該面向的另一個方面涉及高品質數據的可用性以及確保數據具有代表性的實踐。值得注意的是, 該面向下的許多指標已經透過不同的指數來衡量, 在準備方法中, 我們將預先填寫答案, 並為各國提供在需要時更新答案的機會。
6.2	技術面向指標
6.2.1 (量化指標)	基礎建設及關連性
6.2.1.1	擁有行動電話的人口比例。
6.2.1.2	擁有固定寬頻電話的人口比例。
6.2.1.3	活躍使用行動寬頻用戶的人口比例。
6.2.1.4	平均國際頻寬。
6.2.1.5	平均固定寬頻下載速度。
6.2.1.6	使用網路的人口比例。
6.2.1.7	3G 以上行動網路覆蓋的人口比例。
6.2.1.8	使用電力的人口比例。

6.2.1.9	使用網路的性別差距。
6.2.1.10	使用行動電話的性別差距。
6.2.1.11	城鄉網路存取差距（家庭）。
6.2.2 （質化指標）	適用標準
6.2.1.1	是否參與 AI 和數位技術的標準化（技術和道德）？（ISO/IEC、IEEE 7000）
6.2.2.1.1	如果沒有，是否申請參與此流程？
6.2.3 （量化指標）	運算技術
6.2.3.1	全國人均資料中心數量
6.2.3.2	到最近資料中心的距離
6.2.3.3	託管資料中心
6.2.3.4 （質化指標）	是否有 AI 驅動的雲端運算政策？如果沒有，是否正在研擬？
評估指標	
6.2.3.4.1	之前是否評估過此類政策的有效性？
6.2.4 （量化指標）	統計表現
6.2.4.1	統計績效指標
6.2.4.1.1	數據產品得分
6.2.4.1.2	資料來源得分
6.2.4.1.3	數據基礎設施得分
6.2.4.2 （質化指標）	是否有任何法律或政策為一致的資料管理和發布提供全面的框架？
評估標準	
6.2.4.2.1	之前是否評估過此類政策的有效性？
6.2.4.2.2	針對政府資料是否有明確規範品質管制流程

聯合國教育、科學及文化組織 (UNESCO) 「倫理影響評估法：AI 倫理建議的工具 (Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence)」

界定範圍	
1.	專案概述
1.1	描述系統
1.1.1	描述規劃設計、開發或部署的 AI
1.1.2	描述該系統的目的或目標。如果目的是解決特定問題，具體說明預計解決的問題。請同時說明系統如何適應更廣泛的工作方案。
1.1.3	參考專案生命週期描述專案的當前狀態。
1.1.4	回答下列問題，描述 AI 的功能請注意（此類問題係基於 OECD 的 AI 分類框架）
1.1.4.1	與此系統互動的用戶是誰（包括他們的能力水準）？
1.1.4.2	使用者有什麼程度的選擇權？
1.1.4.3	該 AI 將應用在哪些領域？
1.1.4.4	該 AI 將用於哪些業務功能？
1.1.4.5	對關鍵職能和活動的影響
1.1.4.6	描述 AI 部署的廣度。
1.2	相關性
1.2.1	該項目是現有項目的擴展或改編嗎？如果有，之前是否進行過評估？如果是這樣，自初次評估以來系統的哪些功能發生了變化？
1.2.2	AI（包括核心模型）是為此特定目的或目標而開發的，還是基於現成的模型（例如 BERT、ChatGPT 等）建構的？
1.2.3	列出系統對其他非直接開發的模型或非直接使用的資料的相關性。
2.	比例篩選及不造成傷害
2.1	在採購 AI 時，必須考慮使用 AI 的目標和相關的特定系統，以及該技術在預期目的是否值得使用方面的相稱性，同時考慮 AI 的風險、不確定性和缺點。這種反思使採購人員能夠在實踐中的手段和預期目標之間取得平衡，以證明使用特定方法或系統的必要性並證明其適用性，確保與 AI 相關或屬於 AI 的一部分的流程不超過實現合法目標所需的範圍。重要的是，該建議強調“對人權和基本自由的任何可能的限制必須有合法基礎，並且合理、必要和相稱，並符合各國根據國際法承擔的義務。此外，雖然建議中闡述的所有原則和價值觀都是重要且可取的，但在實踐中，它們有時可能會發生衝突，例如，當對透明度和可解釋性的需求可能會影響保護隱私和資料保護的能力時，就會出現這種情況因此，比例原則也可以在需要時發揮關鍵作用，幫

	助根據具體情況調和不同道德原則和/或優先事項之間的緊張關係，同時仍然尊重人權和基本自由。
2.2	建立比例性
2.2.1	是否仔細考慮可用於實現相同目標的非演算法選項？如果有，為什麼涉及 AI 的選項受到青睞？
2.2.2	是否考慮不同的 AI？選擇這種特定方法的理由是什麼？
2.2.3	該項目的範圍是否已明確界定？為了確保其與既定目標保持相稱，對該專案的範圍施加了哪些限制？
2.2.4	[用於事後分析] 此系統在實現其既定目標方面效果如何？
3	專案治理（確定角色和職責）
3.1	確保確定參與者以實現透明度並避免專案團隊內任何令人困惑的責任擴散至關重要。由於 AI 具有固有的風險，因此確定誰對 AI 的哪個方面負責非常重要。此外，專案團隊應特別注意利害關係人（包括潛在用戶，特別是來自邊緣社群的用戶）的代表性是否不足。專案團隊內部缺乏多元化意味著可能會缺少某些觀點，這可能會對無人代表或代表性不足的社群造成更大的傷害。相較之下，專案團隊內部的更多多樣性可能有助於及早發現偏見並減輕危害。此外，允許使用者提供回饋以促進模型開發至關重要，因為使用者通常比開發人員更加多樣化，並且可能會更早注意到這些問題。
3.2	角色和職責
3.2.1	負責此 AI 的專案團隊中誰擁有最終決策權？
3.2.2	請描述誰負責專案的主要工作流程，包括第三方或外部組織的任何代表。包括團隊內角色和職責的完整描述，以及所涉及的不同個人和組織的全貌。
3.2.3	是否考慮 AI 專案團隊的多樣性，特別是在（但不限於）性別、年齡、種族、膚色、血統、語言、宗教、國籍、民族、社會出身、經濟或社會條件方面，殘疾和性取向，包括這如何反應預期用戶群體的複雜性和多樣性，以及這如何引入偏見？
3.2.4	為解決這些問題並考慮專案團隊可能缺少哪些觀點，請以個人或理想情況下以團隊的身份進行立場反思。
4	多利害關係人治理
4.1	專案團隊應在系統設計的早期階段製定利害關係人參與計劃。此利害關係人參與計畫將允許專案團隊制定他們的參與目標，應定期審查該目標，以確保利害關係人參與不僅作為清單練習完成，而且構成決策的一個組成部分和變革性面向過程



4.2.1	哪些利害關係人最有可能受到 AI 部署的影響？在這裡，專案團隊列出了受影響的利害關係人並確定了顯著的利害關係人（考慮受保護的特徵和背景脆弱性如何相互交叉，從而使特定利害關係人更容易受到不利影響。
4.2.2	基於上一個問題和立場反思，在 AI 的開發、部署和使用過程中，將涉及或諮詢哪些利害關係人？
4.2.3	讓這些利害關係人參與其中的目標是什麼？（例如，確保所有不利影響都得到識別，提高社區對正在設計的系統的信任程度
4.2.4	吸引利害關係人的計畫是什麼？
<b>實施聯合國教科文組織原則</b>	
5	安全與保障
5.1	然存取更多資料通常被認為是增強安全性和準確性的機會，但這意味著攻擊者還可以更快地學習並使用 AI 來不斷改進他們的攻擊，將速度與上下文相結合。AI 中的安全意味著，就像新車在被允許上路之前必須經過安全測試一樣，新藥物在銷售給消費者之前必須滿足嚴格的安全標準一樣，AI 也具有遵守監管、技術和社會標準。許多 AI 的黑盒子特徵進一步強調了遵守安全標準的必要性，因為這些技術很複雜並且通常結合了多個系統。AI 通常處理大量（有時是敏感）資料。這對資料安全構成重大威脅。鑑於敏感的個人資料往往是網路攻擊的目標，資料外洩的風險是一個主要問題。攻擊的主要類型包括：資料中毒：透過更改訓練資料或其標籤來操縱模型的行為；輸入操縱：此技術需要輸入惡意內容來欺騙系統。考慮到大型語言模型的興起，這是一個非常相關的問題，因為它包括直接攻擊的情況，即有人在 ChatGPT 或 Bing Chat 中插入提示以嘗試使其以不同的方式運行，以及依賴數據的間接攻擊。這些資訊可以讓攻擊者在資訊片段和個人身分之間建立更緊密的聯繫，從而可能推斷出敏感的個人資料。為了緩解這些問題，AI 需要定期進行測試和重新驗證，特別是當它們用於醫療保健等敏感環境中的決策時。
5.2	程序評估
5.2.1	採取了哪些措施來確保 AI 的安全並防止其受到系統操縱？
5.2.2	採取了哪些措施來確保 AI 訓練資料的安全性和安全性，防止資料中毒/損壞？
5.2.3	採取了哪些措施來確保 AI 處理資料的安全性？
5.2.4	如果 AI 正在處理的訓練資料或資料被中毒或損壞，或者您的系統被操縱，您怎麼知道？
5.2.5	AI 在使用前是否經過測試？
5.2.6	如果 AI 已經投入使用，AI 投入使用後是否進行了進一步

	的測試和重新驗證？5.3					
5.3	識別和減輕影響					
5.3.1	正面影響：考慮 AI 的設計、開發、部署和使用可能產生的所有結果，無論是正面的還是負面的。					
	該系統對安全和安保的預期積極影響是什麼？	評估預期正面結果的規模	評估預期正向結果的範圍	評估預期正面結果發生的可能性		
		顯著水準 • 非常高 • 高 • 中 • 中/次要	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非預期 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	發生的可能性 • 低 • 中 • 高 • 非常高		
5.3.2	負面影響					
	該系統對安全和安保的預期負面/不利影響是什麼？	評估預期負面影響的規模	評估預期負面影響的範圍	評估預期負面影響的可補救性	評估發生預期負面影響的可能性	上述程序保障措施在多大程度上減輕了這種影響？需要實施哪些額外的緩解和補救策略來應對這種潛在危害？
		嚴重程度： • 災難性 • 嚴重 • 重大	影響程度 <u>受影響方</u> • 主要 • 次要	• 非常低 • 低 • 中 • 高	• 低 • 中 • 高 • 非常高	

		<ul style="list-style-type: none"> <li>• 中/輕微程度</li> </ul>	<ul style="list-style-type: none"> <li>• 意外 / 非預期</li> <li>• <u>時間尺度</u></li> <li>• 短期</li> <li>• 中期</li> <li>• 長期</li> <li>• 跨世代</li> </ul>			
6	公平、不歧視、多元化					
6.1	<p>部分分析系統的分析發現，膚色較深的人，尤其是女性，整體而言更有可能被錯誤分類。對於某些種族、口音很重或不會說母語的使用者來說，語音辨識系統也可能不太容易使用。研究人員也強調了模擬世界中的偏見已轉移到 AI 甚至被 AI 放大的許多其他方式，使系統難以存取並導致歧視性結果。這種情況往往源自於 AI 領域的資料集、資料來源和技術團隊往往缺乏多樣性，並因這一事實而加劇 因此，為了維護公平並防止 AI 永久存在歧視，採購團隊應確保有適當的流程來測試偏見，例如對資料集進行交叉演算法偏見審計並明確公平性如何在演算法中解決，同時也積極努力促進多元和包容性。</p>					
6.2	<p>程序評估</p> <p>在本節中，在回答有關特定群體測驗的問題時，專案團隊應特別考慮（但不僅限於）種族、膚色、血統、性別、年齡、語言、宗教、政治觀點、國籍、民族、社會出身、出生的經濟或社會狀況以及殘疾。請具體說明是否對結合了其中幾個標準的組進行了測試，即係統是否已根據交叉性進行了測試。</p>					
6.2.1	防止歧視的努力					
6.2.1.1	該演算法是否經過不同組別的測試？					
6.2.2	資料品質及防止歧視偏差					
6.2.2.1	是否有適當的流程來測試數據是否存在偏差？					
6.2.2.1.1	是否對數據進行了分析以防止數據出現社會和歷史偏差？					
6.2.2.1.2	數據是否均衡，是否反映了目標最終使用者群體的多樣性？					
6.2.2.1.3	可以預見用於訓練的資料和 AI 處理的資料之間是否存在任何差異，這可能導致 AI 產生歧視性結果或針對不同群體表現出差異？					
6.2.2.1.4	是否制定了流程來記錄如何在設計過程中解決資料品質問題？					
6.2.2.1.5	是否採取了教育和意識舉措來幫助 AI 設計師和開發人員					

	認識到他們在 AI 的設計和開發中可能引入的偏見？				
6.2.3	防止可接觸方面的歧視				
6.2.3.1	該設計是否允許所有人，尤其是邊緣群體，存取 AI 並與之互動？請指定可訪問性方面的任何限制				
6.2.3.2	從技術角度如何體現公平原則？例如，您能否具體說明 AI 校準的公平性技術概念是什麼？（例如，個人公平、人口平等、機會均等等）。				
6.2.3.3	AI 將應用於哪一部分人群？受影響的人群是否特別邊緣化？				
6.3	確認及減輕影響				
6.3.1	正面影響				
	該制度對公平、非歧視和多樣性的預期積極影響是什麼？	評估預期正面結果的規模	評估預期正面結果的範圍	評估預期正面結果發生的可能性	
		顯著水準 • 非常高 • 高 • 中 • 中/次要	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非預期 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	發生的可能性 • 低 • 中 • 高 • 非常高	
	該制度對公平、非歧視和多元化的預期負面/不利影響是什麼？	評估預期負面影響的規模	評估預期負面影響的範圍	評估預期負面影響的可補救性	發生的可能性
		嚴重程度： • 災難性 • 嚴重 • 重大	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/	• 非常低 • 低 • 中 • 高	• 低 • 中 • 高 • 非常高

		• 中 / 輕 微程度	非預期 時間尺度 • 短期 • 中期 • 長期 • 跨世代		
7	永續發展				
7.1	AI 可用於建立改進氣候變遷模型，既可用於排放預測，也可透過太陽預報等支援緩解措施。演算法可以在低碳能源上運行，同時支援對環境不利的環境，例如，基於 AI 的推薦系統可能會鼓勵資源的過度消耗。但是，另外一方面來看，研究表明，訓練大型語言模型的碳足跡相當於約 30 萬公斤的二氧化碳排放量。機器學習模型的設計、開發和使用通常也需要大量的電力和水資源。意識到此類影響並採取行動減輕影響非常重要。例如，系統使用的能源網對排放有顯著影響，因為不同地區由再生和不可再生能源的不同組合供電。因此，可以透過選擇碳排放量最小的能源網絡來部分減輕模型訓練的碳影響。				
7.2	程序評估				
7.2.1	AI 是否曾進行環境影響評估？				
7.2.2	是否使用負責任創新的問責指標（SDG、ESG）來預測 AI 如何促進環境繁榮（長期永續性），而不是僅僅避免直接的、可能的區域和短期損害？				
7.2.3	AI 在其整個生命週期中可能會以不同的方式損害環境和生態系統。其中一些問題可能更直接適用於嵌入式 AI（AI 演算法和模型在設備層級的應用）。請回答以下問題。				
7.2.3.1	研究/設計/開發階段				
7.2.3.1.1	是否估計過 AI 硬體製造過程中涉及的原材料提取、加工和運輸對環境的影響？				
7.2.3.1.2	您測量過系統的耗電量嗎？結果如何？				
7.2.3.2	使用階段				
7.2.3.2.1	根據 AI 的用途，它可能會鼓勵排放或資源密集型活動。例如，基於 AI 的推薦系統可能會增加資源的消耗。AI 驅動的自動駕駛汽車可能會阻止人們選擇更環保的公共交通選擇。是否具體考慮了您的 AI 所促進的用例對環境的影響				
7.2.3.3	使用結束/卸載/終止階段				
7.2.3.3.1	一旦系統退役，將如何處理廢棄 IT 硬體的拆卸、回收和/或處置過程？				
7.3	確認和減輕影響				
7.3.1	正面影響 下列問題並非詳盡。考慮 AI 的設計、開發、部署和使用可能產生的所有結果，無論是正面的還是負面的。				

	該系統對環境和生態系統的繁榮有哪些預期的正面影響？	評估預期正面結果的規模	評估預期正向結果的範圍	評估預期正面結果發生的可能性		
		顯著水準 • 非常高 • 高 • 中 • 中/次要	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非預期 時間尺度 • 短期 • 中期 • 長期 • 跨世代	發生的可能性 • 低 • 中 • 高 • 非常高		
7.3.2	負面影響					
	系統對環境和生態系統繁榮的預期負面/不利影響是什麼？	評估預期負面影響的規模	評估預期負面影響的範圍	評估預期負面影響的可補救性	發生的可能性	上述程序保障措施在多大程度上減輕了這種影響？需要實施哪些額外的緩解和補救策略來應對這種潛在危害？
		嚴重程度： • 災難性 • 嚴重 • 重大	影響程度 <u>受影響方</u> • 主要 • 次要	• 非常低 • 低 • 中 • 高	• 低 • 中 • 高 • 非常高	

		<ul style="list-style-type: none"> <li>• 中/輕微程度</li> </ul>	<ul style="list-style-type: none"> <li>• 意外/非預期</li> <li>• <u>時間尺度</u></li> <li>• 短期</li> <li>• 中期</li> <li>• 長期</li> <li>• 跨世代</li> </ul>			
8. 隱私及資料保護	應在國家或國際層面以多方利益相關者的方式建立適當的資料保護框架和治理機制，並受到司法系統的保護，並在 AI 的整個生命週期中得到保證。					
8.1	<p>AI 模型通常是透過累積大量的數據來訓練。當 AI 與大數據和物聯網結合時，這種做法會引起人們對用戶隱私和資料濫用的擔憂。大數據對 AI 的影響通常用三個 V 來表徵：數量（用於訓練的資料量）、多樣性（實現新的和意想不到的推論的組件）；和速度（促進即時分析和共享的組件）。我們與連網裝置的持續連結和互動意味著，如果隱私得不到適當保護，我們的運動、個性、習慣和品味的完整表現就可能被創造和利用。為敏感環境（例如醫療保健）訓練高效的 AI 模型需要大量的隱私敏感數據，而在使用 AI 時，始終存在從所謂的非敏感數據或匿名數據推斷出敏感數據的風險。發生這種情況是因為能夠在多個非敏感屬性之間進行連結並推斷出敏感資訊。從法律角度來看，與其他原則相反，隱私和資料保護領域在世界各地相對受到更多監管，儘管這種保護的力度和廣度各不相同。在國際層面，「世界人權宣言」第 12 條將對權利的保護納入一份具有約束力的文件中。類似的保護也可見於「歐洲人權公約」8 條。就其地位而言，雖然隱私是一項國際公認的人權，但資料保護卻不是一項受到嚴格監管的主題：例如，請參閱《歐洲通用資料保護條例》，該條例啟發了其他幾個國家。資料保護法通常包括有關所收集資料類型的規範、資料儲存時間長度以及同意規範等要求。除了法律要求之外，從技術角度來看，在隱私設計或隱私保護機器學習的保護下，還有不同的方法，試圖盡可能減少用戶隱私的風險、識別個人身份的可能性、或洩露個人信息。其中一種方法稱為差異隱私，它向現有資料添加噪音，以便在資料外洩時將個人與其可識別資訊分開。</p>					
8.2	程序評估					
8.2.1	資料保護					
8.2.1.1	AI 可以存取哪些類型的個人資料？					
8.2.1.2	數據和輸入是由人類、自動感測器還是兩者收集的？					

8.2.1.3	專家提供的資料和輸入是提供的、觀察到的、合成的還是衍生的？								
8.2.1.4	如果資料來自外部實體，是否有書面協議詳細說明資料共享的條件？								
8.2.1.5	資料儲存的安全等級是否與其敏感性相稱？								
8.2.1.6	是否應用了資料最小化原則？換句話說，是否對系統中包含每種資料類型的相關性和必要性進行了事前評估？								
8.2.1.7	如果資料是個人資料：								
8.2.1.7.1	不同類型的個人資料是否受到不同的處理標準（特別是敏感類型的資料）？								
8.2.1.7.2	資料是否經過匿名或假名處理？								
8.2.1.7.3	系統是否主動連結不同資料庫？								
8.2.1.7.4	人們是否主動同意 AI 處理他們的資料？								
8.2.2	隱私								
8.2.2.1	是否對 AI 進行了隱私影響評估？								
8.2.2.2	訓練資料的品質是否經過公平性和非歧視性的評估？								
8.2.2.3	系統中是否應用了隱私設計？請詳細說明如何？								
8.2.2.4	用戶能否請求刪除其資料並停止 AI 的處理？								
8.2.2.5	如果第三方可以存取數據，是否有相關規定來防止惡意行為？								
8.3	確認及減輕影響								
8.3.1	考慮 AI 的設計、開發、部署和使用可能產生的所有結果，無論是正面的還是負面的。								
	<table border="1"> <tr> <td>正面影響該系統對個人和團體隱私的預期正面成果是什麼？</td> <td>評估預期正面結果的規模。</td> <td>評估預期正向結果的範圍。</td> <td>評估預期正面結果發生的可能性</td> </tr> <tr> <td></td> <td>           顯著水準           <ul style="list-style-type: none"> <li>• 非常高</li> <li>• 高</li> <li>• 中</li> <li>• 中/次要</li> </ul> </td> <td>           影響程度  <u>受影響方</u> <ul style="list-style-type: none"> <li>• 主要</li> <li>• 次要</li> <li>• 意外/非預期</li> </ul> <u>時間尺度</u> <ul style="list-style-type: none"> <li>• 短期</li> <li>• 中期</li> <li>• 長期</li> <li>• 跨世代</li> </ul> </td> <td>           發生的可能性           <ul style="list-style-type: none"> <li>• 低</li> <li>• 中</li> <li>• 高</li> <li>• 非常高</li> </ul> </td> </tr> </table>	正面影響該系統對個人和團體隱私的預期正面成果是什麼？	評估預期正面結果的規模。	評估預期正向結果的範圍。	評估預期正面結果發生的可能性		顯著水準 <ul style="list-style-type: none"> <li>• 非常高</li> <li>• 高</li> <li>• 中</li> <li>• 中/次要</li> </ul>	影響程度 <u>受影響方</u> <ul style="list-style-type: none"> <li>• 主要</li> <li>• 次要</li> <li>• 意外/非預期</li> </ul> <u>時間尺度</u> <ul style="list-style-type: none"> <li>• 短期</li> <li>• 中期</li> <li>• 長期</li> <li>• 跨世代</li> </ul>	發生的可能性 <ul style="list-style-type: none"> <li>• 低</li> <li>• 中</li> <li>• 高</li> <li>• 非常高</li> </ul>
正面影響該系統對個人和團體隱私的預期正面成果是什麼？	評估預期正面結果的規模。	評估預期正向結果的範圍。	評估預期正面結果發生的可能性						
	顯著水準 <ul style="list-style-type: none"> <li>• 非常高</li> <li>• 高</li> <li>• 中</li> <li>• 中/次要</li> </ul>	影響程度 <u>受影響方</u> <ul style="list-style-type: none"> <li>• 主要</li> <li>• 次要</li> <li>• 意外/非預期</li> </ul> <u>時間尺度</u> <ul style="list-style-type: none"> <li>• 短期</li> <li>• 中期</li> <li>• 長期</li> <li>• 跨世代</li> </ul>	發生的可能性 <ul style="list-style-type: none"> <li>• 低</li> <li>• 中</li> <li>• 高</li> <li>• 非常高</li> </ul>						
9	員國應確保始終可以為 AI 生命週期的任何階段賦予道德和法律責任。它強調 AI 永遠無法取代人類的最終責任和義								



	務。人的監督不僅指個人的監督，也酌情包括包容性的公共監督					
9.1	人類監督對於支持和尊重人類自主至關重要，有助於解決基於流程的問題，包括透過分配自由裁量權和減少歧視性決策，幫助減少自動化/演算法決策的非人性化影響，確保透明度和可解釋性，以及基於結果的問題。					
9.2	程序評估					
9.2.1	模型是否透過與現場數據互動而發展和/或獲得能力？					
9.2.2	AI 是否 (a) 取代現有的電腦系統？ (b) 取代人類； (c) 增加新功能或補充現有功能？					
9.2.3	如果 AI 接管了以前由人類執行的任務，那麼知識轉移如何保存？先前執行任務的人員在 AI 的開發和訓練中的參與程度如何？					
9.2.4	AI 是否有權做出影響人們的決定？					
9.2.5	是否總是可以將 AI 生命週期任何階段的道德和法律責任歸咎於自然人或現有法律實體？					
9.2.6	是否存在讓人類實體推翻 AI 所做的決策的機制？					
9.2.7	是否存在過度依賴 AI 而導致人類自主權受到不利影響或損害的風險？					
9.3	確認及減輕影響					
9.3.1	正面影響 慮 AI 的設計、開發、部署和使用可能產生的所有結果，無論是正面的還是負面的。					
	該系統對人類監督的預期正面影響是什麼？	評估預期正面結果的規模。	評估預期正向結果的範圍。	評估預期正面結果發生的可能性		
		顯著水準 • 非常高 • 高 • 中 • 中/次要	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非預期 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	發生的可能性 • 低 • 中 • 高 • 非常高		
9.3.2	負面影響					
	該系統對人類監督的影響的	評估預期負面影響的	評估預期負面影響的	評估預期負面影響的	發生的可能性	上述程序保障措施在

	預期負面影響是什麼？	規模	範圍	可補救性		多大程度上減輕了這種影響？需要實施哪些額外的緩解和補救策略來應對這種潛在危害？
		嚴重程度： • 災難性 • 嚴重 • 重大 • 中/輕微程度	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	• 非常低 • 低 • 中 • 高	• 低 • 中 • 高 • 非常高	
10	透明度和可解釋性；問責制和責任					
10.1	<p>透明且可解釋的 AI 是負責任的，並具有責任歸屬機制，對於尊重、保護和促進人權、基本自由和道德原則至關重要。這包括制定適當的監督、影響評估、審計和盡職調查機制，包括保護舉報人，以確保問責制。</p> <p>重要的是要確保所有 AI，包括機器學習或機器人系統，無論自然人參與的程度如何，都受到精確的監管以及透明度和問責制要求。此類機制應以系統和環境相關監管的形式出現，能夠促進對 AI 做出的結果和決策背後的邏輯和原因的解釋，從而確保此類資訊易於獲取。解釋不僅對於防止結果和流程中的錯誤至關重要，而且對於建立最終使用者的信任也至關重要。應注意避免不透明的機制和系統，包括避免使用暗示 AI 代理的黑盒決策演算法和設計選擇系</p>					

	統意識、穩健的審計、演算法的可解釋性和透明度以及影響減輕程序（例如上訴和投訴）。
10.2	程序評估
10.2.1	系統意識
10.2.1.1	當使用者與 AI（而不是人類）互動時，他們是否充分意識到了這一點？
10.2.1.2	受 AI 影響的個人（直接或間接）是否充分意識到（影響他們的）決策是由 AI 或 AI 演算法通知或做出的？
10.2.1.3	是否已製定適當的解釋來幫助使用者和其他受影響的個人了解決策過程或系統在需要時如何運作
10.2.1.4	是否制定了適當的解釋，以幫助負責監管的政府機構了解決策過程或系統在需要時如何運作？
10.2.1.5	採用 AI 的決定是否已記錄並在線上傳達？
10.2.1.6	AI 是否可以做出任何負責該系統的自然人或法律實體缺乏專業知識或能力來批評、修改或推翻的決定？
10.2.2	審計
10.2.2.1	已經採取了哪些技術和制度設計，以確保 AI 的問責制、可審計性和可追溯性？
10.2.2.2	是否有指定的董事會、委員會或人員或類似機構來審查問責制和責任問題以及其他道德問題？
10.2.2.3	系統有審核流程嗎？
10.2.2.3.1	誰監督這個審計過程？
10.2.2.3.2	這是否涉及內部、外部或第三方查核人員？
10.2.2.3.3	是否進行相關檢查以確保審計人員不存在潛在或現有的利益衝突？
10.2.2.3.4	查核過程是否涵蓋整個專案生命週期？如果不是，該流程涵蓋哪些階段？
10.2.2.3.5	多久查核一次？
10.2.2.4	是否有審計追蹤記錄 AI 所做的所有決策？
10.2.2.5	審計追蹤中是否可以識別所有關鍵決策檢查點？
10.2.2.6	責任如何歸屬？
10.2.3	演算法可解釋性
10.2.3.1	該演算法（包括其內部工作邏輯）是否向公眾或任何監督機構開放？ AI 的程式碼是開源格式的吗？
10.2.3.2	公共機構可以索取代碼副本嗎？
10.2.3.3	用於訓練系統的資料集是否已知且可追蹤？
10.2.4	緩解評估
10.2.4.1	如果演算法造成不當行為，是否有關於責任分配的協議？
10.2.4.2	是否有指定的專案團隊成員或公共部門機構可以審查投訴、向受影響的個人提供解釋並在需要時糾正決定？
10.2.4.3	受 AI 影響的個人是否可以向該專案團隊成員提出索賠、投訴或要求解釋如何做出決定？

10.2.4.4	個人可以對 AI 做出的決定提出上訴嗎？				
10.2.4.5	是否有適當的監控機制？				
10.2.4.6	是否存在撤銷個人對系統的存取權限的機制（包括推翻決策的能力）？				
10.2.4.7	是否有適當的程序來調查公眾、研究人員或媒體提出的有關該系統的主張？				
10.2.4.8	對檢舉人保護有何規定？				
10.3	確認與減輕影響				
10.3.1	正面影響 考慮 AI 的設計、開發、部署和使用可能產生的所有結果，無論是正面的還是負面的。				
	該系統對透明度和可解釋性、問責制和責任有哪些預期的正面影響？	評估預期正面結果的規模。	評估預期正向結果的範圍。	評估預期正面結果發生的可能性	
		顯著水準 • 非常高 • 高 • 中 • 中/次要	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非預期 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	發生的可能性 • 低 • 中 • 高 • 非常高	
10.3.2	負面影響				
	與透明度、可解釋性、問責制和責任或缺乏相關的預期負面/不利影響是什麼？	評估預期負面影響的規模	評估預期負面影響的範圍	評估預期負面影響的可補救性	發生的可能性
					上述程序保障措施在多大程度上減輕了這種影響？需要實施哪些額外的緩解和補救策略

						來應對這種潛在危害？
	嚴重程度： • 災難性 • 嚴重 • 重大 • 中/輕微程度	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非預期 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	• 非常低 • 低 • 中 • 高	• 低 • 中 • 高 • 非常高		
11	<p>意識和素養</p> <p>為了 AI 的道德發展和部署，應培養公眾對 AI 技術和數據價值的認識和理解，並以它們對人權、基本自由以及環境和生態系統的影響為基礎。這可以透過開放和無障礙的教育、公民參與、數位技能和 AI 道德培訓、媒體和資訊素養以及政府、政府間組織、民間社會、學術界、媒體、社區領袖和私營部門聯合領導的培訓來促進。</p>					
11.1	<p>至關重要的是，每個人，包括最終用戶、民間社會、政策制定者和學生，而不僅僅是開發商和採購員，都接受 AI 教育，並最終具備 AI 素養。廣泛的 AI 素養將確保使用者能夠成為關鍵消費者：有效地使用 AI 作為工具，評估 AI 並讓開發商/採購商承擔責任。它還可以促進創新、支持和促進透明度和問責制、平息用戶毫無根據的恐懼並防止誤解。培養 AI 意識和素養與確保透明度、可解釋性和問責制密切相關。</p>					
11.2	程序評估					
11.2.1	是否已經公開宣布設計這個 AI 的意圖？					
11.2.2	可以在網路上找到有關係統、其功能、用途和功能的資訊嗎？如果沒有，是否有計劃在專案生命週期的特定階段發布此資訊？					
11.2.3	該系統將供公眾使用還是僅供內部使用？					
11.2.4	是否制定了任何計劃來幫助用戶和受影響的群體了解該系統及其部署背後的原因？					

11.3	確認與減輕影響					
11.3.1	正面影響					
該系統對 AI 意識和素養 有哪些預期的 正面 成果？	評估預期正 面結果的規 模。	評估預期正 向結果的範 圍。	評估預期正 面結果發生 的可能性			
	顯著水準 • 非常高 • 高 • 中 • 中/次要	影響程度 <u>受影響方</u> • 主要 • 次要 • 意外/非 預期 <u>時間尺度</u> • 短期 • 中期 • 長期 • 跨世代	發生的可能 性 • 低 • 中 • 高 • 非常高			
11.3.2	負面影響					
該系統 對AI意 識和素 養有哪 些預期 的負面 成果？	評估預 期負面 影響的 規模	評估預 期負面 影響的 範圍	評估預 期負面 影響的 可補救 性	發生的 可能性	上述程 序保障 措施在 多大程 度上減 輕了這 種影響？ 需要實 施哪些 額外的 緩解和 補救策 略來應 對這種 潛在危 害？	
	嚴重程 度： • 災難 性 • 嚴重 • 重大 • 中/輕	影響程 度 <u>受影響 方</u> • 主要 • 次要 • 意外	• 非常 低 • 低 • 中 • 高	• 低 • 中 • 高 • 非常 高		

		微程度	/非預期 時間尺度 • 短期 • 中期 • 長期 • 跨世代			
--	--	-----	---	--	--	--

## 歐洲審計院（ECA）部署 AI 的一般風險描述、影響及其因應對策

風險	描述及影響	可能性/影響性	因應對策
有限的透明度和對 AI 邏輯的監管	許多 AI 系統因其複雜和不透明的決策過程而被稱為“黑箱”。理解和解釋系統用來產生特定輸出的過程可能是困難的。這限制了系統的透明度，並使在審計證據的背景下評估輸出可靠性變得更加困難。	高/高	<ul style="list-style-type: none"> <li>● 選擇本質上更加開放、透明且可解釋的 AI 模型。</li> <li>● 確保模型提供達成結果的逐步過程描述和所使用的軟體代碼（以便允許重新執行）。</li> <li>● 為審計人員提供培訓和方法指導，幫助他們理解輸出，並記錄輸出在審計工作中的使用情況。</li> </ul>
責任缺口	在 AI 系統作出決策的過程中出現失誤時，可能無法明確識別應承擔責任的實體，因為法律和倫理框架尚未跟上發展。	中/高	<ul style="list-style-type: none"> <li>● 不要將任何具有實質意義的決策交給 AI 系統。</li> <li>● 部署 AI 系統來協助人類員工，由人員保留監督權和承擔責任。</li> </ul>
偏見及歧視的產出	AI 系統可能沿襲其訓練數據的偏見，從而導致有偏頗和/或歧視性的輸出。這可能會影響其輸出的可靠性。高	高/高	<ul style="list-style-type: none"> <li>● 為員工提供關於這些風險的培訓和方法指導。</li> <li>● 制定明確的規則和界限，以確保輸出是可用的。</li> <li>● 定期重新評估模型，並優先選擇那些訓練數據偏見較少的模型。</li> <li>● 在適用的情況下，使用事先知</li> </ul>



			識較少的模型，並用 ECA 的知識庫（如 AWARE 等）來進行訓練。
隱私及資料保護議題	AI 系統使用和生成大量數據，其中可能包括個人數據（包括敏感信息）。這些信息有可能未按照現行法規進行處理，或者沒有得到充分保護。此外，這些數據可能不夠準確。	高/高	<ul style="list-style-type: none"> <li>●對於本地部署的 AI 模型，確保模型及其支持的基礎設施和治理符合相關的數據保護法規。</li> <li>●對於基於雲端的商業服務，確保供應商符合 GDPR 規定並提供充分的合同保證。</li> <li>●儘可能優先選擇歐盟的模型和供應商。</li> <li>●出於工作目的，只允許通過企業許可證使用商業服務（具有更強的合同保證）。</li> <li>●應用從其他 IT 系統中獲得的經驗和最佳實踐。</li> </ul>
安全議題	AI 系統可能容易受到攻擊，包括病毒入侵。這可能導致惡意使用、操縱或數據洩漏，從而造成損害或聲譽損失。	高/高	<ul style="list-style-type: none"> <li>●在部署 AI 模型之前進行專門的安全評估。</li> <li>●在 IT 安全計畫中考慮針對 AI 模型特有威脅的額外減緩措施。</li> <li>●出於工作目的，只允許通過企業許可證使用商業服務（具有</li> </ul>

			更強的合同保證)。
依賴與自主權喪失	過度依賴 AI 可能導致人類專業知識和決策自主權的喪失。	中/中	<ul style="list-style-type: none"> <li>●將 AI 的應用範圍限制在特定的應用領域，以便技術能夠協助員工並提高效率。</li> <li>●投資於培訓，以保持員工的專業知識。</li> </ul>
環境影響	AI 系統需要大量計算，這導致高能耗並可能對環境造成影響。	高/低	表現最佳的商業服務供應商有動機開發更高效的模型，以便能夠擴展他們的業務。

## 歐洲審計院（ECA）部署 AI 的業務風險描述、影響及其因應對策

風險	描述及影響	可能性/影響性	因應對策
取得及維持成本太高	由於預算有限，獲取和維持最先進的 AI 成本可能會過於高昂。	高/高	<ul style="list-style-type: none"> <li>● 探索機構間合作的可能性，以共享基礎設施和授權成本。</li> <li>● 儘管並非業界最先進的技術，一些更便宜或資源消耗較少的 AI 模型可能仍足以滿足使用需求。</li> </ul>
法規限制使用 AI	遵守即將實施的 AI 法案及其他適用法規（包括資料保護）可能會非常複雜，並且可能會減緩採用進程。此外，法規環境的快速變化可能會中斷系統的部署。	中/中	<ul style="list-style-type: none"> <li>● 同時尋求不同的解決方案，並為每個使用案例選擇最合適/風險較小的方案。</li> <li>● 儘可能優先選擇開源模型和歐盟供應商。</li> <li>● 持續監測法規的發展，定期審查和更新內部政策。</li> </ul>
過度依賴單一 AI 供應商	供應商的競爭優勢可能使其擁有主導市場份額或建立壟斷，從而導致 ECA 過度依賴單一供應商。	中/中	<ul style="list-style-type: none"> <li>● 即使主要使用商業供應商，仍然考慮部署替代的開源模型以降低依賴性。</li> <li>● 如果性能相當，則優先選擇可以在本地運行或輕鬆運行於不同供應商雲基礎設施上的開源模型。</li> <li>● 有前景的歐盟供應商正在獲得重要資金，並進入市場。</li> </ul>

<p>內部技能不足以部署生產工具</p>	<p>內部可能缺乏足夠的技能來將初步試點或概念驗證轉化為所有員工可用的產品，從而減緩採用進程或增加成本。</p>	<p>中/高</p>	<ul style="list-style-type: none"> <li>●將在 ECA 建立 AI 能力中心。</li> <li>●機構間的合作，特別是與歐盟委員會的合作，可能有助於填補技能差距。</li> </ul>
<p>AI 模型的快速過時</p>	<p>在過去兩年中，AI 領域的發展速度非常快。由於準備、測試和運行試點需要時間，因此存在部署一個已經過時的模型進入生產環境的風險。</p>	<p>高/中</p>	<ul style="list-style-type: none"> <li>●在本地基礎設施及相關流程中保持足夠的靈活性。</li> <li>●儘管部署的模型不是表現最好的，但仍可能足以滿足我們的使用需求。</li> <li>●對於商業產品而言，此風險的影響較小，因為升級到更高性能的模型可能僅需購買新許可證。</li> </ul>
<p>使用 AI 造成聲譽損害</p>	<p>因誤用 AI 而導致的錯誤/問題可能會對 ECA 造成聲譽損害。</p>	<p>中/高</p>	<ul style="list-style-type: none"> <li>●為員工提供培訓和明確的指導，說明 AI 的可接受和正確使用方式。</li> <li>●為員工進行具體的檢查和培訓。</li> </ul>
<p>員工的負面觀感和反對聲</p>	<p>員工對 AI 的採用可能持負面看法，認為這是一種自動化人類任務的嘗試，因此對勞動力構成威脅。這可能導致反對聲音和缺乏採用。</p>	<p>低/中</p>	<ul style="list-style-type: none"> <li>●在所有層級進行適當的溝通（明確律定 AI 作為人類員工的輔助角色）。</li> <li>●提供培訓和意識提升課程。</li> </ul>
<p>在內部流程中誤用 AI</p>	<p>在內部流程中不當使用 AI 可能</p>	<p>低/高</p>	<ul style="list-style-type: none"> <li>●歐盟 AI 法案明確列出被禁止</li> </ul>

	會導致非法或有害的做法。		的有害做法。 ●現有的規則和保障措施同樣適用於新技術。
--	--------------	--	--------------------------------