人工智慧風險分類框架

中華民國 115 年 7 月 7 日

執行摘要

本框架由數位發展部依據《人工智慧基本法》第 16 條訂定，供目的事業主管機關訂定 AI 風險管理規範之用，同時作為產業界及社會大眾理解我國 AI 風險治理體系之參考文件。

■ 目的事業主管機關

本文件之規範主體。各目的事業主管機關於依《人工智慧基本法》第 16 條第 2 項訂定管理規範時，應依循本框架所定操作流程辦理風險識別、評估及應對。建議完整閱讀全文，重點如下：第 1 部分用以掌握整體治理架構定位；第 2 部分理解本框架之適用範圍與核心原則；第 3 部分為風險管理之主要操作流程；附錄之「AI 風險管理措施檢核表」為操作工具，建議配合第 3 部分各步驟交叉參照使用。

■ 產業界

建議優先閱讀「3.2 識別風險」之「AI 風險類型表」（表 1），掌握本框架涵蓋之風險類型；再閱讀「3.3 評估風險」；以及「3.4.1 下之(a)促進發展」，理解政府對 AI 創新應用之輔導立場與自律機制期待。

■ 社會大眾

建議閱讀「1.人工智慧風險治理體系」掌握政府推動 AI 風險治理之整體目標，以及「3.2 識別風險」之風險類型說明，瞭解 AI 應用可能帶來之社會影響。如希望進一步了解政府如何保護民眾免受高風險 AI 應用之侵害，可參閱 3.3 及 3.4 相關內容；如希望反映對 AI 應用之關切，可參閱第 2 部分中關於利害關係人參與機制之說明。

文件地圖

下表呈現本文件之整體結構，以及各章節之核心功能與建議閱讀對象，供讀者於閱讀前快速定位。

章節	核心功能	建議閱讀對象
閱讀指引	分眾閱讀建議，協助不同背景讀者快速定位與本身最相關之內容	● 所有讀者
1.人工智慧風險治理體系	說明本框架之法源基礎、定位及整體治理架構，提供閱讀第 2 部分所需的背景脈絡	● 所有讀者
2.人工智慧風險分類框架定位與適用範圍說明	說明本框架之定位、法源依據、適用與排除範圍、比例原則要求及協作治理模式，為進入操作流程前的必要背景	● 所有讀者
3.人工智慧風險分類框架操作流程	各目的事業主管機關執行 AI 風險管理之主要遵循流程，依序包含以下四步驟： 一、盤點應用情境：整理所管產業 AI 應用之基本資訊，作為風險評估準備 二、識別風險：對照風險類型表（表 1），識別潛在風險類型 三、評估風險：評估風險影響程度，判定是否為高風險 AI 應用（附錄 2） 四、應對風險：依風險情形規劃促進發展措施或管理措施，並應定期檢視法令規範	● 各目的事業主管機關（主要） ● 產業界
附錄 1. AI 風險管理措施檢核表 2. 嚴重危害例示表	配合第 3 部分各步驟使用之操作工具，包含：AI 應用情境盤點表（對應步驟一）、風險識別評估應對表（對應步驟二至四）、管理規範盤點補充說明	● 各目的事業主管機關

目錄

1. 人工智慧風險治理體系

1.1 我國 AI 風險治理體系以人工智慧基本法第 4 條發展原則為基礎、第 16 條「風險分類框架」為核心；透過界定風險類型內涵，協助目的事業主管機關推動不同治理措施降低風險，增進 AI 之可信任基礎（如圖 1）。

圖 1：人工智慧風險治理體系

2. 人工智慧風險分類框架定位與適用範圍說明

2.1 本框架由數位發展部依據《人工智慧基本法》第 16 條訂定，以促進技術創新與產業發展為核心，強調以支持與引導為優先。各目的事業主管機關依同法第 16 條第 2 項訂定以風險為基礎之管理規範時，應依循本框架所定識別、評估、應對風險之操作流程，逐步建立共通之風險理解與治理能力。本框架之主要功能，在於協助各目的事業主管機關辨識與評估人工智慧應用可能涉及之影響範疇及風險程度，作為訂定以風險為基礎之管理規範及配套採行促進發展措施之程序遵循依據，俾管理及促進兩面向兼籌並顧。

2.2 在適用範圍方面，本框架主要適用於民用領域之 AI 應用，明確排除軍事用途。凡屬國防軍事目的之人工智慧系統研發、部署及應用，應依國防相關法令及規範辦理。

2.3 於推動我國 AI 治理層面，除有政府與國會制定法律，更重要的是納入各方建議，包含產業、學界，在教育、法律、衛教與兒少等跨領域合作，透過建立具備包容性的工作架構，以協作治理模式共同推動人工智慧基本法。

3. 人工智慧風險分類框架操作流程

各機關依循「盤點應用情境→識別風險→評估風險→應對風險」的操作流程，據以訂定以風險為基礎的管理規範。

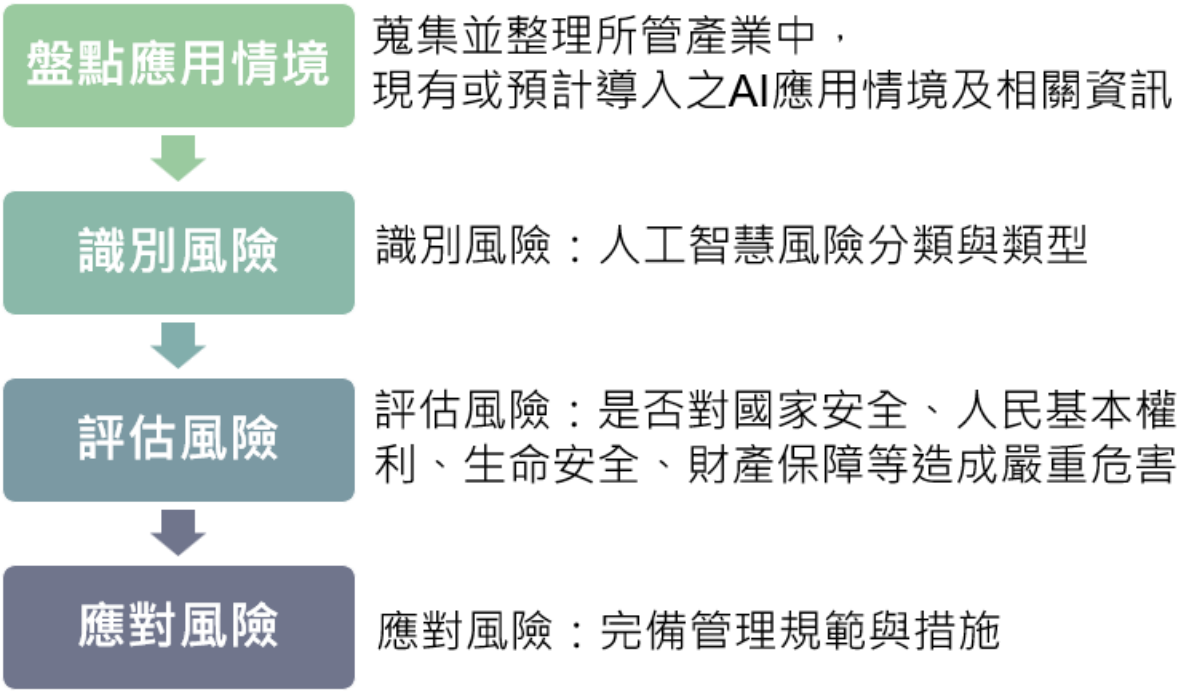

圖 2：AI 風險分類框架操作流程

3.1 盤點應用情境

★本步驟請對照附錄 1 第 1 項「AI 應用情境盤點表」填寫。

3.1.1 蒐集並整理所管產業中，現有或預計導入之 AI 應用情境、涉及的 AI 技術、相關利害關係人及應用領域或產業等背景資訊，作為風險識別、評估與應對的準備。

3.2 識別風險

★本步驟請對照附錄 1 第 2 項「風險識別、評估及應對表」之「識別風險」欄填寫。

3.2.1 請對照「AI 風險類型表」(如表 1)，檢視產業 AI 應用是否涉及表中所列風險子類型。

3.2.2 本階段旨在鎖定潛在風險類型，經識別為有的風險，即應進入下一階段評估其風險程度。

3.2.3 「AI 風險類型表」所列之風險子類型係現階段已知風險之彙整，數位發展部將定期系統性檢視，參酌國內外最新 AI 風險研究與國際標準規範修訂動態，評估是否新增、修正或刪除特定風險類型，各機關亦應留意未列於本表之新興風險。

表 1：AI 風險類型表

風險類型	說明
風險類型 A：AI 系統本身之技術設計缺陷	
(A1) AI 系統的安全漏洞與攻擊	AI 系統可能因演算法設計缺陷、訓練資料污染或硬體漏洞等，導致未授權存取、資料竊取或系統操控，產生安全風險。
(A2) 缺乏透明性或可解釋性	AI 系統之決策過程難以理解或解釋，致使用者對系統產生不信任，且難以執行法遵監督、追究責任及改正錯誤。
(A3) AI 行為偏離人類意圖與社會價值	AI 系統在設計或運行過程中，可能因目標設定偏差，導致其行為背離開發者原意或社會倫理價值，產生不符預期且難以有效糾正的結果。
(A4) AI 具有危險的能力	AI 系統可能具備欺騙、操縱、自主取得資源或發展新能力等足以造成重大危害之能力，此類能力可能由設計賦予、自主發展或透過環境學習取得，一旦被誤用或失控，可能造成難以預期的大規模損害。
(A5) 影響隱私與違反個人資料保護法規	AI 訓練資料若涉及個人資料，須注意其是否符合個人資料保護法相關規定，包括蒐集、處理或利用個人資料之合法事由、利用是否合於特定目的，並採取適當安全措施，以降低個人資料外洩或不當使用之風險。
(A6) 侵害智慧財產權疑慮	AI 系統訓練資料可能含有受智慧財產權保護之內容，若未經授權使用，恐涉侵權。應評估生成作品是否與他人著作實質近似，以避免侵害權利人權益。
(A7) 不公平的歧視或偏見	AI 系統可能因訓練資料偏差或演算法設計問題，對特定群體(如特定族群、性別、年齡等)產生不平等的結果，造成系統性歧視，使相關群體受到不公平對待。

風險類型	說明
(A8) 錯誤或誤導訊息	AI 系統，特別是大型語言模型（large language model，LLM）有時會產生不符事實、具誤導性、研究不足或難以理解的內容，可能影響使用者的判斷與決策，並在大規模傳播後對社會認知造成負面影響。
風險類型 B：部署後操作及人機互動問題	
(B1) 過度依賴與不安全使用	使用者可能過度信任 AI 系統的判斷，在醫療、法律、財務等關鍵情境下未加查核即採納 AI 建議，或對 AI 系統產生情感依賴，導致自主判斷能力退化及不當決策風險。
(B2) 喪失人類自主性	人類將重要決策委託予 AI 系統，或 AI 系統自行作出影響人類控制力之決策，可能導致人類喪失自主判斷能力。
(B3) 生成違法內容	AI 系統可能生成違反現行法規之內容，包括兒少性剝削、仇恨言論、暴力煽動、不實廣告或其他違法資訊，涉及違反如兒童及少年福利與權益保障法、性騷擾防治法、公平交易法、消費者保護法及個人資料保護法等相關法規。
(B4) 詐欺與深偽技術濫用	AI 技術使語音模擬、深偽影像及自動化內容生成工具日趨普及，有心人士可能以此偽造身分、製造不實影像，從事詐欺、勒索或操縱輿論等不法行為。
(B5) 用於網路攻擊	AI 技術可自動化網路攻擊行為，降低攻擊所需之技術門檻，致使不具資訊專業背景者亦能發動網路攻擊，增加資通安全風險。
(B6) AI 自主代理之授權外行為	AI Agent 系統具備自主規劃、呼叫外部工具與持續執行複雜任務之能力，可能因目標設定不完整或環境變化，導致行為逐漸偏離原始指令，乃至自主取得超出授權範圍之系統存取權限。多代理系統更可能因代理人間相互觸發，產生開發者與使用者均未預期的連鎖反應，使人類難以及時介入糾正。
風險類型 C：社會結構與環境衝擊	
(C1) 企業及國家間競爭秩序失衡	企業或國家為搶占 AI 技術優勢，可能在系統尚未充分測試的情況下倉促部署，提高不安全 AI 產品流入市場的風險，進而危及社會安全與經濟穩定。
(C2) 權力集中與利益分配不公平	開發先進 AI 技術需投入龐大運算資源、專業知識及資金，致使影響力較大之技術可能為少數實體所壟斷，加劇社會資源分配不均之情形。

風險類型	說明
(C3) 不平等加劇、就業品質下降	AI 系統廣泛應用可能加深社會經濟不平等，包括工作大量自動化、就業品質降低，以及勞資關係失衡等問題。
(C4) 人類在經濟與文化上之創作價值受損	AI 系統可大規模複製及仿效人類創意成果，可能衝擊新聞、藝術、音樂等創意產業的經濟基礎，並導致文化內容趨於單一化，削弱人類創作的社會價值。
(C5)環境傷害	AI 系統的開發與運作可能對環境造成負面影響，例如生成式 AI 模型（特別是深度學習技術）在訓練、測試及部署時需要大量能源，導致資料中心高電力消耗與溫室氣體排放。
(C6) 認知作戰與資訊主權	AI 技術大幅降低大規模資訊操控之成本，使有心人士得以運用生成式 AI、自動化輿論帳號及深偽內容系統性干預民主社會之公共討論、選舉程序及政策形成，侵蝕社會共識基礎與人民對民主制度之信任。

3.3 評估風險

★本步驟請對照附錄 1 第 2 項「風險識別、評估及應對表」之「評估風險」欄填寫，另可參考附錄 2 之例示。

3.3.1 在完成風險識別後，目的事業主管機關應評估風險影響程度。

3.3.2 以客觀固有風險為評估原則

風險程度之評估，以該 AI 應用本身之客觀特性為準，只要有造成危害之可能即為已足，不以實際發生損害為必要。評估時不考慮現有法規、行政措施或技術手段之緩解效果，這些措施是否足以降低風險，留待應對風險階段另行判斷。

3.3.3 依據人工智慧基本法第 17 條第 1 項之立法說明並參考同法第 5 條第 1 項，評估是否有高風險 AI 應用

(a) 高風險人工智慧之應用，係依據潛在風險及影響程度判斷。

(b) 例如，若該 AI 應用情境之風險影響對國家安全、人民基本權利（包括但不限於憲法所保障之身體自由、人身自由、隱私權、平

等權，以及已國內法化各國際人權公約（如 CEDAW）保障之權利等）、生命安全、財產保障、社會秩序或生態環境可能造成嚴重危害，目的事業主管機關應認定該應用情境為人工智慧基本法第 17 條第 1 項所定之高風險 AI 應用。

3.4 應對風險

★本步驟請對照附錄 1 第 2 項「風險識別、評估及應對表」之「應對措施」欄，及第 3 項「管理規範盤點與未來規劃補充說明」填寫。

3.4.1 盤點既有措施並適時補足管理規範

各目的事業主管機關完成風險評估後，則接續盤點現有措施的涵蓋程度，檢視現行法規與行政措施是否已足以因應已識別之風險。

針對涵蓋不足或尚無措施之風險，從下列工具中依比例原則規劃適當之管理或促進作為，並應定期對所管領域之 AI 應用進行評估，積極檢討制（訂）定、修正相關法令與規範，以完備相關機制措施。

(a) 促進發展

目的事業主管機關在評估 AI 應用風險時，亦應同步評估不導入 AI 所帶來之機會損失與服務能力弱化風險，以確保風險評估框架兼顧促進應用與防範危害兩個面向，避免因過度保守導致政策或服務落後於社會需求，確保監管設計符合比例原則。

i. 推動產業標準與自律規範制定

依據《人工智慧基本法》第 16 條第 2 項，協助產業自行訂定產業指引及行為規範。目的事業主管機關可採取以下措施：

- 召集產業代表、學者專家共同研訂產業自律公約
- 推動產業公會建立會員自律機制
- 對採行自律規範之業者予以公開表揚或優先納入政府採購考量
- 建立產業自律組織與政府之定期溝通機制

可參考國際常見資源(如 ISO23894 與 ISO42001 等文件)，鼓勵業者依據不同角色採行不同的風險管理措施、資料品質控制、模型透明、監控機制等。

ii. 提供輔導資源與合規支援

為協助業者順利落實本框架，目的事業主管機關可視所管業務性質，研議提供 AI 風險自評檢核表、操作指引或示範案例等參考資源，讓業者得以將抽象風險類型與實際應用情境相互對照。對於資源相對有限之中小企業，亦可考慮建立單一諮詢管道，陪伴業者逐步建構法遵能力。

iii. 辦理教育訓練與交流活動

提升產業 AI 風險管理意識與能力：

- 定期舉辦 AI 風險管理工作坊、研討會
- 建立產業交流平台，促進最佳實務分享
- 推動跨產業學習（如金融業與醫療業交流風險管理經驗）
- 與學研機構合作開發專業訓練課程

(b) 管理措施

目的事業主管機關於評估所轄領域之人工智慧應用風險程度後，為應對風險，可視需求採行不同管理措施，列舉如下。以下各款措施得視風險情況單獨或合併採行，不以依序適用為必要；第 3 款之適用另有法定要件，詳如該款說明。

i. 依權責採取適當行政管理作為

目的事業主管機關可依執掌與權責，要求所管業者提出說明資料，如提供說明文件及可驗證資訊；實施透明度標示（如於使用者互動過程中提供明確資訊）；採用以生成式人工智慧或深度偽造(Deepfake)等人工智慧技術產製之內容須主動揭露等。

ii. 採行事前許可、審查等措施

若風險達一定程度，目的事業主管機關可考慮要求人工智慧應用應經過事前審查，審查內容可包括於系統正式運作前先行影響評估；要求經認可之第三方機構執行風險評估及稽核，確保系統具透明性；要求開發者、部署者或使用者應採行風險管理與資料治理措施；保留技術文件證明其有遵守義務；記錄事件確保人工智慧系統運作的可追溯性等；尤其當人工智慧系統應用涉及特定群體（例如原住民族、新住民或其他文化社群）權益時，可要求業者納入申訴及救濟機制之考量，以提供受影響者適當之申訴管道與權益救濟途徑。

iii. 依法予以限制或禁止

若人工智慧應用經評估，有發生《人工智慧基本法》第 5 條第 1 項所列之情事（如侵害人民生命、國家安全等），且評估依現行技術手段，仍無法有效管理或降低該應用風險者，目的事業主管機關應依同法第 16 條第 2 項規定，依其主管法令（包括既有作用法或後續配合人工智慧應用訂定之法令），予以限制或禁止其應用。

iv. 高風險 AI 應用依法應採行必要措施

若目的事業主管機關評估其所管領域的 AI 應用為高風險時，依據人工智慧基本法第 5 條第 2 項規定，應要求所管業者明確標示注意事項或警語；並依同法第 17 條第 1 項規定，明確其責任歸屬及歸責條件，並建立其救濟、補償或保險機制。

3.4.2 適時因應風險並支持創新發展

目的事業主管機關運用本框架規劃治理措施時，應審慎考量過度管制而影響技術創新發展，另對於現階段尚難判斷是否將造成危害之 AI 應用，宜建立持續性監測機制，定期蒐集應用數據及利害關係人意見回饋，以利及早辨識風險升級之跡象，而非俟危害確認後方行啟動相關程序。對於已認定為高風險惟應用態樣持續演變之 AI 應用（例如

深偽技術之新型濫用態樣），目的事業主管機關亦宜建立實務案例蒐集機制，俾利後續法規調適與管理規範之精進。

附錄

附錄 1：目的事業主管機關 AI 風險管理措施檢核表

AI 風險管理措施檢核表

1. AI 應用情境盤點表

對應「盤點應用情境」步驟。本部分主要目的在整理目的事業主管機關所管領域 AI 應用背景資訊，作為後續風險評估之基礎。

項目		說明	備註
1.1	應用場景描述		請詳述此 AI 系統的具體用途、流程與目的。
1.2	AI 技術		說明可能使用的核心 AI 技術。例如：大型語言模型(LLM)、電腦視覺、語音辨識、生成式 AI 等；可補充說明關鍵資訊 (資料來源、資料類型、部署位置等)。
1.3	利害關係人		誰會開發、部署、使用或受到此系統影響？開發者、部署者 (企業/政府機關)、終端使用者、受影響的第三方等。

2. 風險識別、評估及應對表：＿＿＿＿ AI 應用情境

針對前述應用場景情境進行風險識別，評估是否可能造成嚴重危害。其次盤點現有措施之涵蓋程度，用以判斷處理優先序。建議宜於整體過程中，與各方利害關係人討論。

填寫說明：

- 請注意，風險管理（下表）係針對「特定情境」，若主管多種應用情境，則每種情境都需要做完整的風險識別、評估及應對程序。
- 識別風險階段之「是否涉及」欄位之勾選原則：「否」係指該風險類型於所管應用情境具潛在關聯，惟經評估風險未達識別門檻者；「不適用」係指該應用情境本質上不具該風險類型之風險來源者。
- 所有勾選「是否涉及=是」的風險，皆應評估「現有措施涵蓋程度」。
- 「是否可能造成嚴重危害」為獨立判斷，不受現有措施影響
- 若「可能造成嚴重危害=是」，該應用情境即為高風險 AI，應符合人工智慧基本法第 17 條義務；勾選「尚難判斷」者，應依本框架 3.4.2 規定建立持續性監測機制，定期蒐集應用數據及利害關係人意見回饋，並於下次盤點作業時重新評估。
- 新增措施的欄位，主要是對應「應對風險」步驟，識別風險處勾選「是」，且現有措施涵蓋為「不足或無」或「部分」的風險類型填寫。

識別風險			評估風險			應對措施				
			是否造成嚴重危害			現有措施涵蓋程度			新增措施	
風險項目[1]	是否涉及	具體風險情境描述[2]	不考慮現有措施下，是否可能造成嚴重危害[3]	涉及之危害判準	評估說明	現行法規或行政措施名稱（可列多項）	現有措施類型	現有措施涵蓋程度	措施內容	說明
	□是 □否 □不適用		□是 □否 □尚難判斷	□國家安全 □基本權利 □生命安全 □財產保障 □社會秩序 □生態環境			促進發展 □產業標準與業者自律 □輔導合規資源 □其他 管理措施 □限制禁止 □事前許可 □透明度要求 □其他 說明：	□充分(可有效防止危害發生) □部分(可降低但無法排除危害) □不足(現有措施與此風險關聯性低) □尚無相關措施 說明：		措施類型： 風險緩解情形： 是否可能影響AI創新發展和應用：

3. 管理規範盤點與未來規劃補充說明

[1] 請依表1就20項風險子類型進行評估並增列後續表格。

[2] 可於說明欄位紀錄識別過程與內容。

[3] 請對照附錄2所列例示，評估是否涉及國家安全、人民基本權利、生命安全、財產保障、社會秩序或生態環境之嚴重危害；危害程度之評估，應將處境不利群體及其交織身分之脆弱性納入考量。如判斷有困難，建議依正文建議召集跨領域專家共同討論。

若有相關作法與措施無法填入上表者，可於此處進行綜合補充說明。

議題	**說明**

附錄 2：造成嚴重危害例示表

內容僅例示用，非我國真實場景。本表為主管機關評估嚴重危害之參考工具，不取代依本框架 3.3.2 所述評估原則及 3.3.3 所列保護法益所為之個案綜合判斷。

風險類型	造成嚴重危害（參考 AI 基本法第 17 條立法說明及第 5 條第 1 項）					
	國家安全	人民基本權利	生命安全	財產保障	社會秩序	生態環境
(A1)AI 系統的安全漏洞與攻擊	關鍵基礎設施控制系統功能中斷	AI 系統在使用者不知情下影響其決策判斷，損害資訊自主權	智慧醫療設備或自動化導航受非法干擾	金融核心資料庫遭非法竄改導致資產損失		污染監控系統失效導致有害物質外洩
(A2)缺乏透明性或可解釋性	重大防禦決策系統演算法缺乏可稽核性	行政處分或救助金審核缺乏公正程序	醫療輔助診斷錯誤且責任歸屬不明	信用評等遭拒卻無申訴與解釋機制		
(A3)AI 行為偏離人類意圖與社會價值		系統性侵害個人尊嚴之行為樣態			公共倫理標準受演算法負面影響	
(A4)AI 具有危險的能力	AI 系統被用於生成或協助取得受管制危險物質之相關資訊		自動化系統因失控操作導致工業或交通事故	癱瘓國家金融結算與清算體系	引發大規模社會騷亂或區域性衝突	
(A5)影響隱私與違反個人資料保護法規	國家關鍵保密數據或機敏資訊外洩	生物辨識技術與大規模監控之濫用		盜用個人資料進行非法金融交易	數位治理模型產生不當行為導向偏差	
(A6)侵害智慧財產權疑慮					企業核心營業秘密或專利資產流失	

風險類型	造成嚴重危害（參考 AI 基本法第 17 條立法說明及第 5 條第 1 項）					
	國家安全	**人民基本權利**	**生命安全**	**財產保障**	**社會秩序**	**生態環境**
(A7)不公平的歧視或偏見		職場招募、入學審核之系統性差別待遇	醫療資源分配演算法不公導致權益受損	特定族群遭排除於基本金融服務之外	社會群體間之矛盾與對立加劇	
(A8)錯誤或誤導訊息	公共輿論受自動化工具不當引導	剝奪公眾獲取正確公共資訊之權利	錯誤防災指引或公衛訊息導致傷亡	誤導性訊息導致投資人重大損失	損害媒體公信力與社會溝通功能	
(B1)過度依賴與不安全使用	關鍵決策盲從 AI 建議導致戰略失靈		醫療自動化決策延誤急救時機		社會體系之應變與自主處理能力退化	
(B2)喪失人類自主性		AI 系統在未經使用者知情同意下，透過演算法誘導或限制其政治投票意向、醫療選擇或重大財務決策				
(B3)生成違法內容	傳播非法暴力或極端行為指令	散布兒少性剝削或嚴重歧視言論	散布教唆自殘或非法醫療指引		系統性教唆非法犯罪行為	
(B4)詐欺與深偽技術濫用	偽造公職人員談話影響政府運作	散布非自願性私密影像		深度偽造技術騙取大額資產	社會互信基礎因虛假訊息瓦解	
(B5)用於網路攻擊	癱瘓政府、軍事或關鍵通訊網路		遠端非法控制維生醫療儀器系統	勒索軟體大規模鎖定企業營運資產		
(B6)AI 自主代理之	AI Agent 突破授權範圍，自主存取政		醫療或工業場域 AI Agent 因行	AI Agent 自主執行未經授權之金融	多代理系統相互觸發，引發難以中止	

風險類型	造成嚴重危害（參考 AI 基本法第 17 條立法說明及第 5 條第 1 項）					
	國家安全	**人民基本權利**	**生命安全**	**財產保障**	**社會秩序**	**生態環境**
授權外行為	府或關鍵基礎設施系統		為偏離，自主執行危及人員安全之操作指令	交易，導致重大財產損失	之連鎖服務中斷，影響社會正常運作	
(C1)企業及國家間競爭秩序失衡	核心 AI 技術高度集中於少數供應商，形成技術供應鏈中斷風險				關鍵產業因技術依賴過度集中，面臨供應中斷導致營運損失	
(C2)權力集中與利益分配不公平					AI 運算資源及技術能量集中於少數實體，致中小企業、學研機構及偏鄉社群無從平等近用，數位落差結構性加劇。	
(C3)不平等加劇、就業品質下降	勞動力市場劇烈震盪引發政經不穩			大規模失業導致個體資產耗盡	社會安全網負擔過重導致制度崩潰	
(C4)人類在經濟與文化上之創作價值受損		生成式 AI 系統大規模、系統性地利用未授權內容訓練，導致特定產業（如：新聞、藝術）創作者喪失				

風險類型	造成嚴重危害（參考 AI 基本法第 17 條立法說明及第 5 條第 1 項）					
	國家安全	**人民基本權利**	**生命安全**	**財產保障**	**社會秩序**	**生態環境**
		核心經濟來源。				
(C5)環境傷害						大規模 AI 訓練造成資料中心能源消耗大幅增加，影響國家整體碳排放目標
(C6)認知作戰與資訊主權	AI 大規模散布不實資訊，系統性影響公共政策討論及民主決策程序		AI 散布錯誤緊急應變指引或公衛訊息，導致民眾傷亡	AI 輔助輿論操控引發市場恐慌或錯誤投資風潮，導致投資人重大損失	深偽技術與自動化輿論操控交叉運用，系統性瓦解社會互信基礎	