



新北市政府
New Taipei City Government

新北市政府使用人 智慧作業指引說明會



Agenda



- ▶ 前言(第壹章)
- ▶ AI應用與技術(第貳章)
- ▶ AI導入評估及管理、規範及內控機制
(第參章及第肆章)
- ▶ 本作業指引之修訂(第伍章)
- ▶ 附件
- ▶ 機關配合事項
- ▶ QA

01

前言(第壹章)

前言-目的

目前面臨的問題

本指引預期的目標

各機關在導入AI技術時常自行摸索，導入方式不一

提升AI系統導入的一致性與規範性，強化AI治理能力

民眾無法得知本府資訊是否由AI生成，可能產生認知錯誤或誤信虛假資訊，導致信任流失。

打造一個民眾可信賴的AI應用環境，落實政府資訊揭露與責信原則

AI系統一旦上線，後續缺乏版本控管、模型偏誤監控或知識更新機制，導致預測失準或違反政策初衷。

確保AI應用具備長期穩定性、準確性、公正性與安全性

前言-訂定依據

① 國家科學及技術委員會
「行政院及所屬機關（構）使用生成式AI 參考指引」

② 數位發展部
「公部門人工智慧應用參考手冊」



前言-適用對象

- ① 本府所屬各機關及學校，辦理AI導入作業應遵循本作業指引。各機關並得視使用人工智慧之業務需求，參酌本作業指引另訂該機關之使用規範或內控管理措施。
- ② 本府及所屬各機關就所辦理之AI相關採購事項，亦應要求得標之事業、法人、團體或個人遵循本作業指引，並恪遵該機關所制定之使用規範或內控管理措施。

舉例來說

使用AI執行業務所需(簡報製作、公文、文稿撰寫等)，使用到chatGPT、Copilot等工具，亦須遵守本指引。

舉例來說

以採購案方式請廠商建置資訊系統其中部分功能有AI相關應用(智能客服)

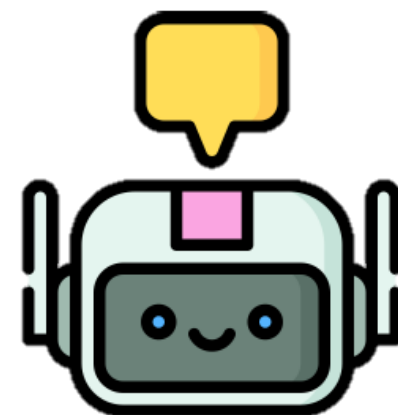
01

AI應用與技術 (第貳章)



什麼是AI?

係指模擬、延伸或強化人類智能之方法，
具備學習、自主推理、決策預測或模式識別等能力。



AI相關技術

凡應用下列任一技術，且系統核心功能實質依賴AI 模型或演算法進行預測、判斷、分類或決策者，即得認定為人工智慧應用範疇。

機器學習

分類、回歸、分群等演算法

深度學習

類神經網路、卷積神經網路、循環神經網路

自然語言處理

涵蓋語意理解、對話生成、語音辨識、文字生成與機器翻譯、文本摘要與資訊擷取等應用。

電腦視覺

物件偵測、人臉辨識、影像分割與光學字元辨識、姿態估計及視覺問答

生成式人工智慧

文本、圖像、語音等多模態資料之自動生成技術，包含大型語言模型（LLM）與其他生成模型。

AI模型管理與應用技術

自動化建模、模型微調、遷移學習、模型部署、模型漂移監控、機器學習 / 人工智慧營運管理平台、檢索增強生成、模型可解釋性工具、A/B測試、漸進式部署等技術。

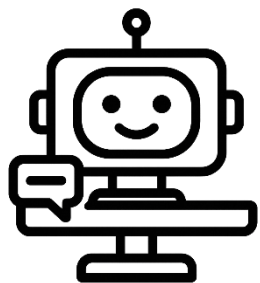
AI系統整合技術

API管理與微服務架構、容器化部署技術、AI模型版本控制等系統整合相關技術

代理式人工智慧

多代理系統、強化學習代理、目標導向對話系統及自主決策框架。

AI應用場景



智慧客服



公文起草與
文字生成



民眾意見
摘要與分析



教育訓練
教材生成



語音與影像
內容製作



多語言翻譯
與在地化



流程自動化
與文件分類



政策模擬
與公眾溝通



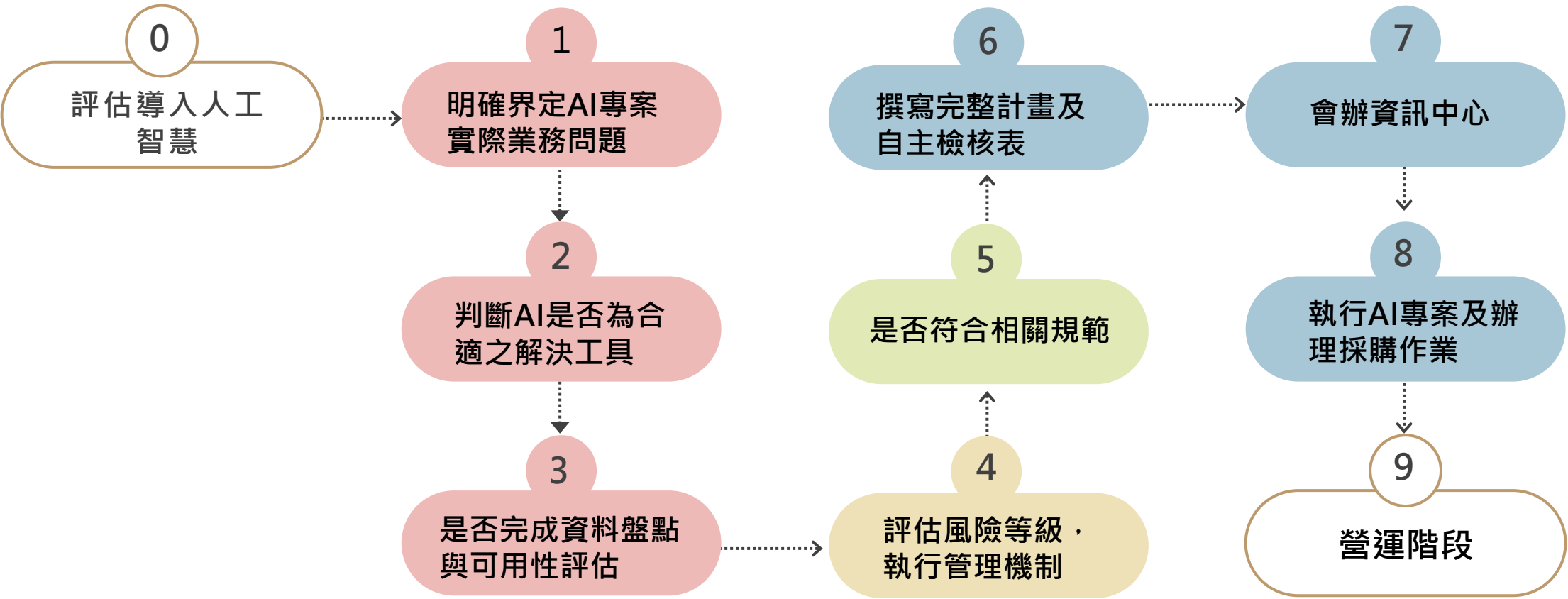
資料視覺化
敘事生成

more...

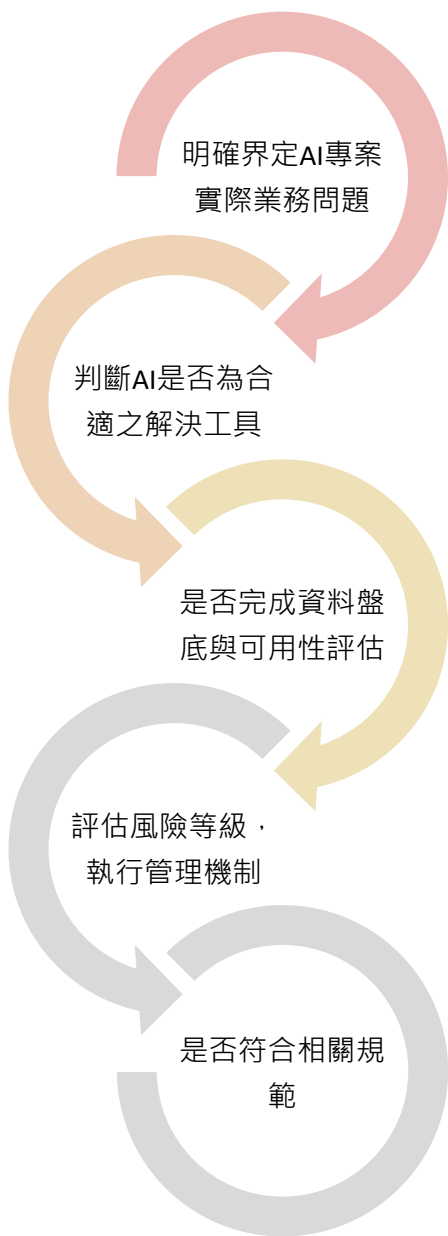
01

AI導入評估及管理、規範及內控機制(第參章及第肆章)

附件二、新北市政府導入AI內部控制流程圖



導入前期評估



導入流程可依循《公部門人工智慧應用參考手冊》，進行前期之服務與資料評估

不要為了AI而AI

- 1 事先應針對實際業務場景進行分析，明確界定待解決之問題
- 2 評估AI技術是否為適切的解決工具
- 3 確認資料之可取得性、品質、完整性及合規性

風險評估

參照歐盟人工智慧法案，依據AI對人類權益、社會福祉、資料安全與基本權利的潛在影響，將AI系統劃分為四大風險等級，並對應不同層級的管理強度

風險等級	說明	管理措施
不可接受風險	涉及對人類尊嚴、自由與民主制度構成嚴重威脅之AI用途	不得開發
高風險	AI系統若直接影響個人安全或基本權利且為社會關鍵領域運作（如醫療、交通、教育、執法），即被歸類為高風險	應送本府資安實驗室或AIEC驗測後始得實施。
有限風險	雖不會直接造成個人基本權利受損，但仍涉及與人互動時的資訊不對稱風險	應明確告知使用者其所接觸內容為人工智慧產出
最低風險	AI用途對個人或社會幾乎不構成風險	

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

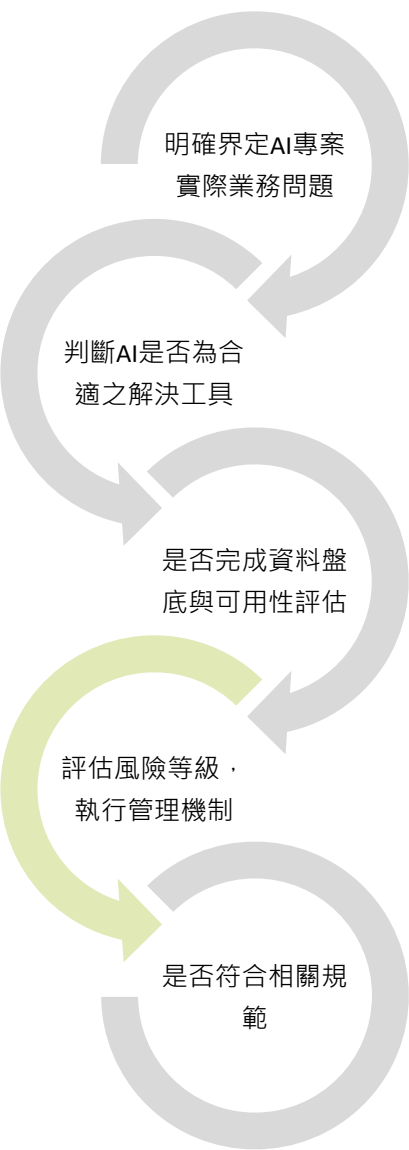
是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

風險評估-管理機制(AIEC驗測)

高風險專案應送本府資安實驗室或AIEC驗測後始得實施。



風險評估-管理機制(AIEC驗測)

數位發展部數位產業署 函

地址：100台北市中正區忠孝西路一段66號
20樓

聯絡人：王崇宇

電話：02-23808344

電子郵件：toast1989@adi.gov.tw

受文者：新北市政府

發文日期：中華民國114年7月28日

發文字號：產經字第1140006792號

速別：普通件

密等及解密條件或保密期限：

附件：

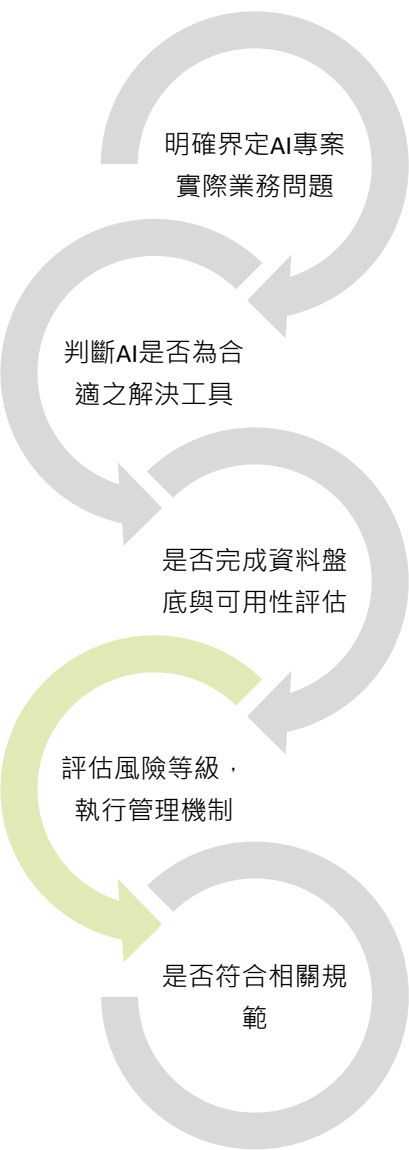
主旨：有關貴府所詢AI產品與系統評測制度相關文件一案，詳如說明，復請查照。

說明：

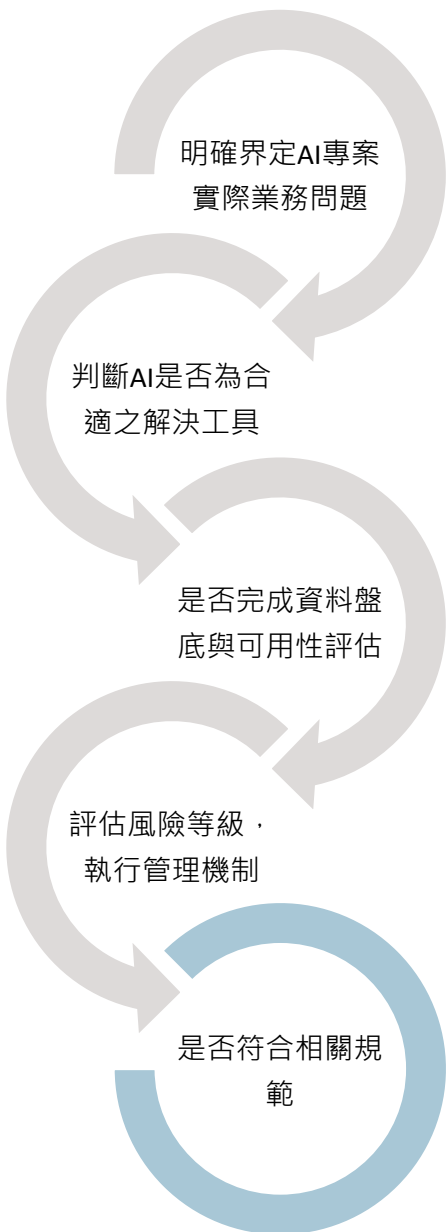
- 一、復貴府114年7月11日新北府研資字第1143473961號函。
- 二、有關「AI產品與系統評測中心」之相關資訊，已登載於評測中心官方網站 (<https://www.aiec.org.tw/>)，敬請參閱網站內容，倘貴府有產品測試需求，建議可送交「AI產品與系統評測中心」進行評測，作為後續應用之參考。

正本：新北市政府

副本：



相關規範



需遵守項目

1. 第4章第1條、第2條、第3條、第4條、第7條1款至第7條3款及第8條
2. 須遵守項目已納入附件三AI專案建置自主檢核表

建議項目

1. 第4章第5條、第6條及第7條第4款

相關規範

需遵守項目-第4章第1條

一、機關應審慎面對 AI 產出之資訊，由機關之業務單位依風險進行客觀且專業之最終判斷，不得取代其自主思維、創造力及人際互動，亦不得完全信任 AI 產出，或逕以未經確認內容作為行政行為或公務決策之唯一依據。

依據行政院及所屬機關（構）使用生成式AI參考指引第2點。

建議項目-第4章第5條

五、如 AI 系統涉及輔助決策功能，建議機關可制定系統標準流程，例如：「AI 輔助決策三階段驗證機制」：

- （一）第一階段：AI 系統產出結果。
- （二）第二階段：業務人員專業審核。
- （三）第三階段：主管覆核確認。

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

相關規範

需遵守項目-第4章第2條及第3條

- 二、使用 AI 應遵守資通安全、個人資料保護、著作權及相關資訊使用等規定，並注意其侵害智慧財產權與人格權之可能性。
- 三、各機關在採購或使用 AI 資通訊產品時，請確依行政院公布之「各機關對危害國家資通安全產品限制使用原則」辦理。應不允許大陸地區廠商及陸籍人士參與，並不得採購及使用大陸廠牌資通訊產品(含軟、硬體及服務)。

- 1.依據行政院及所屬機關（構）使用生成式AI參考指引第7點。
- 2.依據行政院秘書長109年12月18日院臺護長字第1090201804A號函規定，禁止使用及採購大陸廠牌資通訊產品（含軟體、硬體及服務）。

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

相關規範

需遵守項目-第4章第4條

四、鑒於近年 AI 技術日益普及，各機關以 AI 為名進行對外宣傳之情形日增。為維護政策推動之專業性與公信力，並避免對外產生誤導或認知落差，於宣傳前，應審慎確認所宣傳專案是否應用前述第貳章第二條所列之 AI 相關技術，以確保宣傳內容之準確性與專業性。

議員提案

新北市議會第 4 屆第 5 次定期會提案單

審查會別：第 1 審查會

案號：第 69 案

議案案別：議員提案

類別：研考

提案人：彭佳芸 張嘉玲

連署人：廖宜琨 鄭宇恩

案由：建請市府擬定各局處活動冠名 AI 之要件

說明：本市多數局處政策與活動皆冠上 AI 名稱，惟本府尚無明確 AI 定義與標準，眾多不具備生成式 AI 內涵之活動也冠上 AI 字詞，導致該活動實際應用內容混淆。建請市府擬定 AI 冠名指引，明確規範應具備之技術條件、資料來源與運算方式，避免誤用 AI 概念。

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

相關規範

建議項目-第4章第6條

六、各機關得提供生成式 AI 使用者基礎教育訓練課程，內容包含：

- (一) Prompt 撰寫技巧與最佳實務。
- (二) 生成內容的驗證與事實查核。
- (三) 常見風險(如內容幻覺、偏誤、錯誤推論)之辨識方法。
- (四) 運用封閉式系統保障機密資訊的技術建議。

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

相關規範

需遵守項目-第4章第7條1款及2款

七、運用生成式 AI 時，應遵循之規範：↵

- (一) 製作機密文書應由機關之業務單位親自撰寫，禁止使用生成式 AI。機密文書，係指「新北市政府文書處理要點」所定之國家機密文書及一般公務機密文書。↵
- (二) 機關之業務單位不得向生成式 AI 提供涉及公務應保密、個人及未經機關（構）同意公開之資訊，亦不得向生成式 AI 詢問可能涉及機密業務或個人資料之問題，但封閉式地端部署之生成式 AI 模型，於確認系統環境安全性後，得依「新北市政府文書處理要點」文書或資訊機密等級分級使用。↵

1. 依據行政院及所屬機關(構)使用生成式AI參考指引第3點及第4點
2. **針對運用生成式AI的相關規範**

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

相關規範

需遵守項目-第4章第7條3款

(三) 各機關使用生成式 AI 執行業務或提供服務輔助工具時，應讓使用者明確知道自己在使用生成式 AI 服務，如透過介面提醒或顯示數位浮水印等方式呈現，並避免 AI 所生成的內容帶有偏見歧視。←

議員提案

新北市議會第 4 屆第 5 次定期會提案單

審查會別：第 1 審查會

案號：第 67 案

議案案別：議員提案

類別：研考

提案人：彭佳芸 張嘉玲

連署人：廖宜琨 鄭宇恩

案由：建請市府運用生成式 AI 推動政策時需註明使用範圍及軟體

說明：本市政策文件或簡報如使用生成式 AI 撰寫，未註明來源或應用範圍將造成民眾對資訊可信度存疑。建請市府於使用生成式 AI 製作之內容中，應標註所使用之軟體名稱與生成內範圍內容，強化資訊透明與公信力

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

相關規範

明確界定AI專案
實際業務問題

判斷AI是否為合
適之解決工具

是否完成資料盤
底與可用性評估

評估風險等級，
執行管理機制

是否符合相關規
範

建議項目-第4章第7條第4款

(四) 如機關使用檢索增強生成 (RAG) 技術，建議針對結合內部資料庫的生成式 AI 系統制定：

1. 知識庫資料品質標準。
2. 檢索相關度閾值設定。
3. 資料來源可追溯性要求。
4. 知識更新機制。

內控機制



預算編列

本府AI專案因涉及智慧城市及資訊相關計畫，預算編列應依「新北市政府資訊計畫先期審查作業要點」辦理

採購前

且於各機關採購前應會辦資訊中心並檢附附件三、AI專案建置自主檢核表

建置完成後

並應提交AI系統架構設計書，包含資料流程圖、模型部署架構、API介面設計等技術文件。

營運管理

應每年檢視，以確保AI系統之穩定性與信賴度。



風險管理

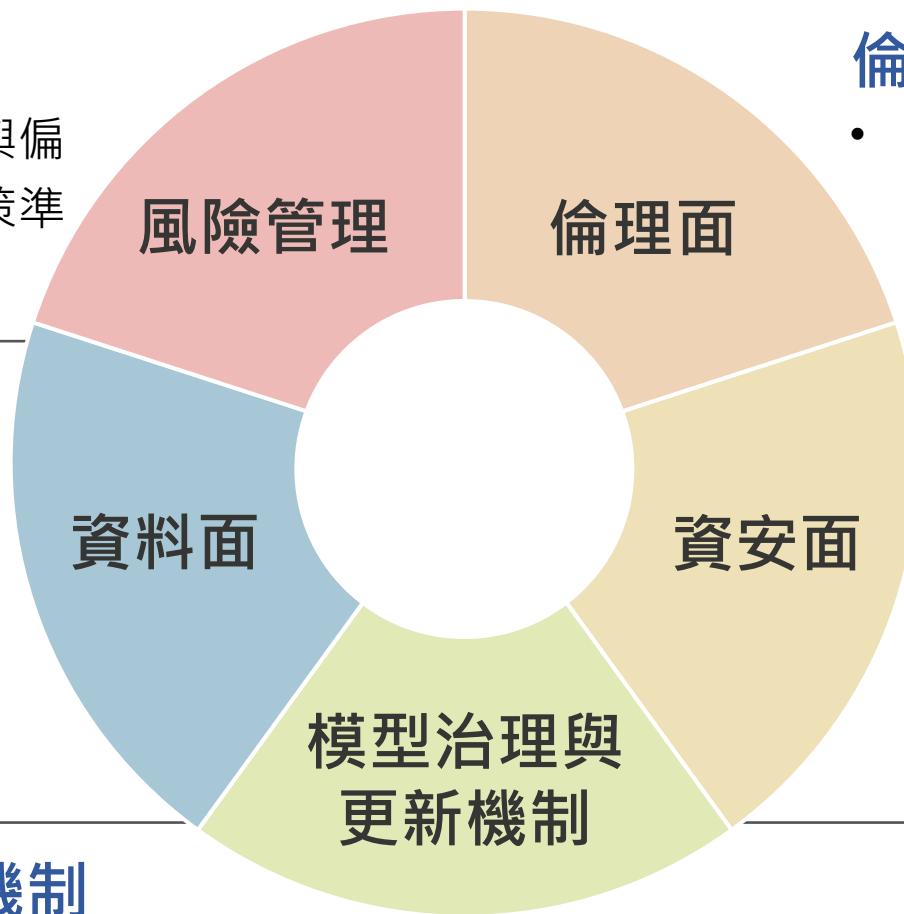
- 定期監控模型效能與偏誤變化，以維持決策準確與一致性

資料面

- 確保合法性、去識別化與資料最小化原則等。

模型治理與更新機制

- 應建立模型生命週期管理制度，包含訓練資料記錄、模型版本控管、偏誤與效能監測等。
- 高風險AI系統應設置自動監控工具，定期偵測模型漂移 (Model Drift) 與推論錯誤率。
- 對於已部署模型，應至少每年進行一次再訓練與驗證，並評估是否應汰換或更新模型。



倫理面

- 持續落實公平性、透明性、可解釋性、問責性、人為監督及隱私保護等原則。

資安面

- 傳統防護機制外，亦應強化對抗樣本攻擊、模型竊取與資料中毒等AI特有風險控管。

01

本作業指引 之修訂

本作業指引之修訂

本作業指引先行試辦，後續將依中央或本府法規之增修、各機關反饋意見或當人工智慧的技術有大幅躍進時，每年進行修訂。

01

附件

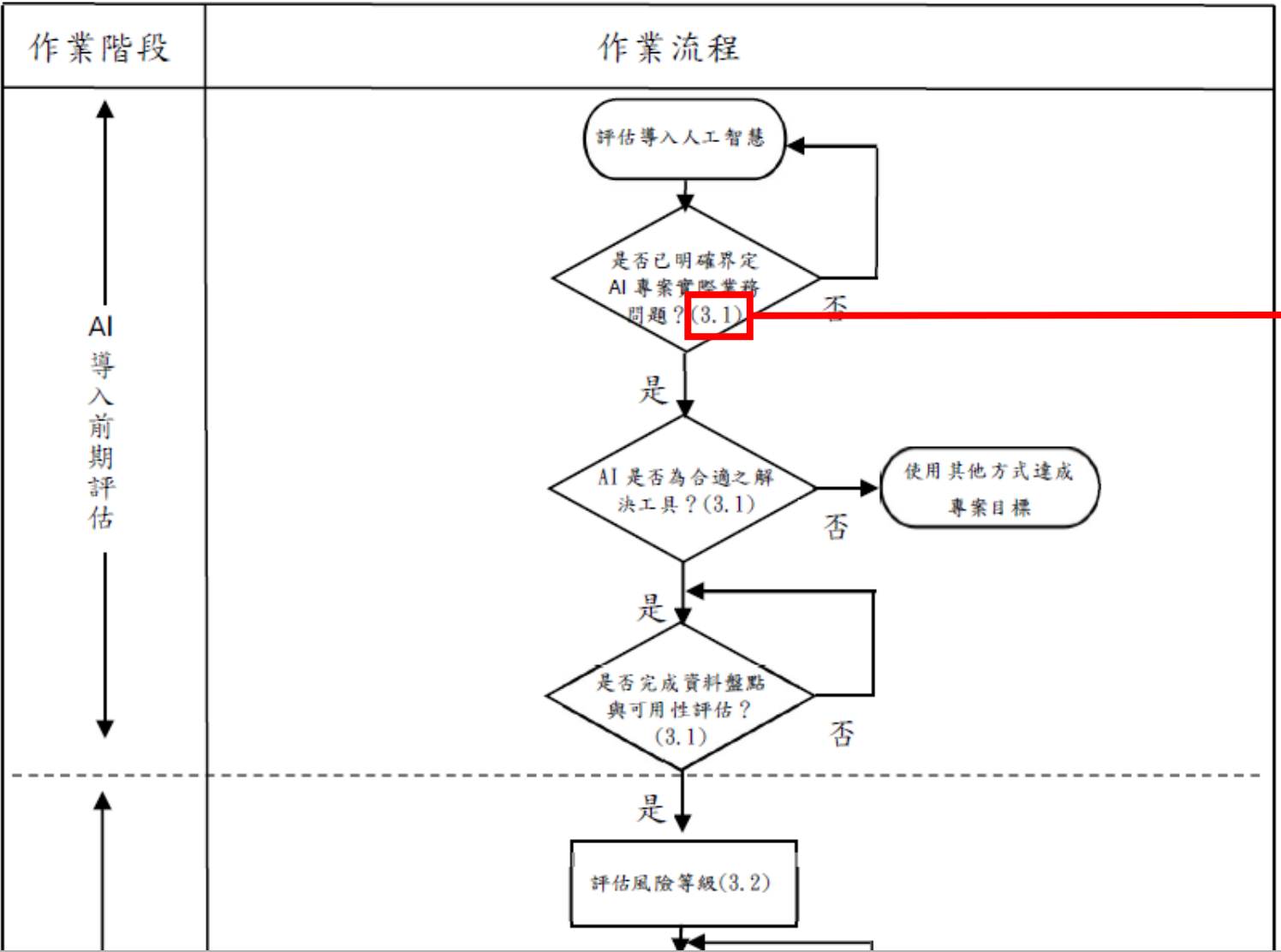


附件一、AI 於政府機關應用場景及風險等級對照表

政府機關可能的應用場景及風險等級對照，及相關判斷說明，**最終判斷仍依本中心認定為準**

應用場景	內容	可能的風險等級	風險等級判斷說明
智慧客服系統	透過生成式 AI 提供即時問答、辦理進度查詢與政策說明，減少人力負擔、提升服務效率。	有限風險	屬於人機互動型應用，可能產生資訊不對稱，使用者應知其對象為 AI 系統。
公文起草與文字生成	協助撰寫公文、新聞稿、政策報告初稿，加速行政流程並提升文字品質與一致性。	有限風險	為內部行政支援用途，屬內容生成型 AI，應提醒使用者注意內容為 AI 產生。
民眾意見摘要與分析	整合線上民意調查、社群反應等，生成摘要報告供決策參考。	有限風險	若僅摘要公開意見並輔助決策，風險較低，但應注意資料來源偏誤與去識別化處理。
教育訓練教材生成	快速製作客製化的內部訓練教材或數位學習資源，提升人員專業能力。	最低風險	屬非關鍵用途，不涉權益判定或安全影響。
語音與影像內容製作	利用 AI 生成語音導覽、宣導短片或手語翻譯影片，加強政策宣傳的包容性與普及率。	有限風險	若含 Deepfake 或生成音訊、手語翻譯等內容，需明示為 AI 產出以避免誤導。

附件二、新北市政府導入AI 內部控制流程圖



可對應本文
第參章第1條

附件三、AI 專案建置自主檢核表

填寫日期： 年 月 日		
AI 專案建置自主檢核表		
承辦機關		
聯絡人	姓名/職稱/電話：	
專案名稱		
專案內容(請簡述)		
專案期程		
風險等級	☐ 不可接受 ☐ 高風險 ☐ 有限風險 ☐ 最低風險 風險評估說明(請簡述)：	
檢核項目	自評結果 (填否及不適用，請 於備註說明原因)	備註
一、AI 導入前期評估(依循《公部門人工智慧應用參考手冊》)		
1. 是否已明確界定 AI 專案所對應的實際業務問題?(分析服務現況與痛點，界定 AI 應用目標)	☐ 是 ☐ 否	
2. 是否已評估 AI 是否為合適之解決工具?(確認 AI 優於其他解法，並具成本效益與技術可行性)	☐ 是 ☐ 否	
3. 是否完成資料盤點與可用性評估?(包括資料可取得性、格式、品質、完整性)	☐ 是 ☐ 否	
二、AI 通用性規範		
1. 是否未直接使用 AI 資訊作為行政行為或決策依據?(需先行審核驗證，避免誤導或錯誤決策)	☐ 是 ☐ 否 ☐ 不適用	

由機關先行自評，實際最終風險等級依本中心認定為準

自檢表已包含本指引中建置時需遵守項目

01

機關配合事項



機關配合事項-系統面

新案(9/30起)

應將不同風險須執行的管理機制(第3章第2條)、須繳交的相關技術文件(第4章第8條) 及營運管理要求事項(第3章第3條)納入契約中

現有案件

仍需填寫自主檢核表，但有關不同風險的管理機制、相關技術文件及營運管理要求事項，可待辦理後續採購案再執行，**另行發函請機關於10/17日前以email方式提供自主檢核表**

機關配合事項-認知面

同仁使用AI執行業務所需(簡報製作、公文、文稿撰寫等)，使用到chatGPT、Copilot等工具，亦須遵守以下規定

第肆章第1條：包含AI 產出之資訊，因由使用人進行客觀且專業之最終判斷。

第肆章第2條：注意個人資料保護、著作權及相關資訊使用等規定。

第肆章第7條第1款：機密文書禁止使用生成式AI。

第肆章第7條第2款：不得向生成式AI 提供涉及公務應保密、個人及未經機關（構）同意公開之資訊。

01

QA



• 01

THE END
