新北市政府人工智慧應用內控管理要點逐點說明

規　　定	說　　明
名稱： 新北市政府人工智慧應用內控管理要點	本要點之名稱。
一、　新北市政府（以下簡稱本府）為確保於運用人工智慧執行業務或提供服務時，能有效管理潛在風險並保障人民基本權利，落實人工智慧基本法第十九條規定，特訂定本要點。	本要點訂定目的。
二、　本要點適用於本府所屬各機關、區公所(包含烏來區公所）及學校（以下簡稱各機關）自行開發、委外建置或採購之各類人工智慧系統與服務。 各機關得依業務特性或特殊需求，另定人工智慧使用規範或內控管理規定，並報送本府備查。	明定本要點之適用對象及其範疇，並考量所屬機關（構）之業務特殊性，得另定嚴格於本要點所定基準之規範，且該規範應提送備查。
三、　本府人工智慧系統內控監督與內控措施推動之統籌權責單位，由本府資訊中心及相關單位另案籌組。	本府人工智慧系統內控之統籌權責單位另案籌組。
四、　各機關自行開發、委外建置或採購各類人工智慧系統與服務，應遵循人工智慧基本法第四條所定七項原則。	本府人工智慧系統與服務於自行開發、委外建置或採購階段，應遵循之核心科技倫理與基本原則。
五、　人工智慧系統與服務之資通安全防護，應依資通安全管理法與其相關法令及各機關資通安全維護計畫辦理；涉及個人資料蒐集、處理或利用時，應依相關法令進行事前評估與控管。	本府人工智慧系統之資安防護與個資保護法律遵循義務。
六、　各機關於規劃自行開發、委外建置或採購人工智慧系統與服務前，或既有人工智慧系統與服務發生重大功能變更時，應填具新北市政府人工智慧應用風險評估檢核表(如附件一，以	本府人工智慧系統於自行開發、委外建置、採購、重大功能變更時應進行風險評估。

下簡稱風險評估檢核表)，識別潛在風險。	
七、　各機關依風險評估檢核表綜合判定結果為甲類者，應辦理下列內控措施： (一)人類監督保留：不得將具法律效果或影響民眾權益之行政處分，完全交由人工智慧自動裁定，應保留適當之人為介入與最終覆核機制。 (二)透明標示：系統與服務自動生成之圖文內容或提供對外互動服務（如客服機器人）時，應於顯著處標示「本內容由人工智慧輔助/生成」等提示。 (三)注意事項或警語：應明確標示注意事項或警語。 (四)建立救濟機制：針對受人工智慧輔助決策影響權益之民眾，應提供明確之申訴管道、決策解釋及權益救濟流程。 (五)強化防護：應於採購契約或系統與服務規格中，加重對供應商之資通安全、資料治理與防範歧視偏見之要求。 (六)上線前檢測：於系統上線前應提供 AI 檢測申請單（如附件二）送至本府資訊中心進行檢測。	經各機關判定為甲類人工智慧風險等級時，應落實執行之六大核心內部控制措施。
八、　各機關依風險評估檢核表綜合判定結果為乙類者，應辦理前點各款內控措施，並依機關資源，優先處理甲類風險部分。	經各機關判定為乙類人工智慧風險等級時，應落實執行前點之六大核心內部控制措施，但優先處理甲類風險。
九、　各機關依風險評估檢核表綜合判定結果為丙類者，應辦理第七點第一款及第二款內控措施。	經各機關判定為丙類人工智慧風險等級時，應落實執行第七點第一款至第二款核心內部控制措施。
十、　各機關依風險評估檢核表綜合判定結果為無顯著風險者，得免辦理第七點所定內控措	經各機關判定為無顯著風險人工智慧風險等級時，應免辦理第七點內部控制措施，惟仍應保留

施。但應依風險評估檢核表填列評估結果，並保留相關紀錄備查。	相關紀錄備查。
十一、　人工智慧系統與服務於運行中，若發生綜合判定結果為甲類、演算法嚴重偏差、生成重大違法內容、資料外洩或對民眾權益產生不法侵害等異常情事，各機關應立即暫停系統與服務部分或全部功能，並依據本府資通安全相關規範進行通報及應變。	建立人工智慧系統運行階段之「異常應變與緊急煞車機制」，依本府資通安全相關規範進行通報及應變。
十二、　各機關應定期辦理並鼓勵同仁參與人工智慧素養、資訊倫理及風險管理之教育訓練，以提升人員對人工智慧技術局限性及資通安全風險之認知。	各機關應定期推動教育訓練，深化內部人員對於人工智慧科技風險及技術倫理之素養。
十三、　本要點由本府資訊中心定期檢視，並適時做滾動式檢討。	由本府資訊中心進行內控機制之定期檢視與滾動式檢討修正機制。