V2026.01.0901

公務AI共通核心能力認證

Government AI Common Core Competence Certification

學習指引

財團法人電腦技能基金會
採認編號：TAI-01-C-N

認證網站

前言

財團法人中華民國電腦技能基金會依據行政院人事行政總處公務人力發展學院所規劃之公務 AI 共通核心能力需求，發展之「公務 AI 共通核心能力認證」，已正式通過行政院人事行政總處採認。為配合政府推動「AI 政府領航人才發展計畫」，基金會將積極發揮民間優質 AI 能力鑑定的專業能量，協助公務人員系統性培養人工智慧知能，共同擴大公務 AI 人才培育與發展效益。

為協助應考人士掌握認證主題方向並有依循準備，本會製作學習指引對本認證各主層面內容以及對應之技能規範，進行重點說明與樣題解析。

本學習指引旨在提供考生學習方向與準備參考，內容並非正式教材或完整題庫，亦不代表實際考試試題，無法作為通過認證之保證。建議考生依據考試簡章所公告之認證主題與內容，進行完整且充分的準備，以提升學習與應試成效。

目　錄

表　目　錄

壹、 公務 AI 共通核心能力認證概述

本認證依據數位發展部「AI 公務人才認定指引」所建構之能力架構進行規劃，旨在培育具備 AI 基礎素養、實務應用及治理管理能力之公務人才。認證內容聚焦於「學 AI、用 AI、管 AI」三大核心能力，除強調公務人員應具備 AI 工具運用與實務應用能力外，亦重視法規遵循、風險管理、資訊安全及公共價值等治理面向，以全面提升公務 AI 應用與治理能力。

認證內容採「AI 產業最新發展內容融入」及「公務專屬內容客製開發」雙軌設計。前者持續納入 AI 技術、應用及產業發展等最新知識，協助掌握 AI 發展趨勢與實務應用；後者則依公務工作特性與實際需求，強化公務情境應用、政策治理及法遵等專屬內容，兼顧 AI 通用能力與公務實務需求，建構完整的公務 AI 能力學習與評量架構。

試題內容依核心能力與主題均衡配置，涵蓋 AI 基礎認知、實務應用、政策治理、法規遵循及公共服務等重要面向，使考生能從不同層面掌握公務 AI 所需具備的知識與能力。各核心主題之試題配置將依年度題庫分析、施測結果及實務發展趨勢持續檢討與優化，適時更新相關內容，以維持認證內容之適切性、代表性及有效性。

表 1. 公務 AI 共通核心能力認證結構一覽表

認證主題		AI產業最新發展範圍	公務專屬情境範圍
主層面	次層面		
1. AI 基礎認知	1.1 AI基礎概念與社會影響	AI概論、數位素養、發展趨勢、AI影響	基礎概論與AI素養/AI發展趨勢與公共治理影響/AI工具發展與系統介接認知
	1.2 生成式AI與人機互動	生成式AI基礎、智慧轉型、未來職場、人機互動	提示詞（Prompt）實務與迭代/人機協作認知與內容把關/任務監督與進階協作
	1.3 AI應用與公共服務轉型策略	IT趨勢、AI發展、AI願景、產業衝擊	預測型vs生成式AI區辨/模型選擇素養/數位平權與科技包容/AI導入與行政流程轉型
2. AI 政策法制	2.1 AI政策與治理指引	AI政策、AI治理指引	政策目標與跨部會指引/風險分級與監理/資料治理與AI訓練使用判斷/AI系統成效評估與持續監測/AI導入治理與生命週期管理
	2.2 AI法規與資安風險管理	AI安全、法規規範	資安規範與資料保護/AI法規與資安風險管理實務/生成式AI使用限制與適法性
	2.3 AI應用下的倫理準則與智慧財產實務	智慧財產權、AI倫理原則	最終問責與行政倫理/智財權、授權條款與使用揭露/資料可信度與透明度

公務專屬情境範圍之規劃設計

本認證之公務專屬情境以公務實務需求為核心進行規劃，著重於公務人員實際運用 AI 工具時所需具備的應用、判斷及治理能力，培養「會使用、能判斷、懂治理」的公務 AI 核心素養。

內容設計以公務場域常見業務情境為基礎，涵蓋公文處理、民眾服務、政策研擬、資料分析、採購作業、跨機關協作及行政流程優化等應用場景，透過情境式及案例導向的內容，引導考生理解如何在實際工作情境中適切運用 AI 工具，並具備辨識問題與判斷潛在風險的能力。

此外，為強化公務機關運用 AI 所需的治理能力，公務專屬內容將納入數位素養、科技包容、資訊偽冒辨識、AI 政策、資料保護、資訊安全、透明性、可解釋性、智慧財產權及倫理責任等重要議題，補足一般 AI 應用能力之外，公務場域所需的特殊能力。

為確保認證內容的適切性與專業性，將邀集產、官、學界專家參與內容規劃、命題及審題作業，檢視內容是否符合公務實務需求、政策發展方向及現行法規規範。同時透過試題分析、專家審議及定期內容更新機制，持續調整試題內容與配置，適時納入 AI 技術發展及政府智慧治理的最新需求，維持認證內容的適切性、專業性與時效性。

表 2. 公務專屬情境題綱對應說明表

認證主題		評測題綱
主層面	次層面	
1. AI基礎認知	1.1 AI基礎概念與社會影響	基礎概論與AI素養、AI發展趨勢與公共治理影響、AI工具發展與系統介接認知
	1.2 生成式AI與人機互動	提示詞（Prompt）實務與迭代、人機協作認知與內容把關、任務監督與進階協作
	1.3 AI應用與公共服務轉型策略	預測型vs生成式AI區辨、模型選擇素養、數位平權與科技包容、AI導入與行政流程轉型
2. AI政策法制	2.1 AI政策與治理指引	政策目標與跨部會指引、風險分級與監理、資料治理與AI訓練使用判斷、AI系統成效評估與持續監測、AI導入治理與生命週期管理
	2.2 AI法規與資安風險管理	資安規範與資料保護、AI法規與資安風險管理實務、生成式AI使用限制與適法性
	2.3 AI應用下的倫理準則與智慧財產實務	最終問責與行政倫理、智財權、授權條款與使用揭露、資料可信度與透明度

貳、 認證主題 1：AI 基礎認知

1.1. AI 基礎概念與社會影響

本內涵旨在建立公務人員對人工智慧(AI)之基礎概念、發展趨勢、應用方式及社會影響的基本認識，並培養 AI 時代所需之數位素養。內容涵蓋 AI 基本技術概念、生成式 AI、公共治理應用、演算法偏誤、資訊辨識及人機協作等，並引導公務人員理解 AI 應用可能帶來的社會影響與風險，在推動智慧行政的同時，兼顧人類自主、資料治理、公平與不歧視、透明與可解釋等原則。

- AI 基礎概念與數位素養

人工智慧已廣泛應用於各領域，公務人員應具備基本 AI 素養，了解人工智慧、機器學習及生成式 AI 等基本概念，以及 AI 如何運用資料進行分析、預測或產生內容，於 AI 時代的數位素養不僅是資訊工具的操作能力，亦包括資訊辨識、理解 AI 產出限制及人機協作等能力。由於生成式 AI 產出可能存在錯誤或虛構資訊，使用時應保持適當的判斷與查證能力。行政院生成式 AI 參考指引亦強調，生成式 AI 產出資訊應由業務承辦人進行客觀且專業的最終判斷，不得完全信任 AI 產出。

- AI 應用與公共治理影響

AI 已逐漸應用於政府業務及公共服務，可協助資料分析、業務處理及服務提供。然而，AI 應用於公共治理時，也可能產生演算法偏誤、歧視、資訊不透明及對民眾權益造成影響等風險。例如，若資料本身存在偏差，可能影響 AI 分析或判斷結果。因此，公務人員應理解 AI 應用可能產生的社會影響，並依應用情境進行適當的風險辨識與管理。

- 生成式 AI 與資料治理

生成式 AI 可協助公務人員進行文字撰寫、資料整理及內容產製等

工作，但其使用涉及資料保護、資訊安全、個人資料及資訊正確性等議題。公務人員使用生成式 AI 時，不得提供涉及公務應保密、個人及未經機關同意公開之資訊；使用 AI 產出之內容，也應進行適當查證。

《人工智慧基本法》將「隱私保護與資料治理」列為政府推動 AI 的重要原則，並強調資料應妥善管理、降低資料外洩風險，同時促進非敏感資料之開放與再利用。政府目前亦持續推動資料開放、共享、再利用及資料管理等相關工作。

因此公務人員應具備基本的資料治理觀念，能依資料性質及業務需求判斷資料是否適合提供 AI 使用，並注意資料的正確性、適法性及安全性。

- 工具發展與系統介接

AI 技術持續發展，應用型態亦由單純的內容生成及問答，逐步擴展至結合不同資料與系統的應用，公務人員應具備 AI 工具及系統介接的基本認識，了解 AI 應用可能涉及資料來源、系統介接、權限管理及資訊安全等事項。

機關導入 AI 應用時，應依業務需求及應用情境進行評估，並注意資料使用、資訊安全及系統管理等相關事項。數發部《公部門人工智慧應用參考手冊》即以 AI 場景評估、專案啟動、資料探索與模型建立、模型迭代與部署，以及風險控制與專案追蹤等階段，提供各機關導入 AI 時參考。

1.1 參考樣題及解析

1	A	科學家在 1950 年代提出用來檢驗「機器表現出的行為是否與人類無法分辨」的圖靈測試（Turing Test），但隨著 ChatGPT 等生成式 AI 能寫出像人類一樣流暢的文字，請問目前生成式 AI 可以具有何項能力？ (A) 具備了強大的多模態識別與語言生成能力 (B) 理解人類的語境、幽默感、情感及邏輯反應 (C) 擁有人類的思考能力或情感 (D) 邏輯推理能力與常識與真實性
解析		生成式 AI 本質上是透過數據統計來模仿人類的語言行為，並不具備真正的理解力、自我意識或情感體驗，尚未達到通用人工智慧（AGI）的思考層次。同時，由於缺乏對物理世界的常識經驗，生成式 AI 在複雜邏輯推理上仍有缺陷，且容易產生「幻覺」而輸出錯誤資訊，無法確保絕對的真實性。
2	D	機關資訊單位於教育訓練中說明機器學習（Machine Learning）的核心概念，機器學習並非單純儲存資料，而是希望系統能透過大量資料自動找出規則並持續改善判斷能力， 請問下列哪一項不屬機器學習的主要作為？ (A) 利用特定演算法從既有的歷史數據中，自動萃取關鍵特徵並建構出具預測能力的規則模型 (B) 對收集到的原始歷史資料進行清洗、去噪與特徵擷取，以建構出適合模型學習的訓練資料集 (C) 使資訊系統具備從大量未知資料中，自主演繹出隱含規律並持續疊代優化規則的演算能力 (D) 由系統工程師事先針對所有可能的輸入組合，在程式中撰寫對應的邏輯條件與輸出規則

解析		機器學習著重於「數據預處理 (B)」、「自動提取特徵與建置模型 (A)」與「自主演繹與疊代優化 (C)」；而由「工程師事先寫死所有條件規則 (D)」則為典型的傳統程式開發，非機器學習範疇。
3	B	市政府擬採購自動駕駛接駁車，廠商在簡報時表示車載系統將運用卷積神經網路（CNN）技術。請問在自動駕駛的實際應用中，此技術最主要扮演下列哪一種角色？ (A) 自動生成路徑規劃 (B) 分析和辨識攝影機拍攝的圖像 (C) 處理語音命令 (D) 優化燃料消耗
解析		CNN 是深度學習領域中專門用於影像處理與電腦視覺（Computer Vision）的核心架構，CNN 專精於「影像與視覺分析 (B)」；而路徑規劃、語音處理與能耗優化則分別屬於控制決策、自然語言處理與系統優化的範疇。
4	B	組織規劃利用 AI 系統分析教育訓練的滿意度問卷，為了避免 AI 模型在分析過程中產生系統性偏見，下列哪一種做法最符合「資料去偏（Data Debiasing）」的原則？ (A) 在滿意度問卷中加入大量無關的干擾數據，降低 AI 的偏見敏感度 (B) 依照資料分布，平衡各層或人數較少部門的資料比例 (C) 讓 AI 在輸出分析結果時隨機替換關鍵字，用隨機性來抵消特定偏向 (D) 升級使用參數量更大的最新 AI 模型，依賴大型模型自動消除資料中的偏見

解析		資料去偏是針對資料本身可能存在的偏差進行適當處理，以降低資料分布不均對 AI 分析結果造成的影響。若不同部門的問卷回覆數量差異過大，可依資料分布情形適度調整或平衡各群體的資料比例，使分析結果較能反映不同群體的實際情況。因此，依資料分布平衡各層或人數較少部門的資料比例，符合資料去偏的原則，故 B 選項正確。
5	B	下列哪一項職場應用情境與機器學習類型的搭配是正確的？ (A) 為了加速處理大量客服信件，主管拿過去幾萬筆已由人工分類好類別的客服信，讓 AI 學習自動分類新信件 — 半監督式學習（Semi-supervised Learning） (B) 人事部門想從上萬封履歷中找出優秀人才，但手邊完全沒有任何標籤或分類標準，於是讓 AI 根據字詞特徵，自動將履歷劃分成若干個特徵群組 — 非監督式學習（Unsupervised Learning） (C) 行銷部門為了預測會員是否流失，主管決定結合極少數以已標註與大量未標註的會員資料來建立預測模型 — 強化式學習（Reinforcement Learning） (D) 業務團隊為了優化官方網站的 AI 智慧客服助理，讓智能客服在與客戶對答中，透過客戶點擊「讚或倒讚」的獎懲訊號來逐步調整對話回覆策略 — 監督式學習（Supervised Learning）
解析		完全有標籤 ➔ 監督式學習 完全無標籤（自動分群） ➔ 非監督式學習 少量有標籤 + 大量無標籤 ➔ 半監督式學習 透過獎懲反饋優化策略 ➔ 強化式學習 資料無任何預先定義的標籤，模型依據文字特徵的相似度進行自我分群（Clustering），屬於典型的非監督式學習應用。

6	A	為了提升為民服務品質，機關規劃利用 AI 系統分析歷年的民眾陳情案件，並從中找出潛在規律，用以提前預警高風險的民怨議題。請問此類運用歷史資料進行未來趨勢推估的作法，主要屬於下列哪一種 AI 應用類型？ (A) 預測型 AI (B) 圖像生成型 AI (C) 音樂生成型 AI (D) 遊戲型 AI
解析		預測型 AI（或稱分析型 AI）的核心範式為「分析歷史數據 ➔ 辨識規律與趨勢 ➔ 進行未來預測或風險預警」。透過統計學、機器學習演算法（如迴歸分析、時間序列分析、分類模型等），對龐大的既有資料進行挖掘，以推估未來發生的可能性，皆屬於預測型 AI (A) 的應用範疇。
7	C	依據《人工智慧基本法》中強調之「透明與可解釋性」精神，下列何者最符合 AI「可解釋性」之概念？ (A) 提升 AI 系統資料處理速度 (B) 增加 AI 模型運算複雜度，讓可解釋性更高 (C) AI 模型提供做出判斷的原因及引用資料來源 (D) 提高 AI 自動生成內容的數量
解析		在 AI 治理與公務應用中，「可解釋性」是指 AI 系統進行預測、分析或輔助決策時，能夠以人類可理解（Human-understandable）的方式，說明其得出該結果的邏輯理由、關鍵影響因子與參考資料來源。其目的在於打破 AI 的「黑盒子（Black-box）」現象，確保決策具備透明度與可預測性。

1.2. 生成式 AI 與人機互動

生成式 AI 的實務操作與人機協作模式，協助組織達成智慧轉型，內容深入公務日常應用場景，如公文撰擬與民眾陳情回應，並教導提示詞（Prompt）設計與迭代優化技巧。同時強調人機協作中的任務監督機制與 AI 產出品質管理，確保公務處理之精確度、合規性與效率，打造流暢且可信賴的人機協作環境。

- 生成式 AI 與組織轉型

生成式 AI（Generative AI）的快速崛起，正驅動政府機關進行深度的數位轉型與組織流程重塑，傳統公務行政長期面臨資料分散、人工比對耗時與行政流程僵化的挑戰；導入生成式 AI 後，運作模式從「被動資料處理」轉化為「主動智慧輔助」，透過語言模型對巨量公務文檔的快速彙整與研析能力，能有效優化跨單位協調效率、簡化行政審查程序，並賦予同仁更多時間從事具高附加價值的政策規劃與決策分析，組織推動轉型的關鍵，在於建立開放且兼顧資安的工具環境，結合流程自動化（Agentic Workflow），將 AI 深度融入日常作業，引導公務體系從傳統的單一「工具使用者」，轉型為具備前瞻思維的「智慧組織」。

- 公務實務場景應用

生成式 AI 在公務體系的實務應用已涵蓋諸多行政場景，顯著提升了為民服務與內部辦理效能，在文書處理方面，AI 能輔助同仁快速擬訂各類公文草稿、研析陳情案件訴求，並將冗長複雜的法令政策轉化為民眾易懂的白話宣傳圖文；在會議與決策輔助上，AI 能將幾小時的語音逐字稿精準彙整為待辦事項清單與會議紀錄；在臨櫃與第一線服務中，結合知識庫（RAG）的智慧 FAQ 機器人，能即時提供民眾正確的申辦指引與資格條件，透過這些具體的應用場景，生成式 AI 不僅有效減輕了基層公務員的重複性行政負擔，更精準提升了公務處理的品質與回應民眾的速度。

- 提示詞設計與迭代優化

提示詞（Prompt）是公務員與生成式 AI 溝通的核心橋梁，提示詞撰寫的精準度直接決定了 AI 輸出內容的實用性與準確度，在公務應用中，良好的提示詞設計需具備結構化思維，明確設定「角色（Role）」、「背景與任務（Context & Task）」、「輸出格式（Format）」及「語氣風格（Tone）」。然而，單次提問往往難以產出完美的公務文本，公務同仁需掌握「迭代修正（Iterative Prompting）」與「思維鏈（Chain-of-Thought）」技巧，透過分步拆解問題、提供少樣範例（Few-Shot Learning）或根據 AI 的初稿給予修正指令，引導 AI 逐步調整與優化內容，確保產出完全符合公務機關業務需求。

- 人機協作與品質監督

在生成式 AI 輔助公務處理的模式中，「人類在環／人類監督（Human-in-the-loop）」是確保行政正確性與法理責任的不可或缺的機制。儘管 AI 具備強大的文本生成與資料整理能力，但仍可能產生「幻覺（Hallucination）」或不準確的法規引用。公務人員的角色必須升級為「任務監督者」，建立嚴格的品質控管與事實查核 SOP。所有由 AI 輔助草擬的公文、陳情回覆或政策分析報告，於最終決行或發布前，皆須經過公務同仁的實質審查、邏輯確認與法規核對。堅守「AI 輔助創作，人類審查決行」的原則，方能在善用科技提升效率的同時，維護公務處分的公正性、嚴謹度與機關公信力。

1.2 參考樣題及解析

1	B	聊天機器人（Chatbot）屬於下列哪一種人工智慧的應用領域？ (A) 影像處理 (B) 自然語言處理（NLP） (C) 文生圖處理 (D) 圖像處理
解析		聊天機器人主要透過自然語言處理（NLP）理解使用者輸入的文字或語音內容，並產生適當的回應，使人與電腦能以自然語言進行互動。因此，聊天機器人屬於自然語言處理的主要應用之一，故 B 選項正確。
2	C	同仁使用生成式 AI 協助草擬「活動延期通知公文」，第一次僅輸入「幫我寫活動延期通知」，結果內容過於口語且格式不完整，若希望 AI 產出包含正式公文格式、通知對象及延期原因等內容，下列哪一種 Prompt（指令）設計方式最適當？ (A) 請幫我寫活動延期通知，不限格式，內容簡短即可 (B) 請協助產生活動延期通知，可自由發揮內容，不需特別說明對象 (C) 請以正式公文格式，撰寫給參訓單位之活動延期通知，並包含延期原因、通知對象與聯絡資訊 (D) 請整理活動延期重點，內容口語化並以社群貼文方式呈現
解析		Prompt 的內容越明確，AI 越能依照使用者的實際需求產生適當結果。撰寫公務文件時，應明確說明文件類型、使用對象、格式及必要內容，例如指定正式公

		文格式、通知對象、延期原因及聯絡資訊，可降低產出內容與實際需求不符的情形。因此 C 選項的 Prompt 設計最為完整且適當，故答案為 C。
3	D	在法律行業中，生成式 AI 可以應用於多個方面，但下列哪一項不是生成式 AI 目前能夠實現的？ (A) 自動生成法律文件草稿，如合約和訴狀 (B) 預測案件結果，基於歷史案件數據 (C) 生成法律意見書，提供客製化法律諮詢 (D) 完全替代人類律師進行法庭辯論和交叉審問
解析		生成式 AI 已可應用於法律文件草擬、案例資料分析及法律研究等工作，並可作為法律專業人員的輔助工具。然而，涉及法庭辯論、交叉詰問及法律專業判斷等工作，仍需要由具備法律專業能力的人員負責，不能完全由 AI 取代。因此，D 選項所述「完全替代人類律師進行法庭辯論和交叉審問」並非目前生成式 AI 能夠完全實現的應用。
4	B	承辦人於審查「活動宣導設計案」服務建議書時，廠商提到將運用 AI 自動生成宣傳圖片與活動主視覺，此類應用最適合使用下列哪一項 AI 技術？ (A) 決策樹（Decision Tree） (B) 文生圖模型（如：Midjourney、Stable Diffusion） (C) 支援向量機（SVM） (D) 強化學習（RL）
解析		專門用於根據文字指令「自動生成圖片與視覺設計」的生成式 AI 技術，B 選項說明符合題幹中生成宣傳圖片與活動主視覺的需求，凡是以文字提示詞產生海報、主視覺或宣傳圖片者，皆屬於文生圖模型；而決

		策樹 (A)、SVM (C) 與強化學習 (D) 則分別用於邏輯判斷、數據分類與策略控制，不具備圖像創作功能。
5	ACD	機關在撰寫「生成式 AI 發展趨勢」的參考資料時，需要評估多模態生成的應用發展，關於生成式 AI 的敘述下列哪些正確？ (A) 生成式 AI 是透過大量資料訓練，學習生成規則，進而自動產出對應內容的技術 (B) 生成式 AI 僅能用於生成圖像資料，無法生成文字或音訊等其他類型的內容 (C) 生成式 AI 常用的技術包括生成對抗網路（GAN）、長短期記憶網路（LSTM）和 Transformer 模型 (D) Transformer 模型是一種在生成式 AI 中相當常見且核心的深度學習架構
解析		生成式 AI 具備強大的多模態處理能力，廣泛應用於文字創作（ChatGPT）、語音合成（Suno）、影像生成（Midjourney）及影片生成（Sora）等，非僅限於圖像生成。
6	C	關於生成式 AI 的應用敘述，下列哪一項是錯誤的？ (A) 生成式 AI 可以應用於程式碼產生、工程、行銷、客戶服務、金融和銷售等多個業務範圍 (B) 生成式 AI 可以用於改善客戶體驗，例如通過聊天機器人、虛擬助理和智慧聯絡中心等功能 (C) 生成式 AI 僅能在程式碼產生方面取得成效，其他應用場景如內容建立和文字摘要無法實現 (D) 生成式 AI 可以加速藝術和音樂中的創意內容製作，如文字、動畫、影片和影像等
解析		生成式 AI 可依據使用者提供的指令或資料產生新的內容，應用範圍相當廣泛，除程式碼產生外，也可

		應用於文字摘要、內容建立、客戶服務、影像、動畫及影片等。因此，選項 C 將生成式 AI 的應用侷限於程式碼產生，並認為其他內容建立及文字摘要無法實現，與實際應用情形不符
7	B	隨著人工智慧（AI）與物聯網（IoT）技術的快速發展，兩者匯流形成的「AIoT」已成為推動智慧應用的核心驅動力，在 AIoT 的架構分工中，物聯網（IoT）的主要功能不包含下列哪一項？ (A) 資料感測 (B) 資料統計與決策分析 (C) 資料收集與彙整 (D) 資料傳輸
解析		IoT（物聯網）的角色→「感知器官與神經網路」：負責實體世界的環境數據擷取（感測）、初步彙整與跨網路傳輸，將數據安全送達雲端或邊端。 AI（人工智慧）的角色→「大腦」：負責對 IoT 所收集到的龐大數據進行學習、統計、演算法模式辨識、預測與智慧決策分析。
8	B	區公所規劃導入「代理型 AI（Agentic AI）」協助同仁處理多步驟的社會福利補助審查作業。該 AI 能自主跨系統比對戶籍資料、計算補助資格並草擬核定公文。為了善盡「任務監督者」的職責，並落實「人類在環／人類監督（Human-in-the-loop）」機制，同仁在設定 AI 的運作邊界與監督節點時，下列哪一種作法最符合人機協作與風險管控原則？ (A) 授權 AI 自主完成資料比對、資格判定及公文自動發送，同仁僅需於事後定期抽查系統紀錄即可 (B) 設定 AI 僅能存取法規授權之資料庫，且於完成資格初審與公文草擬後，必須由同仁進行實質審查與核決，方可發文 (C) 要求 AI 在處理每一筆民眾資料時，皆須停下並

		由同仁手動輸入下一命令，確保 AI 不會自動執行下一個步驟 (D) 將所有審查標準與例外狀況完全撰寫為固定程式碼，取消 AI 的自主規劃能力，改回傳統的自動化流程
解析		代理型 AI 可自主執行多步驟任務，但涉及民眾權益或正式行政行為時，仍應設定適當的資料存取權限及人類監督節點。透過限制 AI 僅存取經授權的資料，並於資格初審及公文草擬完成後，由承辦人進行實質審查與核決，再進行後續行政作業，可兼顧 AI 的自動化效益與人類監督、責任及風險管控。因此，B 選項最符合人機協作與風險管控原則。
9	C	在機關同仁使用生成式 AI 協助整理跨部會會議紀錄，於 AI 自動產生之內容後，下列哪一項作法最適當？ (A) 不經人工確認，直接將 AI 產出內容作為正式會議紀錄發送各單位 (B) 請 AI 再次生成會議紀錄內容，比對確認後即可發送各單位 (C) 由承辦人逐項確認發言內容與結論後，依權責批核後發文 (D) 將 AI 產出內容作為初稿，並直接依據內容彙整會議結論
解析		跨部會會議紀錄涉及各機關權責分工與後續列管事項，具備法律與行政效力，依據機關使用 AI 應注意事項與公務倫理，AI 僅能作為「初稿整理工具」，最終的內容正確性查核、法規與權責核對，以及公文簽核發布，必須由人類承辦人負擔實質與最終的行政責任。

1.3. AI 應用與公共服務轉型策略

人工智慧已逐步應用於政府治理與公共服務，協助機關改善行政流程、提升服務效率及回應民眾需求。推動 AI 應用時，應以實際業務需求與服務問題為出發點，評估 AI 是否適合作為解決方案，避免為導入 AI 而導入 AI。數發部《公部門人工智慧應用參考手冊》將 AI 導入區分為「AI 場景評估、AI 專案啟動、資料探索與模型建立、模型迭代與部署、風險控制和專案追蹤」五大階段，提供機關規劃及導入 AI 應用時參考。

在公共服務轉型方面，政府可運用 AI 協助重新檢視行政流程及服務方式，從流程改善、服務創新等面向提升政府服務品質。國發會近年推動 AI 時代政府服務創新，亦強調運用 AI 科技賦能公共服務，重新思考行政流程，以更貼近民眾需求。

- AI 模型選擇與區辨

在公務體系推動 AI 應用時，應依業務需求及應用情境，選擇適當的 AI 技術。公務人員應具備區辨預測型 AI 與生成式 AI 的基本能力，理解兩者的主要用途與適用情境。

預測型 AI 主要利用既有資料進行分類、預測或分析，例如車流量預估、風險預測及案件分類等，可作為業務分析及決策支援工具；生成式 AI 則可依據使用者提供的指令或資料產生新的內容，例如協助撰寫公文草稿、整理資料、摘要文字或產製其他內容。

機關選擇 AI 技術或工具時，應依業務目的、資料特性及應用情境進行評估，並兼顧資訊安全、個人資料保護及相關法規要求，以選擇適合的 AI 應用方式。

- AgenticAI 與智慧行政

隨著 AI 應用發展，AI 可進一步結合多步驟任務處理及既有業務流程，協助機關執行資料整理、資訊分析及其他業務工作。機關如導

入具自主執行能力的 AI 應用，應依實際業務情境設定適當的使用範圍、權限及人類監督機制，並依應用可能造成的影響進行風險辨識、評估及應對。

對於涉及民眾權益或重要行政業務之 AI 應用，仍應保留適當的人類判斷及監督，避免將重要決策完全交由 AI 自動處理。數發部「AI 風險分類框架」即要求機關依序進行盤點應用情境、識別風險、評估風險及應對風險，並由各目的事業主管機關依其業務權責進行治理。

- 技術導入與數位包容

政府推動 AI 應用及公共服務轉型時，除考量行政效率，也應兼顧社會公平及數位平權，避免因數位能力、資源或使用條件不同而造成服務落差。《人工智慧基本法》已將「永續發展與福祉」列為政府推動 AI 的原則，並明定應降低數位落差、提供適當的 AI 教育及培訓；同時，AI 應用亦應兼顧人類自主、公平與不歧視等原則。

因此機關規劃 AI 公共服務時，應考量不同族群的使用需求及服務可近性，並視業務性質提供適當的服務方式，使民眾不因數位能力或使用條件差異而無法取得必要的公共服務。

1.3 參考樣題及解析

1	B	科技公司規劃開發智慧物流配送系統，希望 AI 系統能依據即時路況，自動且動態地規劃出最佳配送路線，此類需要即時應變與決策的 AI 應用，最適合採用下列哪一項技術範疇？ (A) 自然語言處理（NLP）與文字生成 (B) 強化式學習（Reinforcement Learning）（如遊戲 AI 或自動駕駛） (C) 影像分類與物件偵測 (D) 非監督式學習（Unsupervised Learning）的資料分群
解析		強化式學習適合應用於需要與環境互動並依據回饋持續調整決策的情境。題目中的智慧物流配送系統需根據即時路況變化，動態選擇適當的配送路線，具有持續決策與調整的特性，因此適合採用強化式學習。
2	A	AI 在社群媒體上的「推播（Push）」應用其實已經非常普遍，而且不只是推送通知而已，而是從內容推薦、廣告投放到互動觸發，幾乎都由 AI 在背後運作，如果從公務機關的角度來看，AI 推播的價值不在於「讓民眾一直收到訊息」，而是在於：在正確的時間，把正確的資訊推送給需要的人，這也是近年許多智慧政府（Smart Government）發展的重要方向。請問下列哪一項是效益最高公務機關的 AI 推播應用？ (A) 篩選符合政策補助資格民眾通知書 (B) 節慶祝賀訊息 (C) 電子報推播 (D) 客服專線異動公告

解析		此作法將 AI 的「數據分析與精準比對」能力極大化，落實「主動為民服務」，能直接保障特定群體的政策權益，並展現智慧政府發揮最大公共價值的核心目的，AI 推播最高效益的體現，在於「精準篩選資格並主動通知對應族群 (A)」；而非用於泛濫的節慶賀詞 (B)、傳統電子報 (C) 或全域性系統公告 (D)。
3	B	機關規劃導入 AI 輔助民眾服務系統，推動 AI 應用時應兼顧「人本價值」、「公平性」與「透明性」等原則，以符合《人工智慧基本法》之核心精神，下列哪一項做法或觀念最符合上述原則？ (A) 為追求全面自動化與效率，行政處分可完全交由 AI 系統獨立判定，不需預留人類介入空間 (B) 機關應明定資訊公開管道，並建立覆核與申訴機制以防範演算法偏見 (C) 只要 AI 系統能大幅提升行政效率，即可豁免或省略資安防護與社會影響評估程序 (D) 為保護合作廠商的商業機密，機關應全面隱瞞 AI 的運行邏輯與任何功能說明
解析		《人工智慧基本法》強調人類自主、公平與不歧視、透明與可解釋及問責等原則。公務機關導入 AI 輔助民眾服務時，除追求行政效率外，亦應兼顧民眾權益與 AI 應用可能產生的風險，透過適當的資訊揭露、人工覆核及申訴機制，提升 AI 應用的透明度與可問責性。因此，B 選項最符合上述原則。
4	B	市政府規劃辦理智慧交通系統建置案，預計透過路口攝影機即時分析車流並動態調整控制號誌，承辦同仁審查廠商服務建議書時，發現系統架構將導入「邊緣運算（Edge Computing）」技術，請問此技術主要是為了解決下列哪一項核心問題？

		(A) 資料儲存空間不足 (B) 資料傳輸頻寬負載與時間延遲 (C) 資料清理與去噪 (D) 模型複雜度與資料前處理
解析		邊緣運算是將資料處理與運算能力移至靠近資料產生端的位置，使資料不必全部傳送至遠端雲端進行處理。智慧交通系統需要即時分析車流並調整號誌，透過邊緣運算可減少資料傳輸量及降低傳輸延遲。
5	AC	小明任職的公司規劃導入生成式 AI 協助產製圖文內容，資訊單位於評估風險時指出，若 AI 在訓練或生成過程中未經授權使用個人資料，恐引發隱私權等重大爭議，請問下列哪些屬於生成式 AI 在道德與法律上的主要挑戰？ (A) 未經同意使用個人資料 (B) 提升圖片壓縮效率 (C) AI 生成內容可能侵犯隱私 (D) 增加硬體儲存容量
解析		生成式 AI 在法律倫理上的核心爭議包含「未經同意使用個資 (A)」與「生成內容侵犯隱私 (C)」，AI 生成的擬真圖像（如 Deepfake）或文字若揭露特定個人隱私、不實合成個人資訊，會構成侵犯隱私權與名譽權之重大法律風險；而圖片壓縮 (B) 與硬體容量 (D) 僅為技術與設備議題，不屬於道德法律挑戰。
6	C	一般民眾使用生成式 AI 協助整理個人資料、撰寫文件或查詢資訊時，若未注意資料使用方式，可能產生隱私權風險，下列哪一項敘述是錯誤的？ (A) AI 可能因缺乏透明度，使用者無法清楚得知資料是否被蒐集與使用 (B) AI 使用資料進行分析時，可能超出原本同意的資料使用範圍

		(C) AI 無法整合不同資料來源進行分析，因此不可能推測特定個人身分 (D) AI 可能透過資料推論與重新識別，造成個人隱私風險
解析		生成式 AI 在處理個人資料時，可能涉及資料蒐集、分析、推論及不同資料來源的交叉比對。若資料使用方式缺乏透明度，或超出原有使用目的與範圍，可能增加個人隱私風險；此外，不同資料來源經交叉分析後，也可能產生推論或重新識別特定個人的情形。因此，認為 AI 無法整合不同資料來源、因而不可能推測特定個人身分的說法並不正確。
7	B	AI 從大量的資料訓練中「模仿」出邏輯推導（Chain of Thought）能力，可以用於協助答覆客戶問題。但若提問者所詢問的事項並沒有涵蓋在該 AI 模型原先的大量訓練資料中，極易產生「幻覺（Hallucination）」而給出不正確答案。為避免或降低產生幻覺的問題，又須考慮算力資源之可及性，在實務上多數採用的是下列哪一項做法？ (A) 增加 AI 模型訓練的頻率，使最新的資料或知識可以及時納入以減低幻覺產生 (B) 收集並整理妥適的問答集或過去機關處理該領域事件之案例，建立機關自有的 RAG（Retrieval Augmented Generation 檢索增強生成）向量資料庫，與大型語言模型（LLM）整合運用 (C) 限制使用者向 AI 提問之內容，避免觸發幻覺的問題 (D) 在 AI 模型產生「幻覺（Hallucination）」尚不能完全解決之前，不導入 AI
解析		生成式 AI 可能因缺乏特定領域資訊或無法取得最新資料而產生幻覺。RAG（檢索增強生成）可將機關既有的問答資料、業務文件或相關案例建立為可檢索的資料庫，使用者提問時先取得相關資料，再提供大型

		語言模型產生回答，使回答更能依據指定資料來源，有助於降低錯誤或虛構內容的風險。因此，B 選項最符合實務上的作法。
8	C	為兼顧高齡者與數位弱勢族群需求，政府機關導入 AI 智慧申辦系統，下列何項作法最合適？ (A) 僅保留線上 AI 申辦服務 (B) 要求所有民眾皆須熟悉 AI 工具操作 (C) 導入智慧語音服務並保留實體窗口等多元服務管道 (D) 完全取消人工客服與臨櫃服務
解析		政府推動智慧行政與 AI 應用的核心目的在於「輔助並優化公共服務」，而非強制替代傳統申辦途徑。高齡者或數位弱勢族群可能因缺乏智慧設備、數位素養不足或閱讀困難，無法順暢使用複雜的線上文字介面，透過「智慧語音服務」降低操作門檻（方便長者用講的申辦），同時「保留實體窗口與人工服務」，兼顧了科技創新與數位包容，為最適當的作法。

參、 認證主題 2：AI 政策法制

2.1. AI 政策與治理指引

本核心內涵旨在引導公務機關在法規合規、風險可控與倫理保障的前提下，負責任地導入 AI 技術，建立公務人員對我國 AI 政策、治理指引、風險分級、資料治理及 AI 導入生命週期管理之理解與法遵意識。

- AI 政策與法制規範

我國 AI 政策與法制規範以《人工智慧基本法》為重要法制基礎，政府推動 AI 應遵循人類自主、隱私保護與資料治理、資安與安全、透明與可解釋、公平與不歧視及問責等原則。

在公務應用層面，並依《行政院及所屬機關（構）使用生成式 AI 參考指引》，規範公務人員使用生成式 AI 的原則與注意事項，包含妥善保護公務資訊、落實人工審查與事實查核，以及遵循資通安全、個人資料保護、著作權及相關資訊使用規範，以兼顧 AI 應用效益與依法行政之要求。公務人員應能理解我國 AI 政策與治理規範，並依業務屬性辨識適用之一般性或專業領域 AI 應用規範，確保 AI 應用符合相關政策與法規要求。

- AI 全生命週期治理與風險管理

公務機關導入 AI 時，應從應用場景評估與專案啟動開始，依序進行資料探索與模型建立、模型迭代與部署，並持續進行風險控制與專案追蹤。依數發部《公部門人工智慧應用參考手冊》，AI 導入可分為「AI 場景評估、AI 專案啟動、資料探索與模型建立、模型迭代與部署、風險控制和專案追蹤」五大階段，並將風險管理、倫理、資料安全、AI 資安及治理架構納入相關管理事項。

公務人員應具備 AI 導入前、中、後之治理意識，包括導入前的風險評估、業務適法性檢核、資料治理、系統選型及採購或委外管理；

導入過程中的權限控管、使用紀錄、人類監督及異常處理；以及導入後的持續監測、改善追蹤與內控檢討，確保 AI 系統持續符合施政目標、資安規範及倫理要求。

- AI 風險管理與持續監測

另依數發部「AI 風險分類框架」，可透過「盤點應用情境、識別風險、評估風險、應對風險」之流程，依 AI 應用情境及可能造成的影響採取適當管理措施。公務人員應能依業務屬性進行初步風險辨識與判斷，並了解 AI 風險分類、監理及相關控管原則，將風險管理落實於採購或招標文件審查、契約條款檢視、政策輔助分析及民眾服務系統導入等實務情境。

AI 系統導入後，亦應持續檢視其準確性、回應穩定性、偏誤情形、資源使用效率、使用者回饋及服務品質等面向，並依評估結果進行必要的調整與改善，以確保 AI 應用持續符合業務需求及相關規範。

- 資料治理與 AI 使用判斷

AI 應用亦應遵循《人工智慧基本法》所揭示之隱私保護與資料治理、資安與安全、透明與可解釋、公平與不歧視及問責等原則，妥善管理資料並促進跨域資料交換與整合。

公務人員應具備資料治理與 AI 使用判斷能力，理解資料之品質、可讀性、可檢索性及適用性，並能判斷公務資料是否適合作為 AI 訓練、測試、檢索或生成服務之資料來源；對涉及機密、未公開公務資訊、個人資料或其他受保護資訊，應能辨識其使用限制，並依相關法規及安全規範進行適當管理，以兼顧資料利用、隱私保護與資訊安全。

在使用生成式 AI 時，業務承辦人應對 AI 產出進行客觀且專業的最終判斷，不得以未經確認的 AI 產出直接作成行政行為或作為公務決策的唯一依據，並應遵循資通安全、個人資料保護、著作權及相關資訊使用規範。

2.1 參考樣題及解析

1	D	聯邦學習（Federated Learning）」在公務機關裡，是一種用來提升 AI 可信度與資料治理合規性的技術，關於它的說明哪一項錯誤？ (A)聯邦學習在公務機關的核心價值，不是「更準」，而是「更可控」 (B) 聯邦學習的本質是：資料不出單位，只交換模型更新，也就是各機關（或各縣市、各系統）各自訓練模型，只上傳參數，不上傳原始資料 (C) 資料留在原機關，可降低外洩風險 (D) 機關須上傳參數，也要上傳原始資料，才能確保完整性
解析		聯邦學習（Federated Learning）是一種分散式機器學習技術，各機關可利用本地資料進行模型訓練，原始資料無須離開原機關，再透過交換模型更新資訊進行模型彙整。其優點在於可降低原始資料集中與跨機關傳輸所衍生的風險。故 D 選項所述「須同時上傳原始資料」與聯邦學習的核心概念相違，為錯誤敘述。
2	D	小美使用生成式 AI 協助製作社群貼文與活動宣傳圖片，雖然 AI 能快速產生內容，但她仍擔心內容可能出現錯誤資訊、不當用語或侵犯他人權益，下列哪一項作法對維護內容品質與合規性最為關鍵？ (A) 僅需於貼文中註明「內容由 AI 生成」即可免除所有法律與倫理責任 (B) 提高提示詞（Prompt）的精準度，即可完全杜絕產出侵權內容的風險 (C) 依賴 AI 工具內建的過濾機制，不需額外耗費人力進行人工複審

		(D) 建立組織內部的合規審查機制，針對法規與智慧財產權進行實質把關
解析		生成式 AI 模型於文字生成時可能產生「幻覺（錯誤資訊）」，於圖像生成時則可能無意間重現特定創作者的圖像風格、商標、肖像或具著作權保護之元素，透過建立專責的內部審查機制（如法律、資安與智財權查核流程），落實人類實質把關，為維護內容品質與合規性最關鍵、最根本的做法。
3	B	機關規劃利用歷年外來公文之人工分文結果資料來訓練 AI 模型，以提升公文自動分類與分文的效率，關於將公務資料作為 AI 訓練來源之運用，下列哪一項敘述最適當？ (A) 所有公務資料皆可直接提供給外部 AI 平台進行訓練 (B) 未公開之機密資訊不得用於 AI 模型訓練，使用個人資料進行訓練前應先進行去識別化並評估是否會造成間接識別之風險 (C) 為追求最佳的模型訓練效果，可適度忽略資料授權與著作權規範 (D) AI 模型之訓練資料因屬內部測試，故不需遵守機關資料治理規範
解析		依據《行政院及所屬機關使用生成式 AI 參考指引》與《個人資料保護法》，公務機關於導入或自行訓練 AI 模型（如公文自動分文分類模型）時，必須遵守嚴格的資安與個資控管，公務資料進行 AI 訓練必須堅守「機密不輸入、個資先去識別化並評估間接識別風險 (B)」之原則；絕不可隨意外傳 (A)、忽略智財權 (C) 或漠視資料治理規範 (D)。

4	B	一般民眾使用生成式 AI 服務時，有時會發現系統雖然能快速給出結果，卻無法清楚說明「為什麼會做出這樣的判斷」，此種情況常被稱為 AI 的「黑箱」問題，請問上述主要反映下列哪一項現象？ (A) 資料隱私與保護的挑戰 (B) 模型的運作和決策過程缺乏可解釋性和透明度 (C) 訓練數據的數量不足以支持模型有效學習 (D) 計算資源不足以進行充分的模型訓練
解析		AI 黑箱問題是指 AI 系統雖能產生判斷或結果，但其內部運作及決策依據不易被使用者理解或說明，因此主要涉及模型的透明度與可解釋性。資料隱私、訓練資料不足及計算資源不足雖亦可能影響 AI 系統的應用，但並非黑箱問題的主要意涵，故 B 選項正確。
5	A	法務助理利用生成式 AI 檢索過去的司法判例，並直接將其引用於訴訟狀中呈交法院，事後法官發現該訴訟狀中所載的判決與案號根本不存在，請問此種實務風險主要是由下列哪一項原因所引起？ (A) AI 產生「幻覺（Hallucination）」現象，虛構了不真實的判例與字號 (B) AI 為了保護當事人隱私，自動將真實判例進行去識別化處理 (C) 內部網路遭受惡意駭客攻擊，導致 AI 的輸出結果遭到即時竄改 (D) AI 系統自動啟用了「同態加密（Homomorphic Encryption）」功能，導致輸出的文字與真實判例不符
解析		大型語言模型（LLM）的運作本質是基於機率模型「預測下一個最可能出現的字詞」，而非真正的資料庫搜尋引擎。如具體的法院判決字號、法規條號或歷

		史文獻時，若模型未搭配外部資料庫進行檢索，虛構出看似極為真實卻完全不存在的案件名稱與字號，此現象即稱為 AI 的「幻覺」，AI 生成看似合理的虛構判例與字號屬於典型的 「AI 幻覺現象 (A)」；這也是公務與法律應用中，要求必須落實 人類實質查核（Fact-checking） 的最主要原因。
6	B	機關在推動「AI 系統成效評估與持續監測」機制時，強調公務人員應對 AI 系統的輸出成效進行日常檢測與品質監控。若承辦同仁在利用內部 AI 工具協助產出標案公告文字後，經由人工監測程序發現 AI 誤植了政府採購法規定的重要日期，請問下列哪一種處置方式最符合人機協作下的輸出成效評估、監測與公務權責劃分？ (A) 僅需針對誤植的日期進行人工修正，其餘由 AI 生成之招標條款與文字不另作實質核對 (B) 由承辦人對全文字句進行實質檢視與確認，並依機關行政程序陳核、經權責主管批核後發布 (C) 由於該工具為機關內部認證之安全 AI，其產出之採購法規判定具備合規性，承辦人應予以尊重直接沿用 (D) 將錯誤回報給機關資訊單位進行 AI 模型修正，並向投標廠商聲明「因系統瑕疵導致公告日期依系統為準」
解析		依據《行政院及所屬機關使用生成式 AI 參考指引》，AI 產出僅為業務輔助草案，不能取代公務員的專業判斷與責任。招標公告涉及採購法與廠商重大權益，承辦人必須對全案進行「完整實質審查」，發現錯誤後對全文字句進行全面校對與實質查核，並循機關行政程序送陳長官核裁，完全落實了「人類在環（Human-in-the-loop）」監督機制與最終行政權責歸屬。
7	BCD	在公務機關導入 AI 服務時，關於訓練資料（Training Data）與推論資料（Inference Data）的處

		理，下列敘述哪些正確？ (A) 為了讓 AI 更了解民眾需求，機關可以未經法律授權或當事人同意，將民眾的陳情內容直接作為開源模型的訓練養分 (B) 當使用外部公有雲 AI 模型（如 ChatGPT API）時，不應輸入涉及國家安全或民眾隱私的推論資料 (C) 訓練模型時，應優先使用經過去識別化或彙整後的統計資料，以降低個資外洩風險 (D) 機關應建立資料盤點機制，區分哪些資料可供內部模型訓練，哪些資料不得用於外網模型訓練
解析		依據《行政院及所屬機關使用生成式 AI 參考指引》，機密性公務資訊與民眾個資嚴禁輸入公有雲模型（防止推論資料被雲端業者留存或用於模型再訓練）；機關必須建立完整的「數據盤點與分級機制」，確保數據流向符合法規與資安邊界，機關應建立數據盤點 (D)，訓練時優先採用去識別化資料 (C)，且使用外部公有雲時嚴禁輸入敏感推論資料 (B)；未授權即直接拿民眾個資進行模型訓練 (A) 為嚴重違規之行為。
8	C	機關導入 AI 輔助公文審查系統，協助同仁初步檢查案件內容與法規適用情形。然而，實務上案件常有資料不完整或語意模糊的挑戰；為克服此類實務限制，該系統必須具備「不確定性處理能力」，其主要原因為何？ (A) 能全面取代人工初審流程，自動駁回所有語意模糊的公文案件 (B) 能大幅縮短結構化欄位的比對時間，加速公文自動分辨的時效 (C) 能在資訊不全或存在歧義的情境下，提升法規適用性的分析與推論能力 (D) 能自動補齊欠缺的公文佐證附件，確保機關審查

		案件的實質合規性
解析		公文審查涉及案件資料、法規及實際情境，當資訊不完整或語意存在歧義時，AI 系統若具備不確定性處理能力，可辨識資訊不足或歧義情形，並在有限資訊下進行適當的分析與推論，協助承辦人判斷法規適用情形。此能力並非取代人工審查或自行補齊缺漏資料，故 C 選項正確。

2.2. AI 法規與資安風險管理

本核心內涵旨在建立公務人員對 AI 應用之資通安全、個人資料保護、機敏資料防護及合規使用之風險辨識與實務判斷能力。公務機關導入及使用 AI 時，應遵循《人工智慧基本法》所揭示之隱私保護與資料治理、資安與安全等原則，以及《個人資料保護法》、《資通安全管理法》等相關規範，並依業務性質及資料敏感程度採取適當之安全與管理措施。

- AI 法遵與公務資料保護

公務人員應了解不同 AI 使用情境之資料安全風險，並能判斷公務資料是否適合提供予 AI 服務使用。依《行政院及所屬機關（構）使用生成式 AI 參考指引》，業務承辦人不得向生成式 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，亦不得以生成式 AI 製作機密文書；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，得依文書或資訊機密等級分級使用。

命題可透過具體情境，評估公務人員對資料使用限制之判斷能力，例如未公開之採購資訊、個人資料、未發布之政策內容、內部簽辦資訊或其他應保密資料，判斷是否適合輸入外部生成式 AI 服務，並辨識可能產生之資料外洩風險。

- 影子 AI 與資安風險防護

公務人員應具備「影子 AI（ShadowAI）」風險意識，了解未經機關核准或未經適當安全評估而使用外部 AI 服務，可能衍生公務資訊、個人資料或其他受保護資訊外洩等風險。機關得依使用生成式 AI 之設備及業務性質，訂定相關使用規範或內控管理措施，並透過教育宣導及適當的技術與管理措施，降低未經授權使用 AI 所產生的資安風險。

- AI 法遵與採購、委外管理

  公務機關導入或採購 AI 系統及服務時，應將相關資安、資料保護及 AI 使用規範納入管理，並依業務性質及風險進行適當之合規檢視。行政院生成式 AI 參考指引亦明定，各機關辦理相關採購時，應要求得標之法人、團體或個人注意該指引，並遵守機關所訂定之相關規範或內控管理措施。

  命題可結合 AI 採購、招標文件、契約條款、委外開發或 AI 服務導入等情境，評估公務人員是否能辨識資料使用、資安責任、使用規範及相關合規要求，並依機關規範進行適當的審查與管理。

- AI 產出查核與安全使用

  使用生成式 AI 時，公務人員應對 AI 產出進行客觀且專業的最終判斷，不可完全信任 AI 產出，亦不得以未經確認的內容直接作成行政行為或作為公務決策的唯一依據；使用 AI 執行業務或提供服務時，並應適當揭露，且遵守資通安全、個人資料保護、著作權及相關資訊使用規定。

  因此，命題可透過 AI 產生錯誤資訊、資料外洩、著作權疑慮或不當使用公務資料等情境，評估公務人員是否能辨識風險並採取適當處理方式。

2.2 參考樣題及解析

1	A	機關導入 AI 協助篩選異常採購案件，為避免系統誤判，同仁特別多次測試以檢查「系統判定為異常案件中，實際確實異常的比例」，上述作法不包括評估下列哪一項能力？ (A) AI 處理大量文件之速度 (B) AI 分析模型篩選結果之穩定程度 (C) AI 自動分類資料之精準度 (D) AI 篩選結果之準確性
解析		題目所述「系統判定為異常案件中，實際確實異常的比例」，主要是在檢視 AI 分類與篩選結果的準確程度，透過多次測試亦可觀察模型結果的穩定性。至於 AI 處理大量文件所需的時間，屬於系統處理效率或效能的評估，與題目所描述的測試目的不同，故 A 選項正確。
2	C	公司的資訊安全長準備升級內部 AI 系統的資訊安全防護，下列哪一項做法「最不符合」資安防護原則，甚至可能帶來資安風險？ (A) 定期安排資安團隊對 AI 模型進行紅隊演練，找出漏洞 (B) 導入去識別化技術，確保員工查詢客戶資料時，顯示或列印出資料中其帶有個人資料部分會被系統自動遮蔽 (C) 透過尚未經安全檢核之 API 或 MCP 等連結串接外部網路資訊，來提升運算性能 (D) 對全體員工進行社交工程與 AI 資安培訓，教導如何辨識新型態的 AI 詐騙信件

解析		MCP（Model Context Protocol）或第三方 API 是 AI 系統用來連結外部網路數據、工具與資料庫的關鍵橋梁。若使用未經安全檢核或驗證的 API/MCP 串接，等於在企業與 AI 系統的資安防線開了後門，極易遭到提示詞注入（Prompt Injection）、中繼攻擊、未授權存取或資料外洩等嚴重威脅，未經資安審查與檢核的 API/MCP 介面屬於重大資安漏洞，會將內部系統暴露於不可控的外網風險中，完全違背資安防護原則。
3	A	機關規劃利用 AI 系統分析大量的民眾陳情與意見資料，其中部分資料已具備明確的分類標籤（Label），部分則尚未分類，關於不同 AI 學習方式之敘述，下列哪一項最正確？ (A) 監督式學習通常需要使用已標籤（標記）的資料進行訓練 (B) 非監督式學習在運作時完全不需要輸入任何資料 (C) 監督式學習因技術限制，無法應用於政府的行政分析 (D) 非監督式學習具備完美準確度，可直接取代所有人工判讀
解析		監督式學習（Supervised Learning）：運作機制：好比「有標準答案（Ans）的練習題」。模型需要輸入「資料（Data）+ 對應標籤（Label）」進行訓練（例如：輸入陳情文本，並給予標籤「路面坑洞」）。模型透過學習資料與標籤之間的關聯，未來便能對未知的資料進行分類或預測。 非監督式學習（Unsupervised Learning）：運作機制：好比「沒有解答，讓 AI 自己抓規律」。模型僅輸入未標籤的資料，自動透過演算法尋找資料內部的相似度或結構進行分群（Clustering）或降維。

4	B	一般民眾在網路上觀看影片或社群內容時，可能會遇到利用 AI 製作的深偽（Deepfake）影片，造成假訊息或冒用他人身分等問題，下列哪一項最適合民眾用來防範生成式 AI 產生的深偽技術？ (A) 增強機關或個人隱私資料保護 (B) 提高警覺，遇有可能是深偽（Deepfake）造成之假訊息或身分冒用之徵兆，應避免回應，或先行向當事人或機關查證 (C) 研發深偽內容偵測技術 (D) 創建更多生成式 AI 應用
解析		生成式 AI 可被用於製作高度逼真的深偽內容，可能造成假訊息傳播或身分冒用。一般民眾面對疑似深偽或可疑訊息時，應提高警覺，避免立即回應或轉傳，並透過當事人或相關機關等可靠管道進行查證，以降低受騙或遭冒用的風險，故 B 選項正確。
5	A	在使用生成式 AI 時，下列哪一種行為可能涉及侵犯智慧財產權？ (A) 用 AI 生成新的藝術作品宣稱是自己的創作，並用於參加競賽 (B) 用 AI 生成隨機數據集進行訓練 (C) 使用 AI 來生成數學公式 (D) 用 AI 生成個人筆記作為學習輔助工具
解析		隱瞞 AI 創作並宣稱為個人作品：將純 AI 產製的作品謊稱「自己親手創作」並參加競賽，除了違反競賽規則、涉及民法與刑法上的欺詐行為外；若 AI 於生成過程中無意間重現了他人受保護作品之「實質相似（Substantial Similarity）」特徵，亦會直接引發侵害他人著作權與肖像權的法律訴訟風險。

6	B	機關規劃對外公開施政成果照片，為防止民眾之個人隱私外洩，資訊人員先行利用 AI 工具自動將照片中的車牌與民眾面部進行模糊化處理，請問上述作法最主要的目的為何？ (A) 提高照片之影像解析度與視覺品質 (B) 落實個人資料保護與去識別化規範 (C) 增加 AI 模型之後續訓練資料量 (D) 提升照片在社群媒體的演算法曝光率
解析		將可識別特定自然人與車輛的特徵進行模糊化，最主要的目的就是落實個資法要求與去識別化保護。
7	C	機關委外建置 AI 智慧客服系統，於招標需求書中特別明文規範廠商「不得將機關資料用於其他 AI 模型之訓練」，且「須完整保留系統操作日誌（Log）以供後續查核」，請問上述規範最主要的目的為何？ (A) 提升 AI 回應內容之即時性與穩定性 (B) 強化 AI 系統跨平台資料整合能力 (C) 落實機密及隱私資料保護與資安管理要求 (D) 增加資料再利用與模型訓練彈性
解析		機關委外導入 AI 系統時，應注意機關資料之使用範圍及安全管理。要求廠商不得將機關資料用於其他 AI 模型訓練，可避免資料遭不當利用；保留系統操作日誌則有助於後續追蹤、稽核及資安事件查核。上述措施主要在落實機關資料保護及資安管理要求，故 C 選項正確。
8	B	機關同仁利用生成式 AI 協助撰寫回覆民眾的正式公文函稿，但 AI 所引用的法規條文與現行規定並不一致，若承辦人未經核對便直接將該公文發出，就可能產生法律爭議；關於下列情境哪一項做法不正確？ (A) 發文前，承辦人須詳實檢視 AI 文稿及引用之法律

		依據 (B) 串聯內外部模型以增強模型解讀法規能力，就可以避免法律爭議 (C) AI 造成錯誤影響民眾權益，承辦人須負法律或行政責任 (D) 回覆民眾的函稿若內容涉及當事人隱私或機密資料不宜使用 AI 模型
解析		生成式 AI 可協助撰寫公文，但其產出內容仍可能存在法規引用錯誤或與現行規定不一致之情形，因此承辦人應於發文前確認 AI 產出及其引用之法律依據。即使透過多個 AI 模型進行交叉處理，也無法保證內容完全正確或避免法律爭議，仍須由承辦人進行查證與專業判斷。涉及個人資料、公務應保密或其他受保護資訊時，亦應依相關規範妥善處理，故 B 選項為不正確的做法。
9	C	機關同仁在撰寫採購案件的招標需求書時，利用生成式 AI 協助整理與參考過往的標案內容，請問下列哪一種資料，基於資安與保密原則，「最不適宜」輸入至外部 AI 工具中？ (A) 已於政府電子採購網公開之招標公告 (B) 已對外公開之履約規範 (C) 尚未公開之採購預估底價與委員評選意見 (D) 已發布並結案之機關新聞稿
解析		依據《政府採購法》第 34 條規範，機關辦理採購，底價於決標前應予保密；評選委員會之評選意見與名單在特定階段亦屬應保密事項。若於招標前洩漏底價或評選細節，將嚴重破壞採購公平性並涉犯公務員洩漏國防以外秘密罪，故 C 選項資訊屬於法定應嚴格保密之公務機密，一旦輸入外部 AI 工具即構成重大資安漏洞與洩密違規，故為最不適宜輸入之資料。

2.3. AI 應用下的倫理準則與智慧財產實務

本核心內涵旨在建立公務人員於 AI 應用過程中，對人工監督、AI 產出可信度、資訊揭露、行政倫理及智慧財產權之理解與實務判斷能力，確保 AI 應用符合人類自主、透明與可解釋、公平與不歧視及問責等原則。

- AI 倫理與公務問責

公務機關使用 AI 協助執行業務或提供服務時，AI 應作為業務輔助工具，不得取代業務承辦人的自主判斷。依行政院及所屬機關（構）使用生成式 AI 參考指引，AI 產出資訊應由業務承辦人進行客觀且專業的最終判斷，不得以未經確認的 AI 產出直接作成行政行為，或作為公務決策的唯一依據。使用 AI 執行業務或提供服務時，並應適當揭露。

公務人員應具備 AI 產出查證及風險辨識能力，對可能出現的錯誤資訊、偏誤、幻覺或資訊過時等情形進行確認，避免未經查證即採用 AI 產出。涉及民眾權益之業務，更應維持適當的人類監督及專業判斷，由具權責之人員進行最終判斷並承擔相應責任，以落實依法行政與公共問責。

- 智慧財產權與第三方權利

公務機關使用生成式 AI 產製文字、圖片、影音或其他內容時，應遵守著作權及相關資訊使用規定，並注意侵害他人智慧財產權及人格權的可能性。行政院生成式 AI 參考指引亦明定，使用生成式 AI 應遵守著作權及相關資訊使用規定。

依經濟部智慧財產局目前說明，AI 生成內容是否受到著作權保護，應視創作過程中是否有自然人實際的創意投入而定；若僅將 AI 作為輔助工具，且有人類實際創意投入，完成的創作成果仍可能受到著作權保護；若完全由 AI 獨立生成、沒有自然人實際創意投入，則該生成內容無法享有著作權。

另 AI 生成內容仍可能與他人原有著作產生實質近似，則不能因內容是 AI 生成，就認定沒有著作權風險。使用 AI 產出內容前，應注意來源、使用規範及相關授權條件；如涉及第三方著作或其他權利，應進行適當確認，以降低侵權風險，智慧財產局亦指出，AI 生成內容是否侵害他人著作權，仍須依個案具體事實判斷。

- AI 產出透明與可信度

《人工智慧基本法》將透明與可解釋、問責及人類自主列為政府推動 AI 的重要原則，並要求對 AI 產出內容進行適當的資訊揭露或標記。

因此公務人員應能辨識 AI 產出的可信度與限制，對重要資訊進行查證，不應因 AI 產出內容完整或表達流暢，即視為正確，行政院生成式 AI 參考指引亦明確指出，生成式 AI 產出可能真偽難辨或產生不存在的資訊，使用者應進行客觀且專業的判斷。

使用 AI 執行業務或提供服務時，應依相關規範適當揭露 AI 的使用情形，使使用者能了解 AI 於業務中的參與程度，提升 AI 應用的透明度與可信度。

2.3 參考樣題及解析

1	B	人工智慧在文學上的表現，不只可以應用於聊天或是寫詩而已，這項技術已經可以應用在產業界了，尤其是新聞界。關於人工智慧的技術，在新聞稿的撰寫方面，下列敘述哪一項錯誤？ (A) 通過分析整理數據，並將訊息放到制式的新聞模板裡 (B) 人工智慧能在短短五分鐘內就生產了大量的新聞，每一篇都能使用 (C) 目前人工智慧可協助新聞稿撰寫與資料整理 (D) 人工智慧生產的新聞稿仍須經專業新聞人員審視或查證方可發布
解析		雖 AI 生成速度極快，但由於生成式 AI 存在「幻覺（Hallucination，虛構事實）」、邏輯不連貫、語意偏差或去脈絡化的致命風險，絕非「產出的每一篇都能直接使用」。新聞產出極度強調客觀真實性與法律責任，因此必須經過人類記者與編輯的嚴格查核、校對與修改（Human Editing）才能發布，否則會引發假新聞與新聞倫理風暴。
2	C	考量到 AI 倫理的不傷害原則，下列哪一種情緒辨識的應用情境應受限制？ (A) 偵測幼稚園兒童是否有不安或慌張 (B) 測謊時偵測測試者是否有說謊 (C) 透過悲傷表情預測被測試者的過往創傷記憶 (D) 在餐廳裡偵測客戶的餐點滿意度
解析		不傷害原則要求 AI 技術的設計與應用不得對個體的身體、心理、尊嚴或權益造成實質或潛在的傷

		害。特別是在心理健康與情緒辨識領域，系統不應主動誘發、揭露或挖掘個人的深層心理創傷，強行將表面情緒推論至深層創傷記憶，缺乏科學精確性，且極易引發個體心理二次傷害與隱私極致侵犯，完全違背倫理上的不傷害原則。
3	ABD	從 AI 倫理中關於「社會福祉」與「勞動轉型」的觀點來看，關於製造業人機關係的敘述，下列哪些正確？ (A) 人工智慧可能會取代部分重複性高且標準化的工作內容 (B) 人工智慧導入後，可能帶動 AI 維運、資料分析與智慧製造等新興職務需求 (C) 人工智慧系統因需投入設備與技術成本，長期而言通常無法提升整體生產效益 (D) 人工智慧能協助不良品檢測，進而提升整體作業效率
解析		AI 導入製造業可能同時帶來工作內容轉變、新興職務需求及生產效率提升等影響。部分重複性高且標準化的工作可能受到自動化影響，但也可能增加 AI 維運、資料分析及智慧製造等相關職務需求；此外，AI 亦可應用於不良品檢測，協助提升作業效率。因此 A、B、D 正確，C 的說法過於絕對，故不正確。
4	B	醫院導入 AI 輔助影像診斷系統，在測試資料中的判讀結果表現非常準確，但實際上線後面對新的病患影像資料時，卻出現誤判情況經分析後發現，模型過度記憶訓練資料特徵，無法有效適應未曾看過的新資料，此現象稱為「過度擬合（Overfitting）」，請問可能造成下列哪一項影響？

		(A) 提高模型的預測精準度 (B) 降低模型對未見資料的泛化（Generalization）能力，導致診斷錯誤 (C) 增強模型對新型疾病的應對能力 (D) 提高資料處理速度
解析		過度擬合（Overfitting）是指模型過度學習訓練資料的特徵，導致模型雖在訓練資料上表現良好，但面對未曾看過的新資料時，泛化能力反而降低。題目中的 AI 影像診斷系統即屬此情況，可能因無法有效適應新的病患影像而產生誤判，故 B 選項正確。
5	C	某出版社舉辦「AI 文學創作論壇」，探討生成式 AI 在文字創作上的應用。依據我國著作權法對「著作」之定義，必須屬於「人類精神上之創作」。關於生成式 AI 之產出是否符合此法規定義與著作權保護要件，下列敘述何者錯誤？ (A) 僅依賴簡單提示詞由 AI 獨立生成的詩作，不受著作權法保護 (B) 若人類將 AI 產出之初稿進行大幅度改寫與潤飾，該最終成品可享有著作權 (C) AI 獨立產出的作品，可直接視為該 AI 軟體開發者的精神創作 (D) 若出資聘請專家利用 AI 創作，其著作權的歸屬應依雙方契約約定
解析		依據我國《著作權法》第 3 條與智慧財產局（TIPO）之解釋，「著作」必須係指「人類精神上之創作」，生成式 AI 所產出的內容是否受著作權保護，應視人類是否具有實際創意投入而定。若僅以簡單提示詞由 AI 獨立生成，通常欠缺人類創作投入；若人類對 AI 初稿進行具有創意之改寫、編排或潤飾，則人類創作部分可能受到著作權保護。至於 AI 獨立產出的作品，不能僅因 AI 軟體由特

		定開發者製作，即認定該開發者具有該作品之人類精神創作。因此 C 選項錯誤。
6	D	在人工智慧的倫理討論中，下列哪一項非相關議題？ (A) 個人隱私與機敏資料保護 (B) 透明度與可解釋性（Transparency & Explainability） (C) 公平性與消除偏見（Fairness & Bias Elimination） (D) 運算晶片的降溫技術
解析		D 選項為硬體工程與散熱技術領域，非屬探討社會價值與道德規範之 AI 倫理議題。
7	B	機關利用生成式 AI 協助產製政策宣導短片與新聞稿，在正式發布前，業務單位主動查核所引用素材的創用 CC（Creative Commons）授權條款，並於文宣結尾明確揭露所使用的 AI 輔助工具，請問上述作法最主要的目的為何？ (A) 大幅增加影片與圖片之影像解析度 (B) 落實智慧財產權與授權規範 (C) 有效提升後台 AI 模型的運算速度 (D) 降低行政流程與政策推動的透明度
解析		公務機關在製作文宣或影片時，若引用外部圖片、音樂或素材，必須嚴格遵守《著作權法》與創用 CC 授權條款（如姓名標示、非商業性、相同方式分享等），確保使用素材之合法性，避免侵害他人智慧財產權。
8	ABD	運用生成式 AI 協助執行業務或提供服務，有助於行政效率之提升，但為保持執行公務之機密及專業性，行政院訂有「行政院及所屬機關（構）使用生成式 AI 參考指引」，以利遵循。下列哪些敘述符合

		參考指引之規定？ (A) 生成式 AI 產出之資訊，須由業務承辦人就其風險進行客觀且專業之最終判斷，不應取代業務承辦人之自主思維 (B) 不得以未經確認之產出內容直接作成行政行為或作為公務決策之唯一依據 (C) 機關所辦採購標案，其得標之法人、團體或個人因非屬公務體系，可參考但無須遵守參考指引訂定之規範 (D) 使用生成式 AI 作為執行業務或生成文件之輔助工具時，應於產出之文件揭露
解析		行政院及所屬機關（構）使用生成式 AI 參考指引強調，生成式 AI 僅作為業務輔助工具，AI 產出資訊應由業務承辦人進行客觀且專業的最終判斷，不得以未經確認之產出直接作成行政行為或作為公務決策的唯一依據；機關辦理相關採購時，亦應要求得標之法人、團體或個人注意並遵守機關所訂定之相關規範或內控措施。此外，使用生成式 AI 執行業務或提供服務時，應依規範適當揭露 AI 的使用情形。因此 A、B、D 正確，C 錯誤。
9	C	有關公務情境中使用生成式 AI 的使用方法，下列何者有誤？ (A) 利用 AI 擬定公文草案，隨後親自對照現行法規，逐條修正不正確的法規字號 (B) 利用 AI 撰寫致詞稿發想靈感，隨後依據長官語氣與需求進行大幅度改寫 (C) 利用 AI 查詢相關司法判例，因相信大型語言模型已學習過大量資料，未經核對便直接引用至專案報告 (D) 利用 AI 為宣導活動發想多個企劃主題，並在會議中與同仁共同評估其可行性
解析		依據《行政院及所屬機關使用生成式 AI 參考指

	引》，AI 產出僅能作為輔助素材或初稿靈感，公務員必須對最終產出的所有事實、法規與判例進行親自對照與核對，並負起最終的行政與法律責任，C 選項未經核對即直接引用極易將 AI 虛構之「假判例」寫入公務報告，此舉不僅違反公務查核義務，更構成重大決策與法律風險。

肆、 參考資料

編號	名稱	來源
1	公部門人工智慧應用參考手冊	https://www-api.moda.gov.tw/File/Get/moda/zh-tw/WwHCroVhwWy52dw
2	人工智慧基本法	https://reurl.cc/vG6gLk
3	營業秘密法	https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=J0080028
4	臺灣 AI 行動計畫 2.0	https://digi.nstc.gov.tw/File/7C71629D702E2D89
5	個人資料保護法	https://law.moj.gov.tw/LawClass/LawAll.aspx?PCode=I0050021
6	專利法	https://law.moj.gov.tw/LawClass/LawAll.aspx?pcode=J0070007
7	人工智慧風險分類框架	https://www-api.moda.gov.tw/File/Get/moda/zh-tw/UieEfux5Yv0iJuR
8	AI 產品與系統評測	https://www.aiec.org.tw/web/guest/service
9	資安院 OpenClaw 與 NemoClaw 資安檢測報告 B	https://reurl.cc/RRybYx
10	行政院及所屬機關（構）使用生成式 AI 參考指引	https://www.nstc.gov.tw/folksonomy/list/c79bf57b-dc94-4aff-8d14-3262b5559cfc?l=ch
11	e 等公務學園+學習平台	人工智慧應用趨勢與技術挑戰 https://elearn.hrd.gov.tw/info/10042450
		政府 AI 政策指引方向 https://elearn.hrd.gov.tw/info/10042886
		AI 倫理與治理基礎 https://elearn.hrd.gov.tw/info/10042888
		生成式 AI 基礎入門 https://elearn.hrd.gov.tw/info/10042887
		AI 公務與創新實務 https://elearn.hrd.gov.tw/info/10042892
		AI 驅動政府數位轉型與智慧治理實踐 https://elearn.hrd.gov.tw/info/10046391

V2026.01.0901

認證網站

CSF 財團法人電腦技能基金

採認編號：TAI-01-C-N