

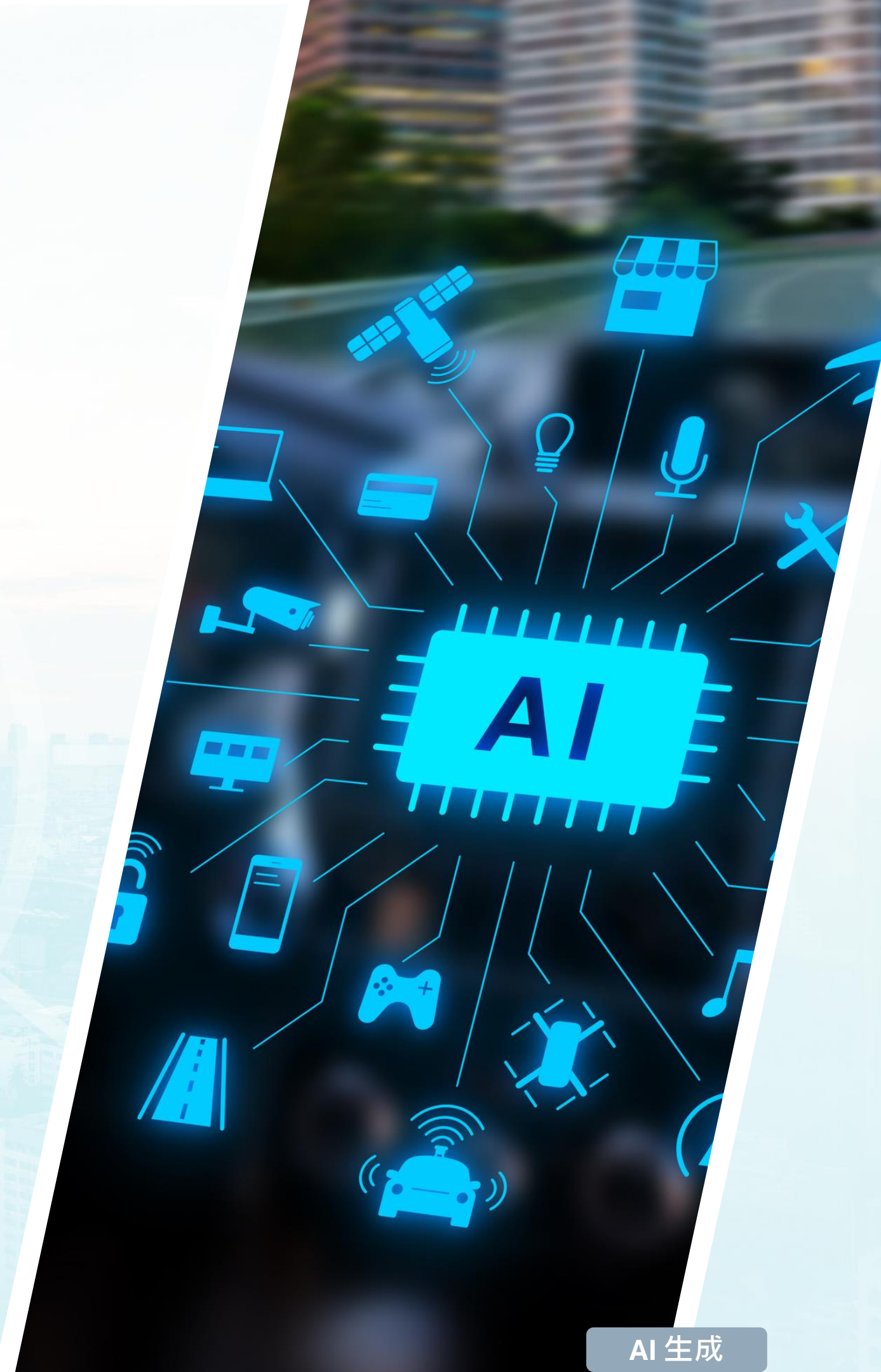


中華人民共和國香港特別行政區政府
數字政策辦公室

「保障個人資料 人工智能沙盒」簡介會

人工智能治理實踐建議： 平衡人工智能發展與責任

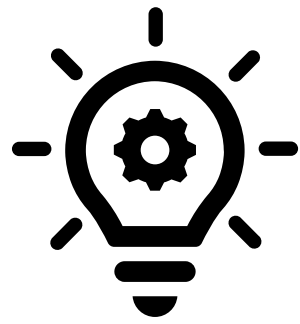
2026年8月28日



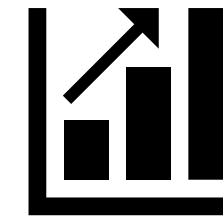
AI 生成

人工智能(AI)在教育中應用的益處

對教師的助益



提高效率

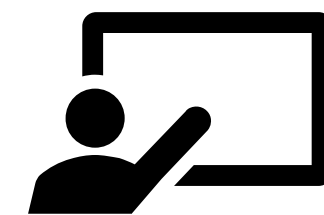


數據驅動教學決策

對學生的助益

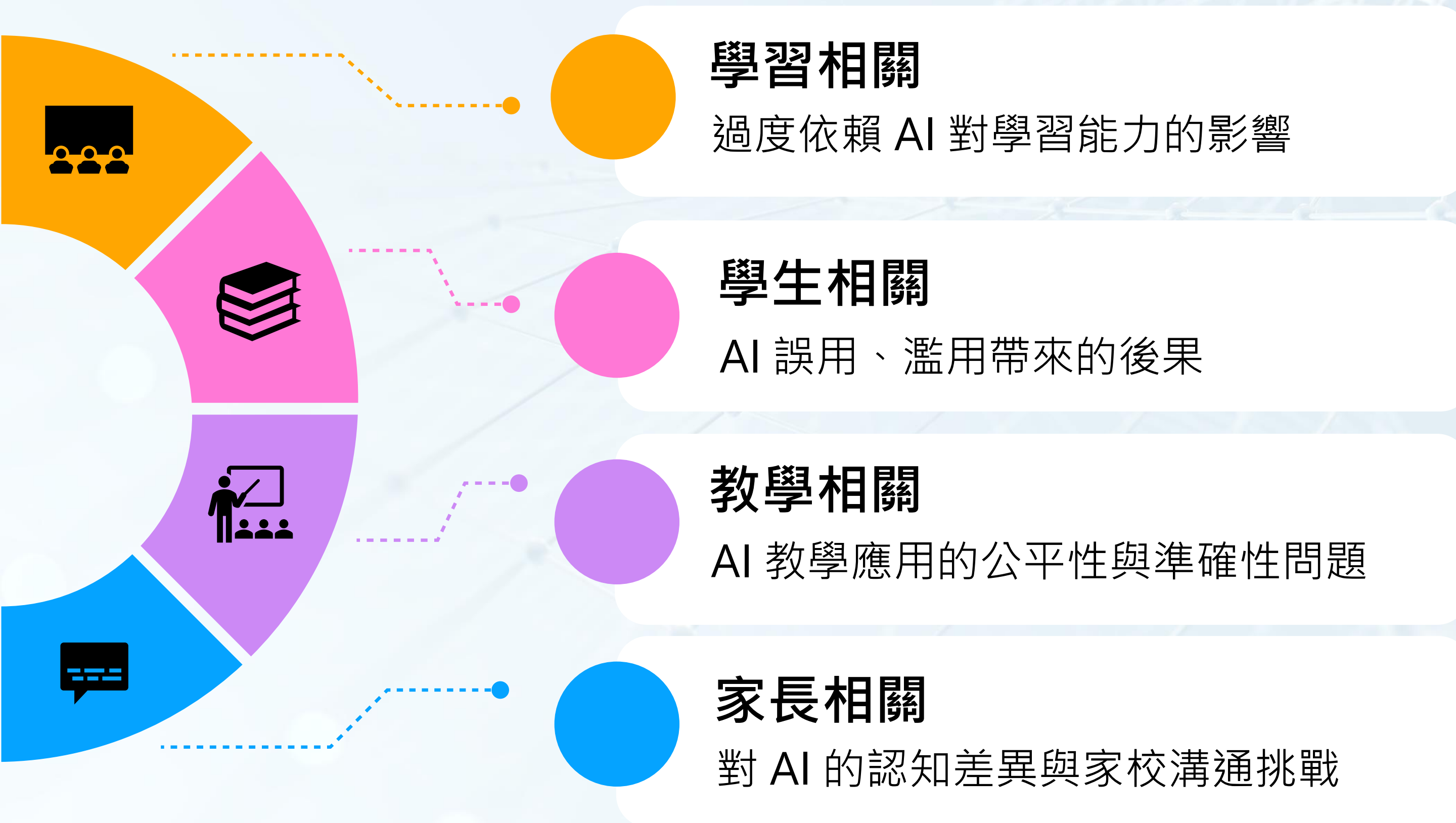


個人化學習體驗



提升學習參與度

教育相關的 AI 道德議題



AI 治理

“AI發展須以安全為舵，以應用為帆”

鼓勵人工智能技術的創新應用

減低人工智能開發和部署
的相關風險

推動人工智能在香港的普及發展

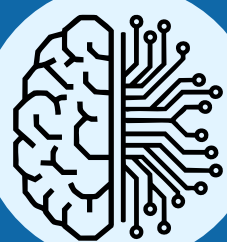
數字政策辦公室的角色

核心工作包括：

制訂指引框架



推動香港
生成式 AI 技術
研發與負責任
應用



支援政府部門
安全應用 AI



支援推廣
AI 科普

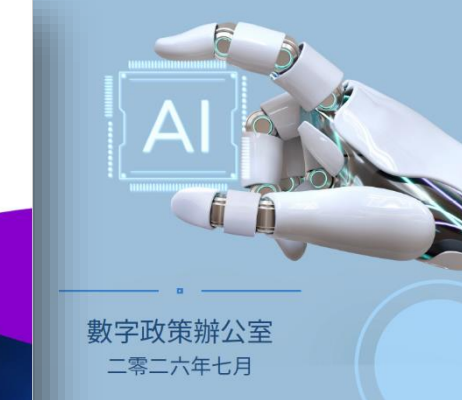


《人工智能道德框架》
《香港生成式人工智能技術及應用指引》
《政府部門人工智能應用指南》



政府部門
人工智能應用指南

人工智能推動政務服務轉型



相關指引



律政司已成立跨部門工作小組，
檢視支持更廣泛應用AI所需的法律配套

《中小學數字 教育發展藍圖》



香港個人資料私隱專員公署
Office of the Privacy Commissioner
for Personal Data, Hong Kong



《開發及使用人工智能道德標準指引》
《人工智能(AI): 個人資料保障模範框架》
《應對濫用人工智能深度偽造技術：
給學校及家長的錦囊》



《香港司法機構法官及司法
人員和支援人員使用生成式
人工智能的指引》



HONG KONG MONETARY AUTHORITY
香港金融管理局

《應用人工智能的高層次原則》
《應用人工智能監察可疑活動》
《應用生成式人工智能的消費者保障》



《生成式AI語言模型的使用》
《加強網絡安全措施
應對人工智能驅動的網絡攻擊
引發的新興風險》

財經事務及庫務局

有關在金融市場負責任地
應用人工智能的政策宣言

香港 AI 治理的核心原則

基於《人工智能道德框架》的 12 項原則



「三道防線」治理架構

資訊科技委員會 / 首席資訊官

- 負責審核、建議及監管高風險AI應用
- 可選擇由人工智能道德委員會(外部顧問)提供支援

第三道防線

項目保證小組 / 項目督導委員會

- 進行獨立檢視及批准

第二道防線

項目團隊

- 負責開發AI應用、風險評估，以及AI評估

第一道防線

校園適用的「三道防線」：

第三道防線（高層監督）
學校管理層

第二道防線（審核把關）
學校行政團隊 / 課程發展組

第一道防線（一線執行）
任課教師 / 資訊科技組



工具一：AI 風險分級表

風險層級	定義	治理策略	校園場景示例	處理規則
 不可接受風險	對社會構成生存性威脅的系統 (例如危害和影響市民安全的用途、潛意識操控)	● 全面禁止 ● 涉及開發或部署的行為須承擔法律責任	用 AI 進行深偽 (Deepfake) 生成 犯罪 / 危害安全的內容	● 絕對禁止 (零容忍) ● 違規者須按相關 法例嚴肅處理
 高風險	關鍵性基礎設施系統 (例如醫療診斷系統、自動駕駛技術)	● 必須進行合規性評估 ● 必須設有人類參與的監督機制 ● 實時監測和持續監控	用 AI 直接學期評核結果、升班留班判定	● 必須由教師進行最終人工審核及把關 ● 受校內機制嚴格監管
 有限風險	帶來中等社會影響的系統 (例如人才招聘系統)	● 履行資訊透明的責任 ● 提供用戶選擇退出的機制 ● 每年定期進行合規審計	AI輔助批改作業、AI 個人化學習建議、AI 輔導答疑	● 必須向師生保持資訊透明 ● 提供非 AI 的替代方案 (退出機制) ● 定期檢視其學與教效能
 低風險	屬於最低風險類別的應用 (例如垃圾郵件過濾系統、創意設計工具)	● 企業自我認證	用 AI 整理筆記、潤色文本、生成備課構思	● 使用者須自行核實內容的真偽及準確性 ● 確保不侵犯知識產權

工具二：生成內容標示

AI 生成

例子

1. 教師準備的教材 / 試卷 / 通告

頁尾標註：

- 「本內容由 AI 輔助生成，並已由教師審核及校對。」

2. 學生作業 / 專題研習報告

標註：

- 「本作業使用 AI 協助資料搜集、文法檢查或語句潤飾。」
- 「本報告使用 AI 協助構思，所有 AI 生成內容均經作者核實、修改，並加入個人分析後完成。」

3. 學校對外發布的宣傳內容

在顯眼位置標註「由 AI 生成」或「由 AI 輔助生成」

什麼時候要標示？

為確保透明度、信賴責任框架和安全，服務使用者應以浮水印、標籤、元數據、數碼簽署或其他方式，明確說明是否使用了生成式人工智能參與內容生成或決策，並且必須承擔其道德和法律影響的責任。

工具三：教學場景提示詞撰寫指南

Prompt Writing

1

設定角色

2

明確目標

3

交代場景

4

規範格式

5

設定限制

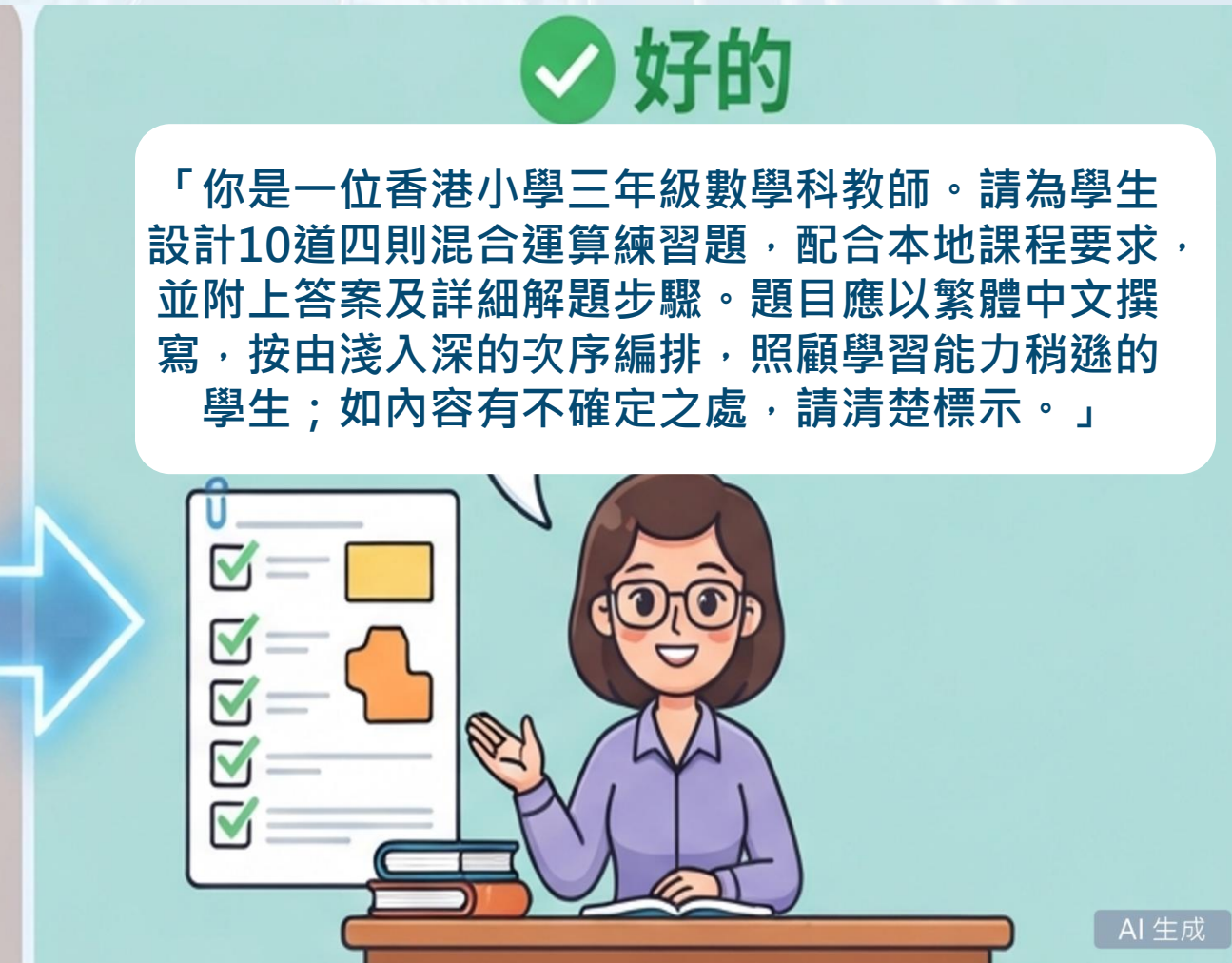
✗ 不好的

「出幾道數學題」



✓ 好的

「你是一位香港小學三年級數學科教師。請為學生設計10道四則混合運算練習題，配合本地課程要求，並附上答案及詳細解題步驟。題目應以繁體中文撰寫，按由淺入深的次序編排，照顧學習能力稍遜的學生；如內容有不確定之處，請清楚標示。」



AI 生成

工具四：AI 生命週期

學校引入AI工具或推行AI項目時，可在以下六個階段持續進行風險評估及治理：



項目策略

為什麼要用這個 AI？
真的能幫助教學嗎？



項目規劃

會用到哪些數據？
可能有什麼風險？



生態環境準備

供應商可靠嗎？符合香港
《私隱條例》要求嗎？



採購與開發

數據會外洩嗎？
結果會有偏見嗎？



系統部署

上線前測試了嗎？
老師學生都會用嗎？



運作與監察

定期檢查有沒有問題，
數據和模型會更新嗎？

工具五： 教師使用 AI 的道德檢視清單

使用任何 AI 工具前，先問自己 5 個問題，有任何一個不確定就先暫停：



1



數據私隱

這個工具會收集學生的個人資料嗎？資料會否用於其他用途？

2



可靠及安全

生成內容會否有錯誤、偏見或不適合學生的內容？

3



公平

不同背景或能力的學生使用時，會否受到不公平對待？

4



人為監督

涉及學生評核、獎懲等重要決定，是否由教師作最終判斷？

5



透明度

工具會否向學生/家長說明此處使用了AI？

應用AI的支援

AI 應用指引



<https://www.smartlab.gov.hk/tc/ai-solution#tabGuideline>

培訓與工作坊



沙盒計劃



AI 科普教育



課外活動例子



沙盒路演

2026年10月6日

下午2時30分至5時

數碼港第5期2樓
「多功能會議廳」

網上報名



港話通



查詢：
ai_sandbox_enquiry@digitalpolicy.gov.hk

AI 生成

謝謝

