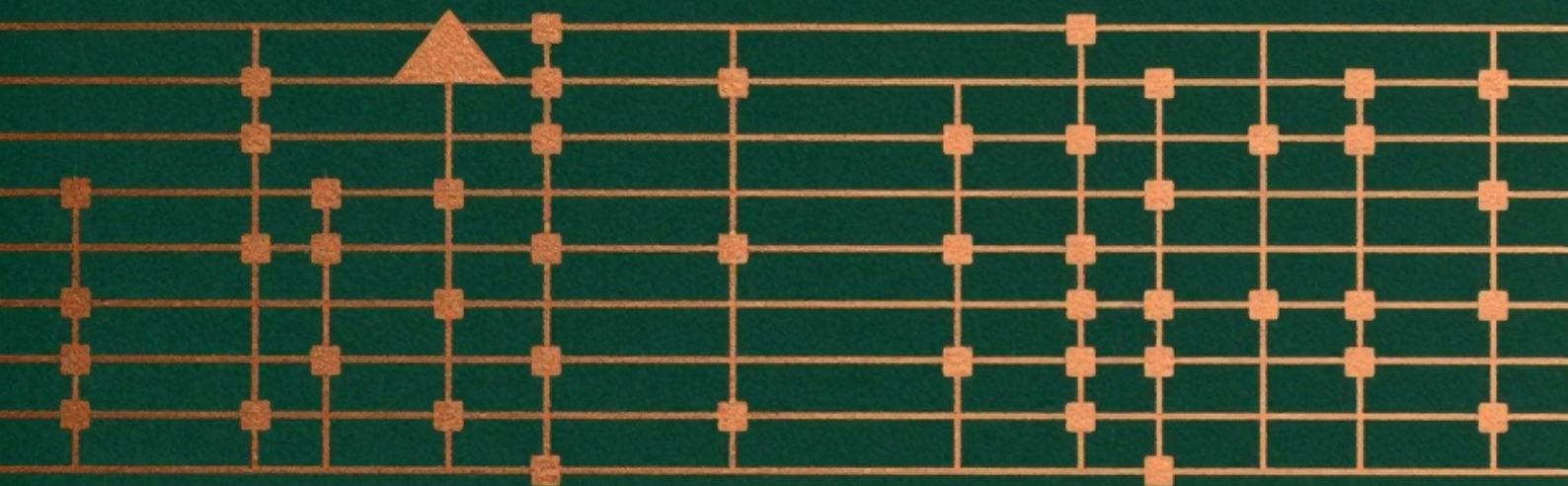


避風港（香港）諮詢有限公司

AI 治理 實務手冊



由系統清冊到董事會一頁摘要

2026 年版 · 免費派發

AI 治理的證據，實際上是什麼

這本手冊說明一件很具體的事：當客戶的採購部門、外聘審計師或董事會問起人工智能，你手上應該已經有哪五份文件，每一份裡面要寫什麼，以及由誰負責。

我們在香港及澳門協助企業準備這些文件。以下內容是我們在項目中實際使用的工作框架，歡迎直接採用。

一、為什麼現在要處理

在香港，AI 治理不是單一條例的要求，而是幾條線同時推進。

- 個人資料私隱專員公署（PCPD）就先進技術及生成式人工智能發出指引。這些指引屬自願性質，但在投訴處理與調查中會被引用。
- 《保護關鍵基礎設施（電腦系統）條例》（第 653 章）自 2026 年 1 月 1 日起生效，指定營運者須履行法定責任。
- 香港金融管理局與證券及期貨事務監察委員會已就人工智能的使用發出通函，屬監管期望。
- ISO/IEC 42001:2023 是可認證的人工智能管理系統標準，正逐步出現在商業合約之中。
- 若集團在歐盟設有業務或用戶，歐盟《人工智能法案》按風險分級分階段實施。

不過對大多數香港企業來說，最先到達的不是監管機構，而是客戶的採購問卷。問卷往往要求你在數個工作天內，說明使用了哪些人工智能系統、處理哪些個人資料，以及誰負責監督。

重點

監管要求通常以年計推進，採購問卷卻以日計到達。把文件準備好，回應時間由數週縮短到一天。

二、五項文件

以下五項構成一套最小可行、且經得起查核的 AI 治理紀錄。它們有先後次序：清冊是基礎，其餘四項都建立在它之上。

01 AI 系統清冊

這是其餘四項的基礎。沒有清冊，風險評估無從開始，問卷也無法回覆。

- 系統或工具名稱、供應商與版本
- 用途，以及它所支援的業務流程
- 具名負責人（職位而非只有部門）
- 處理的資料類別，並標明是否包含個人資料
- 資料儲存與處理的位置
- 使用者範圍與授權方式
- 上線日期與最近覆核日期
- 風險分級

常見問題

清冊最常見的漏洞，是只記錄資訊部門核准的工具，忽略了同事自行訂閱的服務。業界稱之為影子 AI，而它往往正是問卷最關心的一項。

02 風險與影響評估

逐項評估每個 AI 系統帶來的風險，同時涵蓋對個人的影響，以及對機構的風險。兩者都要寫，因為監管機構與客戶關注的重點不同。

- 風險情境（什麼情況下會出問題）
- 發生可能性與影響程度
- 受影響的個人類別與人數級別
- 現有的緩減措施
- 剩餘風險，以及由誰接受
- 覆核日期與觸發重新評估的條件

常見問題

只做技術風險（例如模型準確度），而忽略對個人的影響，正是監管機構最關注的缺口。

03 使用政策與人工覆核

政策要能回答兩個問題：同事可以用什麼，以及在什麼情況下必須由人決定。

- 可接受的使用範圍與禁用用途
- 核准工具清單，並註明版本與用途
- 輸入資料的規則，例如不得輸入客戶個人資料
- 輸出必須經人工覆核的情境
- 對外披露要求，例如是否需要告知客戶
- 覆核責任人，以及覆核紀錄的保存方式

常見問題

政策寫得完整，但沒有指定誰覆核、在什麼時點覆核，實際上等於沒有覆核。

04 審核紀錄

紀錄的作用，是在被問到的時候，可以在一個工作天內拿出來。

- 工具引進的審批紀錄

- 員工確認政策的紀錄
- 高風險用途的人工覆核紀錄
- 事故與異常通報紀錄，包括處理結果
- 定期覆核清冊與風險評估的紀錄

常見問題

紀錄散落在電郵與即時通訊之中，無法在期限內整理出來，是最常見的失分位。

05 董事會一頁摘要

董事會需要的是風險與決定，不是設定細節。一頁為限。

- 機構使用了哪些人工智能，以及用途
- 風險最高的兩至三項，以及原因
- 已經採取的措施
- 尚未解決的項目，以及需要董事會決定的事項
- 下一次覆核時間

常見問題

把技術文件直接交給董事會，結果是沒有人讀，也沒有留下監督紀錄。

三、客戶問卷通常會問什麼

下表是我們在採購問卷、盡職調查問卷與審計師查詢中最常遇到的問題，以及對應的文件。

問卷常見問題	對應文件	準備後的回應時間
貴機構使用哪些人工智能系統？	01 系統清冊	一個工作天內
由誰負責監督？	01、05	一個工作天內
處理哪些個人資料？	01、02	二至三個工作天
資料儲存在哪裡？	01	一個工作天內
高風險用途如何控制？	02、03	二至三個工作天
有沒有事故通報機制？	04	一個工作天內
管理層如何監督？	05	一個工作天內

四、九十天計劃

如果現在手上什麼都沒有，以下是我們建議的次序。第一項通常在兩週內可以完成，而且是其餘各項的前提。

時間	工作	產出
第 1-2 週	盤點所有 AI 系統，包括同事自行訂閱的服務	AI 系統清冊初稿
第 3-6 週	逐項評估風險與對個人的影響，界定高風險用途	風險與影響評估
第 7-10 週	發布使用政策與核准工具清單，指定覆核責任	使用政策與人工覆核規則
第 11-13 週	建立紀錄與事故通報流程，向董事會提交摘要	審核紀錄與董事會一頁摘要
之後	每季覆核清冊與風險，每年全面檢視	覆核紀錄

五、第三方與供應商

大部分 AI 風險並非來自你自行開發的模型，而是來自你使用的服務。客戶問卷通常會追問這一部分。

- 供應商是否說明其模型訓練方式與資料處理安排
- 你的資料是否會被用於訓練，以及可否關閉
- 資料儲存與處理所在的司法管轄區
- 服務水平、事故通知義務與合約終止安排
- 供應商自身的認證，例如 ISO/IEC 27001 或 ISO/IEC 42001
- 分包處理者清單

實務做法

在採購階段就把這些問題放進供應商問卷，答案通常完整。事後補問，往往只能得到部分回覆。

六、事故處理的五個步驟

AI 相關事故的形式與傳統資訊保安事故不同：輸出偏誤、資料外洩、防護被繞過，都屬於此類。以下是我們建議的流程。

- 一、即時隔離：暫停該系統或該用途，並保留現狀。
- 二、評估範圍：涉及哪些個人資料、哪些客戶、多少人。
- 三、通報：確認是否觸發 PCPD、金管局、證監會或客戶合約下的通報義務，以及各自的時限。
- 四、補救與覆核：修正成因，重新評估風險，更新清冊。
- 五、紀錄：把上述四步寫成一份可查核的紀錄。這份紀錄本身就是第 04 項文件的一部分。

提醒

通報時限通常由發現起計，而非由確認成因起計。先通報、後補充，是較安全的做法。

七、澳門

澳門沒有獨立的 AI 專門法例。個人資料保護辦公室（PDPB）發布的 AI 與私隱資料明確標示不具約束力，但仍是監管期望的指標；第 8/2005 號法律（個人資料保護法）則適用於在澳門處理個人資料的活動。

若你在澳門設有營運或客戶，實務上可與香港一併處理：同一套五項文件，加入澳門的資料處理與通報要求即可。

八、常見陷阱

- 由政策開始，而不是由清冊開始。政策會寫得很空泛，因為沒有具體對象。
- 把 AI 治理當成資訊部門的事。問卷問的是機構的責任，由誰回答都要有管理層授權。
- 一次過追求完美。先發布一頁的版本，比半年後仍未發布好。
- 只覆蓋生成式人工智能，忽略已在使用的傳統預測模型與自動化決策。
- 政策沒有覆核紀錄，無法證明已執行。

九、與 ISO/IEC 42001 的關係

ISO/IEC 42001:2023 是可認證的人工智能管理系統標準。上述五項文件已覆蓋標準中大部分與營運有關的核心要求；若要取得第三方認證，仍須補上組織脈絡、領導承諾、內部審核與管理審查等管理系統元素。

若目的只是回覆客戶問卷，五項文件通常已經足夠。是否取得認證是一個獨立的決定，不必與問卷準備綁在一起。

關於界線

我們是就緒準備與顧問夥伴。我們不執行認證審核，亦不簽發證書。認證審核由認可認證機構執行。

十、詞彙表

詞彙	說明
AI 系統	以機器學習或其他自動化方法，產生輸出（例如預測、建議或內容）並用於業務的系統。
影子 AI	未經機構核准、由員工自行訂閱或使用的 AI 工具與服務。
風險分級	按影響程度與可能性，把系統歸入不同等級，以決定所需的控制措施。

詞彙	說明
人工覆核	在自動化輸出實際影響個人或機構之前，由具權限的人檢視並作決定。
管理系統	把政策、程序、責任與紀錄組織起來，使要求可以持續被執行，而不依賴個人記憶。
認證機構	經認可、可對管理系統進行第三方審核並簽發證書的機構。

十一、如何選擇標準：42001、27001、27701

這三個標準經常被混為一談，職責其實相當清楚。

- ISO/IEC 27001 是資訊安全管理系統，處理資訊保安風險，通常是起點。
- ISO/IEC 27701 是私隱資訊管理延伸，處理個人資料控制者與處理者的責任。
- ISO/IEC 42001 是人工智能管理系統，處理 AI 特有的風險：資料來源、模型、影響評估與人工監督。

若客戶只要求一套，多數情況由 27001 起步。若客戶明確要求 AI 治理，則以 42001 為主；兩者一併實施可以共用大量文件，額外工作量通常在兩至三成之間。

附錄 A：AI 系統清冊範本

以下欄位可直接用於你的清冊。前兩行是填寫示例，不屬於你的實際紀錄。

系統／供應商	用途與負責人	資料類別與地點
示例：Microsoft 365 Copilot	撰寫內部文件・ 資訊部主管	內部文件，不含個人資料・ 香港及新加坡
示例：客戶服務聊天機械人	回覆客戶查詢・ 客戶服務主管	姓名、聯絡方式、對話紀錄・ 香港

建議一併記錄：上線日期、最近覆核日期、風險分級，以及是否已列入核准工具清單。

附錄 B：客戶問卷回覆範本

以下是一段可直接改寫使用的回覆文字，適用於「貴機構如何使用人工智能」一類的問題。

範本

本機構已建立人工智能系統清冊，記錄現時使用的工具、用途、具名負責人，以及所處理的資料類別。所有工具均經核准工具清單審批，並受書面使用政策約束；涉及個人資料或影響個人的輸出，須經指定人員人工覆核。我們每季覆核清冊與風險評估，並保留覆核紀錄。清冊及風險評估摘要可於一個工作天內提供。

附錄 C：使用政策條文範本

以下條文可直接採用或改寫，建議與人力資源守則一併發布，並要求同事確認。

- 一、適用範圍：本政策適用於全體員工、承包商及代理人，以及其在業務中使用的所有人工智能工具。
- 二、核准工具：同事只可使用核准工具清單上的工具。清單由資訊部會同資料保護負責人每季覆核。
- 三、禁止用途：不得用於作出對個人有法律或重大影響的自動化決定；不得用於生成或散播歧視性、誤導性或侵權內容。
- 四、輸入資料：未經書面批准並完成風險評估，不得輸入客戶或員工的個人資料、機密資料、未公開財務資料，或第三方的知識產權內容。
- 五、人工覆核：凡輸出將影響客戶、僱員或第三方，或擬對外發布，必須由指定人員在採用前覆核，並記錄覆核結果。
- 六、對外披露：如客戶要求，本機構會說明所使用的人工智能系統及其資料處理安排。
- 七、事故通報：同事發現任何異常輸出、偏誤或資料外洩，須於同一個工作天內通知資訊部及資料保護負責人。
- 八、違反本政策：可能構成紀律事宜。

提醒

政策條文之後要附上一頁「誰負責什麼」，列明每個角色的具體動作。否則條文無法執行。

附錄 D：董事會一頁摘要範本

欄目	內容
使用情況	本季在用 AI 系統 X 個，較上季增加／減少 Y 個。
最高風險	列出兩至三項，並簡述原因。
已採取措施	已完成的控制措施，以及其實際成效。
未解決事項	需要董事會決定的事項，並附建議方案。
合規狀態	相關監管要求的符合情況，以及 ISO/IEC 42001 的準備程度。
下次覆核	日期與負責人。

篇幅以一頁為限。若需要第二頁，通常代表仍未決定要董事會決定什麼。

十二、關於避風港

避風港（香港）諮詢有限公司是一間香港顧問公司，專注 AI 治理、ISO 認證與 IT 審計，服務香港及澳門企業。

我們公開顧問的資歷、我們的方法，以及每一項監管立場的來源與日期。

- 電話：+852 5382 0328
- 電郵：info@safeharbour.hk
- WhatsApp：+852 5382 0328
- 地址：香港上環永樂街 235 號 235 商業中心 22 樓 A 室

本手冊僅供一般參考，不構成法律意見。監管要求會隨時間變動，引用前請先核實最新版本。