職務代理檢查器，從規格書做到技能

講義 · v0.1

國泰世華基層主管訓講義｜2026 年 11 月 18 日（三）09:00–16:30

架構與樣張

全班用同一個工具走完一整天：照規格書做出來、讓 AI 反問補完、把判斷寫成規則、去識別化之後交給 AI，最後放進專案、轉成技能。



春天的分行街口。今天的練習都用合成資料，真實資料留在行內。

這份講義怎麼用

講義依今天上課的順序排列。每一段標出時間，並用三種標記說明這一段在做什麼：

- 聽 這一段要聽懂的觀念，講師說明時可以對照著看。
- 做 自己動手的步驟，提示詞旁邊有「複製提示詞」按鈕。
- 查 做完之後的檢查，逐項確認再往下走。

網頁打不開時，可以看同一份內容的 PDF 版，或從課堂白板複製提示詞。所有練習一律使用合成資料，行內的真實資料不上傳到任何雲端服務。

今天的地圖

08:50–09:00	**開始之前：設定與檢查** 留下：test.html
09:00–09:30	**開場與暖身：一句話做出陽春版** 留下：陽春版知識圖譜
09:30–10:15	**核心練習一：照規格書做出職務代理檢查器 v0** 留下：工具 v0、驗收紀錄
10:15–10:25	**休息**
10:25–10:45	**ChatGPT 帳號設定：讓 AI 先認識你** 留下：自訂指示
10:45–11:25	**核心練習二：交辦，讓 AI 先反問** 留下：工作說明書 V1.0、工具 v1、HANDOFF.md
11:25–11:55	**核心練習三：同一題問三次，答案會一樣嗎** 留下：判定規則表、工具 v2
11:55–12:00	**上午收尾：建立專案** 留下：專案資料夾
12:00–13:00	**午休**
13:00–14:00	**核心練習四：去識別化與還原** 留下：去識別化工具、對照檔、還原後的信件與說明
14:00–14:10	**休息**
14:10–15:00	**核心練習五：專案＝核心提示詞＋參考資料** 留下：「職務代理小幫手」專案
15:00–15:30	**核心練習六：轉成技能** 留下：tool-page-builder 技能包
15:30–15:40	**休息**
15:40–16:10	**終幕：MVP（最小可行的實作）與三句話向長官說明** 留下：MVP 實作說明
16:10–16:30	**問答與結語**

開始之前：設定與檢查

08:50–09:00 10 分鐘

做 10 分

用到 手機、ChatGPT、Copilot、記事本、Edge　**留下** test.html

上課前十分鐘，先把今天會用到的東西確認一次。11/3 那一班有人進不了 Teams、有人打不開講義網頁，開場的時間花在處理連線上，所以這一段請在 09:00 以前完成。

做 五個步驟

1. 用手機掃描課堂白板的 QR code。再把白板網址用郵件或 Teams 傳到自己的電腦，之後複製提示詞比較方便。
2. 在桌面建立一個資料夾，例如「AI練習」。今天做出來的檔案都存在這裡，檔名一律用英文。
3. 登入行內的 ChatGPT 企業版帳號，確認左側看得到「專案」。每換一件事，就開一個新對話。
4. 打開 Copilot，確認可以選「深度思考」或推理模型。功能名稱以你的畫面為準。
5. 測試存檔與開啟：打開記事本，貼上下面這段文字。選「另存新檔」，存檔類型選「所有檔案」，編碼選「UTF-8」，檔名輸入 test.html，存到「AI練習」資料夾。接著在檔案上按右鍵，選「開啟檔案」，再選 Microsoft Edge。

```
<!DOCTYPE html><html lang="zh-Hant"><meta charset="utf-8">
<title>test</title><body style="font-family:sans-serif;padding:40px"><h1>
存檔與開啟都正常</h1><p id="t">如果看得到這一行，代表程式沒有執行。</p>
<script>document.getElementById('t').textContent='程式也可以執行：'+new
Date().toLocaleString('zh-TW');</script></body></html>
```

查 五項都打勾再開始

- [ ] 白板網址已經在自己的電腦上。
- [ ] 桌面有「AI練習」資料夾。
- [ ] ChatGPT 左側看得到「專案」。
- [ ] Copilot 可以選深度思考或推理模型。
- [ ] test.html 打開之後顯示「存檔與開啟都正常」，下一行顯示現在的時間。

如果第 5 步卡住，記下卡在哪一步：存不了檔、打不開、打開是空白，或是只看到「程式沒有執行」那一行。把這個情況告訴同桌或講師，下午的工具會用到同樣的存檔方式。

開場與暖身：一句話做出陽春版

09:00–09:30 30 分鐘

聽 15 分 | 做 10 分 | 查 5 分

用到 Copilot　**留下** 陽春版知識圖譜

聽 今天的規則與三個成品

今天的練習都用合成資料，行內的真實資料不上傳。講師會打開三個成品：一份有完整規格書的卡車裝載模擬器、一張只用一句提示詞做出來的知識圖譜，以及用同一種規格書寫法做的銀行版臨櫃等候模擬器。三個成品都是 Copilot 用深度思考做出來的單一 HTML 檔。

做 一句話知識圖譜

1. 在 Copilot 開一個新對話，選深度思考。
2. 複製下面的提示詞，再把白板上的《金融服務業公平待客原則》全文貼在最後一行。
3. 等它寫完，照「開始之前」第 5 步的方式存成 knowledge_graph.html，用 Edge 打開。

P0｜一句話知識圖譜（Copilot，深度思考）

```
將以下文件的所有知識點，以 SVG 繪製圖文並茂的知識圖譜，做成一個單一 HTML 檔。不要載入任何外部程式庫。
---
（貼上白板上的《金融服務業公平待客原則》全文）
```

查 數一數重複的說明

數一數有幾個知識點的說明是同一句話。一句話只能得到陽春版，想要真的能用的工具，就要先寫出規格書。下一段就從規格書開始。

核心練習一：照規格書做出職務代理檢查器 v0

09:30–10:15 45 分鐘

聽 5 分	做 20 分	查 15 分	補充 5 分

用到 Copilot　**留下** 工具 v0、驗收紀錄

聽 規格書的九個部分

規格書寫清楚九件事：輸出形式與限制、背景、基本設定、資料表、規則、呈現、數據面板、匯出、假設說明。銀行版另外加上三項限制：不載入外部程式庫、只用合成資料、檔名用英文。不載入外部程式庫，可以避免行內網路擋住外部資源，也確保工具不會把資料送出去。

做 貼上規格書，存檔，打開

1. 在 Copilot 開一個新對話，選深度思考，貼上 P1。
2. 生成需要幾分鐘，等待時先讀下面的驗收表。
3. Copilot 沒有提供下載時，把程式碼全部複製到記事本，存成 UTF-8 編碼的 leave_coverage_v0.html，再用 Edge 打開。
4. 生成失敗或程式不完整時，改用課堂白板上的備用版工具。

P1｜職務代理與休假排定檢查器（Copilot，深度思考）

直接產出一個「單一 HTML 檔」（HTML、CSS、JavaScript 全部寫在同一個檔案裡），不要問我問題，程式要完整可以執行，不要省略、不要留 TODO。介面全部用繁體中文。不要載入任何外部程式庫或字型，不要連線到任何網址，只用原生 JavaScript、SVG 或 Canvas。所有資料都是合成資料。

【背景】
我們是銀行的基層主管。單位裡每一項業務都要有人會做，同仁休假時也要有合格的代理人。每個月排休假時，常常到了當天才發現某項業務沒有人能代理。請做一個「職務代理與休假排定檢查器」：輸入同仁會哪些業務、誰要休假，就能看到每一天、每一項業務的代理狀況，哪幾天有缺口，以及可以怎麼調整。

【基本設定】
畫面上可以修改，旁邊註明「預設值，請依單位實際情況修改」：
- 檢查月份：預設 2026 年 12 月。
- 上班日：週一到週五。國定假日可以在畫面上新增或刪除，預設不列任何假日。

- 每位同仁每天最多計入幾項業務：預設 2 項。

【資料】
畫面上有三張可以編輯的表格，每張都可以新增、刪除列，可以貼上 CSV（第一列是欄位名稱），也可以下載 CSV 範本。
1. 業務清單：業務名稱、每天最少在位人數、說明。
2. 能力表：同仁對每一項業務的等級，2＝可獨立辦理、1＝需有人陪同、0＝不會。用「同仁 × 業務」的格子呈現，點一下格子就能切換等級，並統計每一項業務有幾位可以獨立辦理。
3. 休假申請：同仁、開始日期、結束日期、假別（特休、事假、病假、公假、訓練）、可否調整（是／否）。
同仁的欄位：代號、姓名、職稱。
預設載入這組合成資料：12 位同仁（代號 A01 到 A12，姓名與職稱自行編寫），6 項業務（臨櫃存提款每天至少 3 人、外匯 1 人、開戶 1 人、放款撥貸 1 人、理財諮詢 1 人、主管授權放行 1 人），12 月的休假申請約 10 筆，其中要有幾天會讓外匯和主管授權放行出現缺口。

【檢查規則】
1. 每一個上班日、每一項業務，計算當天在位、而且等級為 2 的同仁可以分派幾人。等級 1 的同仁不計入人數，但在明細裡列為「可以陪同支援」。
2. 分派時先處理合格人數最少的業務，每位同仁每天計入的業務數不能超過基本設定的上限。
3. 分派後的人數低於最少在位人數就是紅燈，剛好等於是黃燈（沒有餘裕），高於是綠燈。
4. 只有一位同仁可以獨立辦理的業務，在業務清單旁標示「只有一人會」。
5. 針對每一個紅燈日，找出「可調整」的休假申請，建議改到同一個月裡、天數相同、而且不會造成任何紅燈的最近上班日。建議可以一鍵套用，也可以復原。
6. 資料在 300 筆以內時，1 秒內算完。

【呈現】
- 月曆熱力圖：橫軸是日期，縱軸是業務，格子用紅、黃、綠表示。滑鼠移到格子上，顯示當天在位的合格同仁、休假的同仁、還缺幾人。
- 同仁時間軸：每位同仁一列，標出休假期間，不同假別用不同顏色。
- 切換按鈕：「顯示全部業務／只看有缺口的業務」、「依業務排序／依同仁排序」。

【數據面板】
- 本月紅燈天數、黃燈天數，以及缺口最嚴重的業務與日期。
- 只有一人會的業務清單。
- 培訓建議：如果把某位等級 1 的同仁提升到等級 2，可以消除幾個紅燈。列出效益最大的前三名（同仁、業務、可以消除的紅燈數）。
- 套用調整建議前後的比較。

【匯出】
- 下載代理缺口清單 CSV（日期、業務、需要人數、合格在位人數、缺口、休假的同仁）。

- 下載調整建議 CSV。
- 可以列印，列印時顯示月曆熱力圖與代理缺口清單。

【頁面最下方】
用一小段文字列出你用了哪些假設，例如：資料是合成的，等級由主管判定，沒有考慮輪調、加班與跨分行支援，調整建議只供參考，排定仍由主管決定。

查 驗收表

每一桌對照規格書逐項檢查，把沒做到的地方記下來。這份紀錄就是下一段的材料。

1. 打開之後不是空白畫面，預設資料已經載入。
2. 三張表格都可以編輯、新增與刪除列。
3. 可以貼上 CSV，也可以下載 CSV 範本。
4. 能力表的格子點一下就會切換等級，每項業務的人數統計跟著改變。
5. 月曆熱力圖有紅、黃、綠三種顏色，滑鼠移上去看得到明細。
6. 「只有一人會」的業務有標示出來。
7. 紅燈日有調整建議，可以套用，也可以復原。
8. 培訓建議列出前三名。
9. 可以下載兩份 CSV，列印時版面正確。
10. 頁面最下方有假設說明。
11. 手算抽查：挑一天、一項業務，自己數一次合格在位的人數，和工具的結果比較。

ChatGPT 帳號設定：讓 AI 先認識你

10:25–10:45 20 分鐘

聽 5 分 | 做 12 分 | 查 3 分

用到 ChatGPT　**留下** 自訂指示

聽 記憶與自訂指示

記憶會記住你和它聊過的基本事實，但不會記住所有細節，所以不要把它當成資料庫。自訂指示是寫在最外層的規則，會影響所有對話。下午會用到的「專案」也有自己的指示，專案有寫指示時，先照專案的指示。

做 五個設定

1. 左下角點頭像，進入「個人化」。企業版的畫面可能和個人版不同，名稱以你的畫面為準。
2. 記憶：確認已經打開。看「記憶摘要」的第一句，有沒有認出你的職務。
3. 基本風格：選「專業」或「效率」。不喜歡 AI 寫太多，就選效率。
4. 自訂指示：從「經常叫 AI 做、卻都做不好的事」開始寫。可以先用下面的範例，再改成你的說法。講師的完整版本在課堂白板上。
5. 儲存之後開一個新對話，請它用三句話說明之後會怎麼回答你。

> 請用臺灣繁體中文回答。先給結論或成品，再補必要的依據。
> 事實、推論與建議分開寫。事實要有來源，查不到就說查不到，不要自己編造數字或規定。
> 我是銀行的基層主管。我貼上的內容如果有〔類別_編號〕這種代號，代表已經在本機去識別化，請保持代號原樣，不要猜測代號背後的資料。
> 資訊不足、而且會影響結果時，先問我一個最關鍵的問題，其他部分照合理的假設完成，並標明假設。

查 新對話有沒有照規則

新對話的回答有沒有先給結論、有沒有分開事實與建議。沒有的話，回到自訂指示，把那一條寫得更具體。

核心練習二：交辦，讓 AI 先反問

10:45–11:25 40 分鐘

聽 5 分 | 做 25 分 | 查 10 分

用到 ChatGPT　**留下** 工作說明書 V1.0、工具 v1、HANDOFF.md

聽 交辦之前先讓 AI 問清楚

自己寫的規格書一定不完整。交辦比較複雜的工作時，先請 AI 反問，把工作說明書寫完整，再請它動手。做完之後留下一份交接檔（HANDOFF.md），寫清楚做了什麼、還沒做什麼、下一位接手的人第一步做什麼。這件事做完一次還沒結束，之後還要持續修改。

做 反問兩三輪，再做出 v1

1. 在 ChatGPT 開一個新對話，貼上 P2，再貼上 P1 規格書與上一段的驗收紀錄。
2. 回答 AI 的問題兩到三輪。它附上選項時，回答 1、2、3 就可以。
3. 輸入「可以了」，請它整理成工作說明書 V1.0。
4. 在同一個對話貼上 P3，下載 ZIP，解壓縮到「AI練習」資料夾，打開 v1。

P2｜先反問，再寫工作說明書（ChatGPT）

以下是我交給 AI 做工具網頁的規格書，以及我對第一版的驗收紀錄。規格書是我自己先寫的初稿，內容一定不完整。

你的目標是先問我問題，協助我把規格書改寫成一份完整的工作說明書，這個階段先不要修改程式。每一輪最多問五個問題，優先問會影響檢查結果的問題，例如：誰會使用這個工具、資料從哪裡來、等級由誰判定、例外情況怎麼處理、結果由誰檢查。問題可以附上選項，讓我回答 1、2、3 就好。

等我說「可以了」，再整理成「工作說明書 V1.0」。

（貼上 P1 規格書）
（貼上驗收紀錄）

P3｜做出 v1 與交接檔（ChatGPT）

請依照工作說明書 V1.0，產出第二版的工具網頁，限制和原本的規格書相同（單一 HTML 檔、不載入外部程式庫、合成資料、繁體中文）。

同時寫一份 HANDOFF.md，內容包含：這個工具的用途、檔案清單、資料欄位、檢查規則、這一版改了什麼、還沒完成的事項，以及下一位接手的人第一步要做什麼。

所有檔名都用英文，打包成一個 ZIP 讓我下載。

查 工作說明書與交接檔

- 工作說明書有沒有寫清楚：誰使用、資料從哪裡來、等級由誰判定、例外怎麼處理、結果由誰檢查。
- HANDOFF.md 有沒有寫「下一位接手的人第一步要做什麼」。
- v1 是否改善了驗收紀錄裡沒做到的地方。

核心練習三：同一題問三次，答案會一樣嗎

11:25–11:55 30 分鐘

聽 5 分 | 做 18 分 | 查 7 分

用到 ChatGPT、Copilot　**留下** 判定規則表、工具 v2

聽 從 1 到 10 選一個數字

先請全場從 1 到 10 選一個數字，再看 AI 怎麼選。同一段提示詞跑三次，結果可能不一樣。職務代理裡真正需要判斷的一步，是誰算「可以獨立辦理」。如果讓 AI 每次憑感覺評，每次的結果都不一樣。把主管確認過的規則寫進工具的程式裡，結果才會穩定。

做 判定三次，寫成規則表

1. 在 ChatGPT 開新對話，模型選「中」，貼上 P4。同樣的步驟再做兩次，每次都開新對話。
2. 用 Copilot 再跑一次 P4。
3. 比較四次的結果，圈出每次不一樣的同仁。
4. 回到其中一個對話，貼上 P5，請它寫成判定規則表。
5. 時間足夠時，回到 P3 的對話貼上 P6，做出 v2。

P4｜判定 8 位同仁的外匯等級（ChatGPT 開新對話跑三次，模型選「中」，再用 Copilot 跑一次）

以下是 8 位同仁在「外匯」業務的資料，全部是合成資料。請判斷每位同仁目前的等級（2＝可獨立辦理、1＝需有人陪同、0＝不會），每位用一句話說明理由，最後用表格列出結果。

A01｜襄理，年資 10 年｜近六個月自己臨櫃辦理 0 件，負責放行外匯交易約 40 筆｜外匯訓練：已完成
A02｜行員，年資 6 年｜近六個月獨立辦理 85 件｜外匯訓練：已完成｜錯誤紀錄：無
A03｜行員，年資 8 年｜三年前常辦外匯，近六個月 0 件｜外匯訓練：證書仍在有效期限內
A04｜行員，年資 2 年｜近六個月獨立辦理 30 件｜外匯訓練：已完成｜錯誤紀錄：一次匯款幣別錯誤，當天更正
A05｜行員，上個月從別的分行調來｜自述在前一家分行辦過外匯，本分行沒有紀錄｜外匯訓練：紀錄待查
A06｜行員，年資 4 年｜近六個月獨立辦理 15 件｜外匯訓練：新版課程尚未完成
A07｜行員，年資 1 年｜近六個月經手 12 件，全部有資深同仁陪同｜外匯訓練：已完成
A08｜新進行員，到職 3 個月｜近六個月 0 件｜外匯訓練：排在下個月

P5｜寫成判定規則表（同一個對話）

請先不要判斷等級。剛才的判斷每次都不一樣，因為「可獨立辦理」沒有明確的定義。請協助我把判定方式寫成一張規則表：
1. 每一個等級要符合哪些條件、要用到哪些欄位，門檻寫成數字。
2. 資料不足或互相矛盾的時候怎麼處理。
3. 哪些情況一定要由主管決定，AI 只能列出來，不能自己判斷。

寫完之後，用這張規則表重新判定 8 位同仁，列出每一位的信心水準，以及需要主管確認的人。

P6｜把規則表加進工具，做出 v2（回到 P3 的對話）

請把下面這張判定規則表加進工具網頁。新增一張「經手紀錄與訓練」表格，欄位是：同仁代號、業務、近六個月獨立辦理件數、近六個月陪同辦理件數、訓練狀態、錯誤紀錄次數。工具依規則表自動算出能力表的等級。主管可以手動修改某一格，但要填寫原因，畫面上標示哪些格子是手動修改過的。其他功能維持不變，並更新 HANDOFF.md。

（貼上主管確認過的規則表）

查 規則表由主管決定

- 四次結果裡，哪幾位同仁的等級不一樣？原因是哪一個欄位模稜兩可？
- 規則表的門檻是 AI 提的建議，你同意嗎？哪一條要改？
- 規則表有沒有列出「一定要由主管決定」的情況？

上午收尾：建立專案

11:55–12:00 5 分鐘

做 5 分

用到 ChatGPT　**留下** 專案資料夾

1. 在 ChatGPT 左側建立一個專案，名稱例如「職務代理小幫手」，記憶選「預設記憶」。
2. 從上午每一段對話的「⋯」選「移至專案」，把對話搬進去。

午休 12:00–13:00。

核心練習四：去識別化與還原

13:00–14:00 60 分鐘

聽 10 分 | 做 35 分 | 查 15 分

用到 Copilot、ChatGPT　**留下** 去識別化工具、對照檔、還原後的信件與說明

▼ **這一段的重點**

問題：行內規定不能把真實資料上傳到雲端 AI。可是主管想請 AI 協助的工作，例如回覆客訴、整理代理缺口、撰寫排假協調訊息，內容裡常常有客戶與同仁的姓名、證號和電話。

做法：在自己的電腦把個人資料換成代號，只把代號版本交給 AI。拿回 AI 的結果之後，再用留在自己電腦的對照檔換回原本的內容。替換的工具由 AI 撰寫程式，真實資料只在你的瀏覽器裡處理，不經過任何 AI。

送出前要記得：規則和字典會漏抓，年齡、居住狀況這類描述組合起來也可能認出是誰，所以送出前一定要自己看過一次。對照檔不可以和去識別後的資料一起上傳或寄出。哪些資料可以交給雲端 AI 處理，仍以行內規範與法遵的判斷為準。

聽 一、為什麼要先去識別化

星期四下午，一位客戶到分行反映，定存提前解約之後拿到的利息比預期少。你想請 ChatGPT 協助寫一封說明信，也想把這一週的三筆客訴分類，看看哪一筆要先處理。這些紀錄裡有客戶的姓名、身分證字號、手機和帳號，依行內規定不能上傳到雲端 AI。

不用 AI，工作仍然做得完，只是比較慢。直接把紀錄貼給 AI，又違反規定。上午的練習已經走過第一條路：請 AI 寫一個在自己電腦執行的工具，資料留在本機，AI 只負責寫程式。這一段介紹第二條路：先把個人資料換成代號，再把代號版本交給 AI。

圖表 1　資料不上雲的兩條路



第一條路：請 AI 寫工具

1. 用 Copilot 描述你要的工具，Copilot 寫出一個 HTML 檔。
2. 在自己的電腦打開工具，載入真實資料。
3. 計算與判斷都在瀏覽器裡完成，資料不離開電腦。

適合規則清楚、要重複使用的計算，例如職務代理檢查。



第二條路：先去識別化，再交給 AI

1. 在本機把個人資料換成代號，對照檔留在自己的電腦。
2. 把代號版本交給 ChatGPT 撰寫、摘要或分類。
3. 拿回結果，在本機用對照檔換回原本的內容。

適合需要 AI 理解文字內容的工作，例如寫信、摘要和分類。

兩條路有一個共同點：AI 負責寫程式，或處理代號版本的文字，真實資料始終留在你的電腦。第二條路用到的去識別化工具，也是請 Copilot 寫出來的。

▼ **本節重點**

- 資料不上雲有兩條路：請 AI 寫在本機執行的工具，或先去識別化再交給 AI。
- 兩條路都只讓 AI 處理程式或代號，真實資料只在自己的電腦裡處理。

聽 二、哪些資料要處理：直接識別與間接識別

可以認出一個人的資料分成兩種。直接識別的資料單獨出現就能認出是誰，例如姓名、身分證字號、手機和帳號。間接識別的資料單獨看不出是誰，但是幾項組合起來，就可能縮小到一個人。另外，健康和家庭的細節即使不足以認出是誰，也不應該交給 AI。

圖表 2　三種要處理的資料

種類	例子	處理方式
直接識別	姓名、身分證字號、居留證號、手機、市話、Email、信用卡號、銀行帳號、地址、員編、出生日期、統一編號	換成代號，例如〔姓名_001〕
間接識別	年齡、居住狀況（獨居、住在分行附近）、職稱加單位、精確到日的日期、精確到元的金額	改成年齡層、只寫到月份，或改成金額級距
健康與家庭細節	住院開刀、帶小孩出國	刪除，或改成「病假」「特休」這類一般描述

以練習資料乙的第 3 筆為例，即使拿掉客戶的姓名，「82 歲、住在分行附近的公寓、獨居、要解約兩張定存共 150 萬元」這幾項資訊放在一起，分行同仁大概都知道是哪一位客戶。

間接識別的資料，工具只能提醒，不能替你決定要不要改寫，因為改寫之後的內容還要能讓 AI 完成工作。例如請 AI 判斷這一筆是不是疑似詐騙時，「高齡客戶」和「要求提領大筆現金」這兩項資訊仍然要保留，精確的年齡和住處則可以拿掉。

▼ **本節重點**

- 直接識別的資料換成代號。間接識別的資料改成年齡層、月份或金額級距。健康與家庭細節刪除，或改成一般描述。
- 改寫到什麼程度，要看 AI 完成這項工作需要哪些資訊。

聽 三、代號與對照檔：可以還原的去識別化



對照檔就是鑰匙：留在自己的電腦，和代號版本分開放。

把「周美玲」改成「〇〇〇」或「周＊＊」，也是一種去識別化，但是這樣做有兩個問題。第一，資料裡有好幾位客戶時，AI 分不出哪一段在講誰。第二，拿回 AI 的結果之後，你沒辦法把名字換回去，只能自己一封一封補上。

本課的工具用代號解決這兩個問題。代號的格式是〔類別_三位流水號〕，例如〔姓名_001〕、〔身分證_001〕、〔手機_002〕。同一個值在整份資料裡一律使用同一個代號，所以 AI 知道兩段話講的是同一位客戶。每一個代號和原始內容的對應，都記在一個對照檔裡。

圖表 3　去識別化前後對照（練習資料乙第 1 筆，節錄）

原文

> 客戶周美玲（身分證 A223456781，手機 0912-345-678）到中山分行反映，帳號 012-34-567890-1 的定存提前解約後，拿到的利息比她預期少了約 1,200 元……承辦行員林怡君。

去識別後，交給 AI 的版本

> 客戶〔姓名_007〕（身分證 〔身分證_001〕，手機 〔手機_001〕）到〔分行_001〕反映，帳號〔帳號_001〕 的定存提前解約後，拿到的利息比她預期少了約 1,200 元……承辦行員〔姓名_003〕。

對照檔（節錄，只存在自己的電腦）

```
{ "token": "〔姓名_007〕", "category": "姓名", "original": "周美玲" }
{ "token": "〔身分證_001〕", "category": "身分證", "original": "A223456781" }
```

對照檔就是鑰匙。拿到對照檔的人，就能把代號換回真實資料。所以對照檔只能存在自己的電腦，不可以和去識別後的資料一起上傳到 AI，也不可以一起寄出。工作完成之後，依行內的資料保存規定刪除。

還原時，工具會比對 AI 回覆裡的代號。AI 有時會改動代號的寫法，例如把〔姓名_001〕寫成 [姓名_001] 或【姓名_001】，工具仍然能辨認並還原。AI 也可能自己編出對照檔裡沒有的代號，例如〔姓名_099〕，工具會用紅色標出來，提醒你這一段內容沒有依據。

▼ **本節重點**

- 代號讓 AI 分得出不同的人，也讓你能把結果換回原本的內容。
- 對照檔就是鑰匙，只存在自己的電腦，不和去識別後的資料一起上傳或寄出。
- 還原報告會指出 AI 改過寫法的代號，以及 AI 自己編的代號。

聽 四、身分證字號的檢查碼

格式像身分證字號的號碼，不一定是身分證字號。申請書編號、交易序號和系統代碼，都可能是一個英文字母加九位數字。反過來說，打錯一碼的身分證字號仍然是個人資料。Microsoft Purview 這類企業級的資料保護工具，會先驗證檢查碼，再看號碼附近有沒有「身分證」這類字詞，判斷這個號碼有多大的可能是身分證字號。本課的工具也採用同樣的做法。

檢查碼怎麼算

身分證字號由一個英文字母和九位數字組成。英文字母是縣市代碼，第二碼代表性別，1 是男性，2 是女性。2021 年起核發給外籍人士的新式統一證號，格式和身分證字號相同，第二碼是 8 或 9。最後一碼是檢查碼，由前面九碼計算出來。

圖表 4　檢查碼的算法（以 A223456781 為例）

位置	字母 A	第 2 碼	第 3 碼	第 4 碼	第 5 碼	第 6 碼	第 7 碼	第 8 碼	第 9 碼	檢查碼
數字	1　0	2	2	3	4	5	6	7	8	1
權重	1　9	8	7	6	5	4	3	2	1	1
乘積	1　0	16	14	18	20	20	18	14	8	1

1. 英文字母依對照表轉成兩位數字：A=10、B=11、C=12、D=13、E=14、F=15、G=16、H=17、I=34、J=18、K=19、L=20、M=21、N=22、O=35、P=23、Q=24、R=25、S=26、T=27、U=28、V=29、W=32、X=30、Y=31、Z=33。A 轉成 10。
2. 十位數乘 1，個位數乘 9：1×1＋0×9＝1。
3. 後面九位數字依序乘 8、7、6、5、4、3、2、1、1，加起來是 129。
4. 兩個結果相加：1＋129＝130。總和是 10 的倍數，所以檢查碼正確。

練習資料乙的第 2 筆有一個爭議款申請書編號 Q100200300，格式也是一個英文字母加九位數字。照同樣的方法計算，總和是 62，不是 10 的倍數，所以它不是有效的身分證字號。

工具怎麼判斷

工具找到一個英文字母加九位數字的號碼之後，會先驗證檢查碼，再看號碼前後 300 個字以內有沒有「身分證」「居留證」「統一證號」這類字詞，把號碼分成高、中、低三種信心。兩個英文字母加八位數字的舊式居留證號，工具只比對格式，不驗證檢查碼。

圖表 5　三種信心與預設的處理方式

信心	條件	預設的處理方式
高	檢查碼正確，附近也有證件的字詞	換成代號
中	檢查碼正確，附近沒有證件的字詞	換成代號
低	檢查碼不符，或第二碼不是 1、2、8、9	仍然換成代號，並列在提醒清單，說明原因

工具預設連信心低的號碼也一起替換，因為打錯一碼的身分證字號仍然是個人資料，寧可多遮一些。替換之後，提醒清單會列出這些號碼和原因。如果你確定它是申請書編號這類和個人無關的號碼，而且 AI 需要看到原始號碼，可以在畫面上改成「只提醒，不替換」。

▼ **本節重點**

- 格式像身分證字號，不一定是身分證字號。打錯一碼的身分證字號，仍然是個人資料。
- 工具先驗證檢查碼，再看附近有沒有證件的字詞，分成高、中、低三種信心。
- 預設連檢查碼不符的號碼也一起替換，提醒清單會列出原因。

做 五、操作步驟

下面八個步驟對應課堂上的練習。練習一律使用合成資料，不使用真實的客戶或同仁資料。

步驟 1　請 Copilot 寫出去識別化與還原工具

打開 Copilot，選擇深度思考，貼上提示詞 P7。Copilot 寫完之後，把程式碼存成英文檔名的 .html 檔，例如 deid_tool.html，再用 Edge 或 Chrome 打開。Copilot 生成失敗或程式不完整時，可以改用課堂白板上的備用版工具。

P7｜去識別化與還原工具（Copilot，深度思考）

```
直接產出一個「單一 HTML 檔」(HTML、CSS、JavaScript 全部寫在同一個檔案裡)，不要問我問題，程式要完整可以執行，不要省略、不要留 TODO。介面全部用繁體中文。

【安全限制】
這個工具會處理個人資料，所以要遵守以下限制：
1. 不要載入任何外部程式庫或字型。
2. 不要使用 fetch、XMLHttpRequest、WebSocket 或任何網址，所有處理都在瀏覽器裡完成。
```

3. 不要把原文或對照檔存進 localStorage，頁面關閉後就清除。
4. 畫面上方固定顯示「對照檔不可以和去識別後的資料一起上傳或寄出」。

【背景】
我們是銀行的基層主管。行內規定不能把真實資料上傳到雲端 AI。想讓 AI 協助寫信、摘要或分類時，要先在自己的電腦把個人資料換成代號，把代號版本交給 AI，拿回結果之後再換回來。請做一個「去識別化與還原工具」。

【功能一：去識別化】
1. 輸入：貼上文字，或從本機選擇 .txt、.csv 檔案。
2. 偵測規則（每一類都可以勾選開關）：
   - 身分證字號與居留證號：一個英文字母加九位數字，或兩個英文字母加八位數字。
   - 身分證字號的檢查碼：一個英文字母加九位數字的號碼要驗證檢查碼。英文字母依 A=10、B=11、C=12、D=13、E=14、F=15、G=16、H=17、I=34、J=18、K=19、L=20、M=21、N=22、O=35、P=23、Q=24、R=25、S=26、T=27、U=28、V=29、W=32、X=30、Y=31、Z=33 轉成兩位數，十位數乘 1、個位數乘 9，後面九位數字依序乘 8、7、6、5、4、3、2、1、1，全部相加是 10 的倍數才正確。再看號碼前後 300 個字以內有沒有「身分證」「居留證」「統一證號」這類字詞，分成三種信心：檢查碼正確而且附近有這類字詞是高，只有檢查碼正確是中，檢查碼不符或第二碼不是 1、2、8、9 是低。信心低的號碼預設仍然替換，並列在提醒清單裡說明原因，畫面上可以改成「只提醒，不替換」。兩個英文字母加八位數字的號碼只比對格式。
   - 手機：09 開頭共十碼，中間可以有連字號或空白。
   - 市話：區碼加號碼，例如 02-2345-6789。
   - Email。
   - 信用卡號：16 位數字，中間可以有空白或連字號。
   - 銀行帳號：10 到 16 位數字，中間可以有連字號，位數範圍在畫面上可以修改。
   - 統一編號：8 位數字，前面出現「統編」或「統一編號」時才算。
   - 出生日期：前面出現「生日」或「出生」時才算。
   - 地址：從縣市名稱開始，到「號」或「樓」結束。
   - 自訂字典：使用者貼上清單，每行一筆，格式是「類別,內容」，例如「姓名,王小明」「員編,104233」「分行,中山分行」。人名只靠字典比對，不要自己猜。
3. 疑似姓名提醒：「客戶」「先生」「小姐」「行員」「主管」前後出現的二到三個中文字，如果不在字典裡，用黃色標出來提醒使用者，但不要自動替換。
4. 代號格式：〔類別_三位流水號〕，例如〔姓名_001〕〔身分證_001〕〔手機_002〕。同一個值在整份資料裡一律用同一個代號。
5. 輸出：去識別後的文字（附複製按鈕）、各類別替換了幾筆、原文與結果並排預覽（替換過的地方用顏色標示）、身分證字號與居留證號的檢查表（號碼、信心、原因、是否替換）。
6. 下載對照檔：JSON 格式，內容包含建立時間與每一筆代號、類別、原始內容。檔名用英文，例如 deid_key_20261118_1405.json。
7. 間接識別提醒清單：列出內容中出現的年齡、「獨居」「住在」這類居住描述、精確到日的日期、精確到元的金額，提醒使用者考慮改成年齡層、月份或金額級距。只提醒，不自動修改。

【功能二：還原】
1. 貼上 AI 的回覆，並載入對照檔。
2. 把代號換回原始內容。如果 AI 改了代號的寫法，例如寫成 [姓名_001]、【姓名_001】、姓名_001，也要能辨認並還原。
3. 還原報告：列出已還原幾個代號、AI 改過寫法的代號有哪些，以及 AI 回覆中出現、但對照檔裡沒有的代號（可能是 AI 自己編的，用紅色標出）。
4. 下載還原後的文字（.txt，檔名用英文）。

【自我檢查】
提供一個「無損還原測試」按鈕：把去識別後的文字直接用對照檔還原，和原文逐字比對，顯示通過或不通過，以及不一致的位置。

【預設資料】
預設載入一段合成的練習資料，內容包含姓名、身分證字號、手機、Email、信用卡號、帳號、地址與員編各至少一筆。身分證字號要用檢查碼正確的號碼，另外放一個格式相同、但檢查碼不符的申請書編號，用來示範信心低的情況。

【頁面最下方】
用一小段文字列出假設與限制：偵測規則會漏抓字典沒有列出的人名、暱稱與特殊寫法的地址，送出前要自己看過一次。另外，去識別化可以降低風險，哪些資料可以交給雲端 AI 處理，仍以行內規範與法遵的判斷為準。

步驟 2　驗收工具

工具打開之後，先用下面的驗收表逐項確認，再放進真實的工作流程。第 1 項最重要，工具不可以有任何對外連線。

1. 用記事本打開 HTML，搜尋 http、fetch、XMLHttpRequest、WebSocket，除了說明文字之外，沒有任何網址或對外連線。
2. 預設資料載入之後，每一類個人資料都有替換。
3. 同一個人在不同段落出現，用的是同一個代號。
4. 疑似姓名只有提醒，沒有自動替換。
5. 身分證字號 A223456781 判定為高信心並且替換。申請書編號 Q100200300 判定為低信心，預設仍然替換並列在提醒清單，改成「只提醒」之後保留原樣。
6. 對照檔可以下載，檔名是英文。
7. 無損還原測試通過。
8. 故意把一個代號改成 [姓名_001]，再自己編一個〔姓名_099〕，還原報告能分別指出來。
9. 頁面上方有對照檔的提醒，頁面最下方有假設與限制。

步驟 3　建立字典

規則抓得到格式固定的資料，例如證號、電話和卡號。人名、員編和分行名稱沒有固定的格式，要靠字典。字典每行一筆，格式是「類別,內容」：

```
姓名,周美玲
員編,104233
分行,中山分行
```

同仁的名單可以從排班表或職務代理檢查器的資料複製過來，客戶的姓名則從要處理的紀錄裡整理。字典漏列的人名，工具會在「疑似姓名」清單裡用黃色標出來。

步驟 4　去識別化，逐一看過三份提醒清單

按下「去識別化」之後，先不要急著複製結果，逐一看過三份清單：

1. **疑似姓名**：前面有「客戶」「行員」「主管」這類稱謂、但字典沒有列出的人名。確認是人名之後，加進字典，再做一次去識別化。
2. **身分證字號與居留證號的檢查**：確認信心低的號碼是不是個人資料。
3. **間接識別提醒**：年齡、居住描述、精確到日的日期與精確到元的金額。決定要不要改寫。

要改寫的內容，在原文欄位裡修改，再按一次「去識別化」。這樣去識別後的文字和對照檔才會一致，無損還原測試也才會通過。

步驟 5　下載對照檔，存在自己的電腦

按「下載對照檔」，檔名類似 deid_key_20261118_1405.json。把對照檔存在自己的電腦，不要和去識別後的文字放在同一封郵件，或同一個要上傳的資料夾。接著按「無損還原測試」，確認結果是通過。

步驟 6　把代號版本交給 ChatGPT

開一個新對話，貼上提示詞 P8，再貼上去識別後的文字。P8 要求 AI 保持代號的原樣、不猜測代號背後的資料，並且在發現仍有個人資料時先停下來。

P8｜把去識別後的內容交給 ChatGPT（每桌任選一個任務）

```
以下內容已經在本機完成去識別化，〔類別_編號〕是代號，例如〔姓名_001〕。請遵守以下規則：
1. 所有〔〕代號保持原樣，不要改寫、翻譯、合併或拆開，也不要加上引號或空格。
2. 不要猜測代號背後的真實資料，也不要自己新增代號。
3. 如果內容裡仍有疑似個人資料（姓名、身分證字號、電話、帳號、地址），請先停止，列出所在位置，不要繼續處理。
```

任務：
（任務甲）依據代理缺口清單，寫一則給相關同仁的排假協調訊息，語氣客氣，說明哪幾天缺人、希望誰協助調整。再寫一段給長官的人力缺口說明，三句話以內。
（任務乙）把三筆客訴紀錄分類，寫出類別、緊急程度與建議處理的單位，並為第 1 筆寫一封回覆客戶的信件草稿。

（貼上去識別後的內容）

步驟 7　還原

把 AI 的回覆貼到工具的還原區，載入對照檔，按下「還原」。還原報告會列出三件事：已經還原幾處、AI 改過寫法的代號，以及對照檔裡沒有的代號。

步驟 8　還原之後的檢查

1. 還原報告裡沒有「AI 自己新增的代號」。
2. AI 沒有改動代號的寫法。如果改過，確認還原的結果仍然正確。
3. 還原後的人名對應正確，例如回覆信是寫給客戶，不是寫給承辦行員。
4. 回頭看送出前的版本，確認沒有哪一段話即使沒有姓名也能認出是誰。

練習資料

兩份資料都是合成的，人名、員編、證號、電話、帳號與地址都不是真實的。

資料甲：12 月代理缺口清單

```
日期,業務,需要人數,合格在位人數,缺口,休假的同仁
2026-12-09,主管授權放行,1,0,1,李建宏（員編 087301，特休，帶小孩出國）、張家豪（員編 098812，病假，住院開刀）
2026-12-10,外匯,1,0,1,張家豪（員編 098812，病假，住院開刀）、林怡君（員編 104233，特休）
2026-12-10,主管授權放行,1,0,1,李建宏（員編 087301，特休，帶小孩出國）、張家豪（員編 098812，病假，住院開刀）
2026-12-11,外匯,1,0,1,張家豪（員編 098812，病假，住院開刀）、林怡君（員編 104233，特休）
2026-12-21,放款撥貸,1,0,1,楊承翰（員編 105588，特休）
2026-12-22,放款撥貸,1,0,1,楊承翰（員編 105588，特休）
2026-12-23,放款撥貸,1,0,1,楊承翰（員編 105588，特休）
2026-12-30,理財諮詢,1,0,1,黃俊傑（員編 099120，特休）、吳佩珊（員編 108876，特休）
2026-12-31,理財諮詢,1,0,1,黃俊傑（員編 099120，特休）、吳佩珊（員編 108876，特休）
```

資料乙：臨櫃客訴紀錄

1. 2026/11/12 14:20，客戶周美玲（身分證 A223456781，手機 0912-345-678）到中山分行反映，帳號 012-34-567890-1 的定存提前解約後，拿到的利息比她預期少了約 1,200 元，要求說明計算方式。承辦行員林怡君。
2. 2026/11/13 10:05，客戶黃志明（手機 0988-765-432，Email chihming.huang@example.com）來電表示，信用卡 4000 1234 5678 9010 有一筆 NT$8,960 的海外刷卡他沒有印象，已經先請客服停卡，希望分行協助申請爭議款，爭議款申請書編號 Q100200300。地址：臺北市中山區南京東路二段 100 號 8 樓。
3. 2026/11/14 16:40，客戶林秀英（82 歲，住在分行附近的公寓，獨居）臨櫃要求解約兩張定存共 150 萬元，要領現金，說要交給「檢察官指定的監管帳戶」。行員依關懷提問流程請主管出面，客戶情緒激動。主管張家豪陪同說明之後，客戶同意先不解約。

▼ **做完練習之後再打開：這兩份資料有哪些地方規則抓不到**

- 資料甲的「住院開刀」「帶小孩出國」是健康與家庭的細節，規則不會替換。送出前改成「病假」「特休」。
- 資料乙第 3 筆即使拿掉姓名，「82 歲、住在分行附近的公寓、獨居」仍然可能認出是誰。可以改成「高齡客戶」，並刪除住處的描述。
- 備用版工具的預設字典沒有列入林秀英，所以她的名字只會出現在「疑似姓名」清單裡，要自己加進字典。
- 資料乙第 2 筆的申請書編號 Q100200300 格式像身分證字號，但是檢查碼不符，工具判定為低信心。

查 六、帶回分行的提醒

- 對照檔就是鑰匙。只要有對照檔，代號就能換回真實資料，所以對照檔要存在自己的電腦，不能和去識別後的資料一起上傳或寄出。
- 規則會漏抓字典沒有列出的人名、暱稱，以及寫法特殊的地址。送出前一定要自己看過一次。
- 年齡、居住狀況、職稱加單位這類資訊組合起來，也可能認出是誰。可以改成年齡層、只寫到月份，或把金額改成級距。
- 系統沒有擋下你貼上的內容，不代表內容可以送出。附錄三 說明系統看得到什麼、看不到什麼。
- 去識別化可以降低風險，但哪些資料可以交給雲端 AI 處理，仍以行內規範與法遵的判斷為準。
- 想在分行推行這個做法時，先用合成資料跑一次完整流程，再請資訊與法遵單位審查工具與流程。

核心練習五：專案＝核心提示詞＋參考資料

14:10–15:00 50 分鐘

聽 10 分	做 30 分	查 10 分

用到 ChatGPT　**留下** 「職務代理小幫手」專案

聽 重複的工作放進專案

專案由兩部分組成：核心提示詞，以及需要時才讀的參考資料。核心提示詞寫這個專案做什麼、資料規則與輸出規則。工作說明書、規則表與交接檔放在參考資料裡。對話分支可以從同一個基礎長出不同的成品。

做 設定專案，用分支產出兩種材料

1. 打開上午建好的「職務代理小幫手」專案，在專案設定貼上 P9。
2. 把工作說明書 V1.0、判定規則表與 HANDOFF.md 上傳到專案的來源。
3. 在專案裡開新對話，貼上 P10 與核心練習四去識別後的資料甲。
4. 從最新一則回覆的「⋯」選「新對話分支」，切出兩條分支：
   - 分支 A：「改寫成給長官的一頁人力缺口說明：先講結論，再列出缺口日期、影響的業務與建議的調整，最後一行寫需要長官決定的事。」
   - 分支 B：「改寫成晨會三分鐘的說明稿，對象是全體同仁，說明這個工具怎麼用，以及下個月排假要注意什麼。」
5. 兩份成品帶回去，用核心練習四的工具還原。

P9｜專案的核心提示詞

```
# 職務代理小幫手

## 這個專案做什麼
協助單位主管處理職務代理與休假排定的周邊工作：撰寫排假協調訊息、人力缺口說明、晨會說明與工具操作說明，並協助維護「職務代理與休假排定檢查器」。

## 資料規則
1. 只處理已經在本機去識別化的資料。看到疑似個人資料（姓名、員編、身分證字號、電話，或請假事由裡的健康與家庭細節），先停止並提醒使用者。
2. 〔類別_編號〕代號保持原樣，不猜測、不新增。
```

```
## 參考資料
- 工作說明書 V1.0：工具的用途、欄位與檢查規則。
- 判定規則表：同仁等級的判定方式，門檻以主管確認的版本為準。
- HANDOFF.md：目前的版本與還沒完成的事項。
需要的時候再讀對應的檔案。

## 輸出規則
- 先給成品，再列出需要主管確認的地方。
- 排假與人力的建議只供參考，最後由主管決定。
- 修改工具時，只改使用者要求的部分，並更新 HANDOFF.md。
```

P10｜在專案裡測試，再用對話分支產出兩種材料

```
依據以下去識別後的代理缺口清單，寫一則給同仁的排假協調訊息。
---
（貼上核心練習四去識別後的資料甲）
```

查 專案有沒有照規則

- 貼上還沒去識別化的內容時，專案有沒有先停下來提醒你？
- 兩份分支的成品，代號有沒有保持原樣？

核心練習六：轉成技能

15:00–15:30 30 分鐘

聽 5 分 | 做 20 分 | 查 5 分

用到 ChatGPT　**留下** tool-page-builder 技能包

聽 確認可行之後轉成技能

專案適合同一件事反覆做。做法確認可行之後，可以整理成技能，帶到別的對話或工具使用。

做 打包技能，再用技能做一題

1. 在專案裡貼上 P11，下載技能 ZIP。
2. 安裝：左側面板「外掛程式」→「技能」→ 新增 → 從電腦上傳。Copilot 目前沒有技能選項。
3. 選做：貼上 P12，用技能做一題可參考的題目。第 5 題（房貸還款試算）最容易成功。

P11｜把今天的做法整理成技能

```
把今天從規格書到工具網頁、再到驗收與改版的做法，整理成一包技能。

技能名稱：tool-page-builder
用途：使用者提供一份工作說明書或規格書，技能就產出單一 HTML 工具網頁，並且遵守以下限制：
1. HTML、CSS、JavaScript 寫在同一個檔案。
2. 不載入外部程式庫，不連線到任何網址。
3. 預設載入合成範例資料。
4. 介面用繁體中文。
5. 表格可以編輯，可以貼上 CSV，也可以下載範本。
6. 提供 CSV 匯出與列印。
7. 頁面最下方列出假設。
產出工具之後，再附上一份驗收表，以及一份 HANDOFF.md。

請打包成可以上傳安裝的 ZIP，檔名用英文。
```

P12｜用技能做一題可參考的題目（選做，也就是下午的「各自發展」）

```
@tool-page-builder 請依照下面的規格書做出工具。
---
```

（貼上講義裡第 1、3、5、6 題其中一題的規格書）

查 技能做出來的工具

用核心練習一的驗收表檢查技能做出來的工具，特別是第 1 項與第 10 項。

終幕：MVP（最小可行的實作）與三句話向長官說明

15:40–16:10 30 分鐘

聽 5 分 | 做 15 分 | 查 10 分

用到 ChatGPT　**留下** MVP 實作說明

聽 先用假資料跑通，再擴大

範例：職務代理檢查器的 MVP

先用自己單位 12 月的休假申請，姓名改成代號，在自己的電腦上跑一次檢查器，看紅燈日和自己過去憑經驗排出來的結果是否一致。工具由我使用，等級由我判定、科長複審，排假的最後決定仍由我負責。希望看到的改變是：缺口在排假的時候就看得到，不必等到當天才發現。如果結果一致，再請另一位科長用同一個工具排他那一科。如果差異很大，先回頭修改規則表，再跑一次。

做 寫你自己的版本

在專案裡貼上 P13，寫你帶回單位的版本。寫完之後同桌互相看，也可以到隔壁桌看看別人帶回去的是什麼。

P13｜寫帶回單位的 MVP（在專案裡）

請依據今天的工具與 `HANDOFF.md`，協助我寫一份帶回單位的 MVP（最小可行的實作）說明：
1. 先用合成資料或去識別後的資料跑通的範圍是什麼。
2. 給誰用、由誰檢查結果、最後由誰決定。
3. 希望看到同仁有什麼行為改變，用什麼方式觀察。
4. 如果要改用真實資料，需要先和誰確認：行內規範、法遵、資料放在哪裡、對照檔由誰保管。
5. 跑通之後怎麼擴大。如果結果不如預期，要先調整哪裡。

最後用三句話向長官說明：要做什麼、為什麼值得做、需要長官決定什麼。

查 三句話說得清楚嗎

三句話裡有沒有說清楚：要做什麼、為什麼值得做、需要長官決定什麼。

問答與結語

16:10–16:30 20 分鐘

聽 20 分

用到 —　**留下** —

問答時間可以提出回到單位之後會遇到的問題，例如合規、資料放在哪裡、對照檔由誰保管。

附錄一：去識別化可以採用的做法

本課的工具是最簡單的做法：一個 HTML 檔，不用安裝，規則和字典都看得見。單位想要擴大使用時，可以依資料量、需不需要還原、有沒有資訊人力，選擇不同層級的做法。金管會在 2024 年 6 月發布的《金融業運用人工智慧（AI）指引》，把保護隱私及客戶權益列為核心原則之一。導入下面任何一項工具之前，都要經過資訊與法遵單位審查。

圖表 6　四個層級的做法

層級	做法	資料在哪裡處理	能不能還原	適合的情境	要注意的地方
一、個人工具	本課的單一 HTML 工具：規則、檢查碼與自訂字典	自己的瀏覽器	可以，靠對照檔	少量文字，個人或小組使用	人名靠字典，間接識別只能提醒
二、開源程式庫	Microsoft Presidio、CKIP Transformers、LLM Guard、ARX	行內的電腦或伺服器	Presidio 與 LLM Guard 可以，ARX 不行	部門共用的服務、大量文件、資料分析	需要 Python 與資訊人力維護，中文人名的辨識要先測試
三、本機語言模型	Breeze 2 這類可以在本機執行的模型	自己的電腦	由第一層或第二層的工具負責	找出規則漏掉的人名、暱稱與描述	模型會漏抓或誤判，清單要人看過，詳見附錄二
四、企業級平台	Microsoft Purview、Google Cloud Sensitive Data Protection、商用 AI 閘道、權杖化產品	依產品而定，有雲端服務，也有部署在行內的版本	依產品而定	全行一致的原則、稽核紀錄、管理員工使用外部 AI	授權費用與導入時間。雲端服務本身也涉及資料上傳，詳見附錄三

Microsoft Presidio：微軟的開源專案，用 Python 撰寫。偵測的部分結合規則、檢查碼與實體辨識模型，處理的部分可以選擇替換、遮蓋、雜湊或加密，加密的結果可以再解密還原。處理中文時，要另外接上中文的斷詞與實體辨識模型。

CKIP Transformers：中央研究院詞庫小組的開源專案，提供繁體中文的斷詞、詞性標記與實體辨識。它可以補強繁體中文人名與地名的偵測，例如接在 Presidio 的偵測流程裡。

LLM Guard：Protect AI 的開源專案，放在使用者和大型語言模型之間。它的 Anonymize 功能把個人資料換成代號並存進 Vault，Deanonymize 功能再把模型回覆裡的代號換回來，概念和本課的工具相同。

ARX：處理結構化資料的開源軟體，用 k-匿名等方法把整批資料改寫到無法認出個人，適合在統計分析或對外提供資料之前使用。改寫之後的資料無法還原。

Google Cloud Sensitive Data Protection：Google 的雲端服務。它可以用確定性加密把同一個值換成同一個代號，金鑰交由雲端的金鑰管理服務保管，之後可以還原。使用這項服務本身就要把資料送到雲端，需要另外評估。

商用 AI 閘道：放在員工與外部 AI 服務之間，偵測員工使用了哪些 AI 服務，並且在送出之前自動去識別化，回覆再自動還原。國內的例子有中華資安國際的 AI Archway，可以處理提示詞、Excel 檔，以及從圖片辨識出來的文字。國外也有 Limina（原名 Private AI）、Protecto 這類產品。

權杖化產品：Protegrity、Thales CipherTrust Tokenization 這類產品原本用來保護卡號與帳號，把敏感欄位換成格式相同的權杖，需要時再換回來。已經導入這類產品的銀行，可以評估能不能延伸到使用 AI 的情境。

附錄二：用 Breeze 2 本機模型找出漏掉的個人資料



本機模型在自己的電腦上執行，只負責列出清單，替換與還原交給工具。

Breeze 2 是什麼

聯發科旗下的聯發創新基地（MediaTek Research）在 2025 年發布 MR Breeze 2，有 3B 和 8B 兩種大小。它以 Meta 的 Llama 3.2 為基礎，加強了繁體中文與臺灣的用語和文化知識。3B 的版本適合在手機上執行，8B 的版本適合在個人電腦上執行。官方發布的版本可以讀圖片，也支援函式呼叫，並且依 Llama 3.2 社群授權條款（Llama 3.2 Community License）發布。

為什麼適合用在去識別化

規則和字典有兩個弱點。第一，字典沒有列出的人名，規則抓不到。第二，「林阿姨」「她女兒」「住在分行對面的那位老先生」這類描述沒有固定的格式。語言模型讀得懂上下文，可以找出這些地方。在行內核可安裝的前提下，Breeze 2 在你的電腦上執行，原文不會送到雲端。

分工：模型找出來，人確認，工具負責替換

圖表 7　Breeze 2 與去識別化工具的分工

1. **Breeze 2（本機）**：讀原文，列出個人資料，每行一筆「類別,內容」。
2. **你**：看過清單，刪掉不對的行，補上漏掉的資料。
3. **去識別化工具**：把清單貼進字典，依字典與規則替換，驗證檢查碼，產生對照檔。
4. **ChatGPT**：處理代號版本的文字。
5. **去識別化工具**：用對照檔還原 AI 的回覆。

替換的工作仍然交給 HTML 工具，原因有三個。第一，模型可能順手改寫其他字句，還原之後就和原文不一致。第二，同樣的輸入，模型每次的結果可能不同，工具的規則每次都一樣。第三，工具會產生對照檔，也能做無損還原測試，模型做不到這兩件事。另外，去識別化工具的設計是不連線到任何網址，所以 Breeze 2 和工具之間用複製貼上，不讓工具直接呼叫模型。

安裝與執行（要先經過資訊單位核可）

官方版本放在 Hugging Face（MediaTek-Research/Llama-Breeze2-3B-Instruct 與 Llama-Breeze2-8B-Instruct），要用 Python 的 transformers 程式庫執行。想在 Ollama 或 LM Studio 這類工具裡使用時，目前要用社群轉換的純文字版本，例如 Ollama 上的 willqiu/Llama-Breeze2-8B-Instruct，模型檔約 4.9 GB。社群版本由個人轉換，頁面註明僅供測試，所以導入之前要請資訊單位評估來源與授權。

```
ollama run willqiu/Llama-Breeze2-8B-Instruct
```

一般來說，8B 的模型在記憶體 16 GB 以上的筆電可以執行，沒有獨立顯示卡時速度比較慢。電腦規格比較低時，可以考慮 3B 的版本。

交給 Breeze 2 的提示詞

```
你是銀行的個人資料檢查員。請找出下面文字裡所有可以認出個人的資料，包括姓名、稱謂加姓氏（例如林阿姨）、身分證字號、居留證號、手機、市話、Email、帳號、信用卡號、地址、員編。
請遵守以下規則：
1. 每一筆輸出一行，格式是「類別,內容」。內容要和原文一字不差，不要加上引號或空白。
2. 類別只能使用：姓名、身分證、手機、市話、電子郵件、帳號、信用卡、地址、員編、其他。
3. 同一個內容只輸出一次。
4. 只輸出清單，不要加上說明或編號。找不到任何資料時，只回答「沒有找到」。
---
（貼上原文）
```

內容一定要和原文一字不差，因為工具的字典是逐字比對。貼進字典之前，刪掉只有一兩個字、容易誤換其他地方的內容，例如單獨的「林」。模型回答「沒有找到」時，這一行不要貼進字典。

限制

- 模型會漏抓，也會誤判。清單一定要人看過，送出前也一樣要自己看一次。
- 社群轉換的版本不是官方版本，品質與授權要自行確認。
- 執行模型要安裝軟體。行內電腦能不能安裝，以資訊單位的規定為準。

值得留意：Breeze Guard 26

聯發創新基地在 2026 年 3 月發表 MR Breeze 3 系列。其中的 Breeze Guard 26 以 Breeze 2 8B 為基礎，再用超過 1.2 萬筆臺灣風險情境的資料訓練，可以辨識詐騙、不實資訊、不雅言語、言語歧視與政治操弄這類有害內容，官方表示將開源。它的用途是檢查內容是否有害，和去識別化不同。例如練習資料乙第 3 筆的「檢察官指定的監管帳戶」是常見的詐騙說法，銀行想用本機模型檢查客戶轉述的內容像不像詐騙話術時，可以留意它開源之後的版本。

附錄三：Microsoft Purview 的設計與原理



系統依格式與規則判斷，認得的格式會停下來，認不出的描述會直接通過。

在 Word 文件上看到「機密」的標示、把客戶資料貼到外部網站時跳出警告、在 Copilot 輸入身分證字號之後沒有得到回覆，這些都可能是 Microsoft Purview 的作用。Purview 是微軟的資料安全與合規平台。行內有沒有啟用、設定了哪些規則，由資訊單位決定。這個附錄說明它的設計原理，讓你知道系統看得到什麼、看不到什麼。

設計：先認出資料，再決定怎麼處理

Purview 的做法分成三段。第一段是分類，系統先認出內容裡有哪些敏感資訊。第二段是保護，系統依照管理者訂好的原則，對這些內容加上標籤、加密、顯示警告或封鎖。第三段是紀錄，系統留下誰在什麼時候做了什麼的稽核紀錄。本課的去識別化工具只做第一段，再把認出來的資料換成代號，Purview 則涵蓋全部三段。

圖表 8　Purview 的三段做法

1. **認出資料**：敏感性資訊類型、完全相符資料、可訓練分類器。
2. **決定動作**：敏感度標籤、資料外洩防護（DLP）原則。
3. **留下紀錄**：稽核紀錄，讓管理者知道發生了什麼。

認出資料的方法一：敏感性資訊類型

敏感性資訊類型（sensitive information type，簡稱 SIT）用格式認出資料。微軟內建三百多種類型，包括各國的身分證字號、護照號碼與信用卡號。每一種類型由四個部分組成：

- **主要元素**：要找的格式，例如一個英文字母加九位數字。有些類型還會驗證檢查碼。
- **支援元素**：佐證用的關鍵字，例如「身分證」「身分證字號」。
- **字元鄰近性**：關鍵字要出現在主要元素前後多少個字以內才算數。
- **信賴等級**：符合的條件越多，信賴等級越高，分成低（65）、中（75）、高（85）。

圖表 9　Purview 的「臺灣身分證字號」類型（依微軟公布的定義）

項目	內容
格式	一個英文字母，接著 1 或 2，再接八位數字
檢查碼	驗證
字元鄰近性	300 個字
高信賴等級（85）	格式與檢查碼都符合，而且附近有關鍵字
中信賴等級（75）	格式與檢查碼都符合
關鍵字	身分證、身分證字號、身份證號碼、國民身分證統一編號、簽名、蓋章、簽章等

本課工具的高、中、低三種信心，就是參考這個設計。兩者有一個差別：Purview 這個類型只認第二碼是 1 或 2 的號碼。新式統一證號的第二碼是 8 或 9，要靠其他類型或自訂的類型處理。

認出資料的方法二：完全相符資料

格式只能判斷一個號碼「看起來像」身分證字號，認不出沒有固定格式的資料，例如客戶的姓名。完全相符資料（exact data match，簡稱 EDM）改用銀行自己的客戶資料表來比對。

做法是：資訊單位在行內的電腦執行微軟提供的 EDM Upload Agent，把資料表裡的每個值加上一段隨機值（salt）之後做雜湊，只上傳雜湊值。雜湊是一種單向的轉換，從結果推不回原本的值。系統比對文件內容的雜湊值和上傳的雜湊值，相同就代表文件裡出現了這位客戶的資料，所以微軟看不到原始的資料。資料表最多可以有一億列，也支援繁體中文。

本課工具的自訂字典是同樣的想法：用你自己的名單比對。差別是字典只存在你的電腦，沒有上傳，也沒有做雜湊。

認出資料的方法三：可訓練分類器

有些敏感資料要看整份文件才判斷得出來，例如履歷、合約或內部稽核報告。可訓練分類器先用一批範例文件訓練，學會辨認這一類的文件。

決定動作一：敏感度標籤

敏感度標籤是貼在檔案或郵件上的分類，例如「公開」「內部使用」「機密」。標籤跟著檔案走，檔案複製或寄出之後仍然存在。標籤可以設定加密，只有獲得授權的人才能開啟，也可以限制能不能列印、複製或轉寄。

Microsoft 365 Copilot 也會遵循標籤的設定。檔案加密之後，使用者要同時有「檢視」和「擷取」（EXTRACT，也就是複製內容）的使用權限，Copilot 才能摘要這份檔案。使用者只有檢視權限時，Copilot 只能提供連結，讓使用者自己打開檔案。

決定動作二：資料外洩防護原則

資料外洩防護（Data Loss Prevention，簡稱 DLP）原則把前面的元件組合成規則，寫明在哪些地方、偵測到哪些內容時，要採取什麼動作。

圖表 10　資料外洩防護原則的組成

組成	說明	例子
位置	規則適用的地方	Exchange 郵件、SharePoint、OneDrive、Teams、員工的電腦與瀏覽器、Microsoft 365 Copilot
條件	偵測到什麼內容	含有臺灣身分證字號或信用卡號，或是貼了「機密」標籤
動作	偵測到之後怎麼處理	只記錄、顯示提示、封鎖但允許填寫業務理由之後繼續、直接封鎖

三個和 AI 有關的情境

情境一：在 Microsoft 365 Copilot 輸入含有身分證字號的提示詞。管理者可以設定 DLP 原則，偵測提示詞裡的敏感性資訊類型。偵測到之後，Copilot 不處理這個提示詞，也不回覆。原則也可以設定成只

禁止 Copilot 拿這個提示詞搜尋網路，或禁止 Copilot 讀取貼了特定標籤的檔案和郵件。微軟在 2025 年底推出這類防護的預設原則。

情境二：把內容貼到 ChatGPT 這類外部網站。員工的電腦納入端點 DLP 管理之後，管理者可以把 AI 網站列入「敏感服務網域群組」，再為「貼到支援的瀏覽器」這個動作設定規則。員工貼上內容的當下，系統比對內容裡有沒有敏感性資訊類型，再依設定只記錄、封鎖但允許填寫理由之後繼續，或直接封鎖。上傳檔案到這些網站，也可以設定同樣的規則。Edge 瀏覽器內建這項支援，Chrome 和 Firefox 要安裝 Purview 的擴充功能。

情境三：掌握全行使用 AI 的情形。Purview 的 AI 資料安全態勢管理（DSPM for AI）彙整員工在各種 AI 服務裡的互動，讓管理者看到哪些敏感資料出現在提示詞裡，再決定要補上哪些原則。

對你的意義

圖表 11　Purview 和課堂工具的對照

項目	Microsoft Purview	課堂的去識別化工具
由誰設定	資訊單位統一設定，全行一致	你自己操作
怎麼認出資料	敏感性資訊類型（格式、檢查碼、關鍵字）、完全相符資料、可訓練分類器	規則、身分證檢查碼、關鍵字、自訂字典
機構自己的名單	完全相符資料：雜湊之後上傳比對	自訂字典：只存在自己的電腦
認出之後怎麼處理	標籤、加密、提示、封鎖、記錄	換成代號，之後可以還原
目的	防止資料流到不該去的地方	讓資料以代號的形式交給 AI 處理

了解這套設計之後，有三件事要放在心上。

1. 系統擋下你貼上的內容時，代表內容裡有系統認得的格式。不要改寫號碼的格式來避開偵測，改用去識別化工具處理，或請教資訊單位。
2. 系統沒有擋下來，不代表內容可以送出。格式比對認不出字典以外的人名，也認不出「82 歲、獨居」這類描述，送出前仍然要自己看一次。
3. 行內有沒有啟用 Purview、設定了哪些規則，以資訊單位的公告為準。

附錄四：課後自己做：Outlook 讀信、寫草稿與排程

這一段在 11/3 的課堂上做過，今天改成講義附步驟，回到單位之後自己做。能用哪些外掛，取決於行內資訊單位開放的權限，每個人看到的可能不一樣。

1. 在 ChatGPT 進入「設定」→「整合」→「外掛程式」，確認有沒有 Outlook。
2. 在對話裡輸入 @ 再打出工具名稱，叫出 Outlook。
3. 請它摘要信件，再請它找出需要你親自處理的信並寫草稿。提示詞最後一定要寫「不要發送」。
4. 確認草稿寫在 Outlook 的草稿匣，或顯示在對話裡。
5. 流程確認可行之後，請它排進固定時間，例如每天早上六點。建好的排程會出現在「排程」頁面，之後要修改就到那裡改。

> `@outlook-email` 到我的信箱裡，找出過去三天的行政信件與政策公告，挑出比較重要的幾封，寫成一份摘要。

> 在過去三天的信件裡，找出必須由我親自處理的郵件，並且為我寫出回覆草稿，不要發送。

> 將以上的工作流程建立排程，每天早上 `6` 點做一次。如果沒辦法發送 `email`，請把結果存成一份檔案留在這個對話裡。

附錄五：可參考的題目

第 1、3、4、5、6 題的完整規格書會放在這裡，寫法和 P1 相同，並一律加上「不載入外部程式庫、只用合成資料、檔名用英文」。

題目	用法
第 1 題：分行臨櫃等候與櫃台人力模擬器	開場由講師展示，說明同一種規格書寫法可以換題目
第 3 題：授信與徵信案件進度看板	下午用技能練習，或回單位自己做
第 4 題：臨櫃疑似詐騙關懷提問輔助頁	分級規則一定要換成行內 SOP
第 5 題：房貸還款情境試算與客戶說明頁	最容易做成功，適合第一次用技能
第 6 題：分行業績目標追蹤與追趕試算	個人業績很敏感，只能用合成資料

參考資料

- Microsoft Learn：Taiwan national identification number 敏感性資訊類型定義
- Microsoft Learn：Learn about exact data match based sensitive information types
- Microsoft Learn：Hash and upload the sensitive information source table for exact data match
- Microsoft Learn：Microsoft 365 Copilot architecture, data protection and auditing
- Microsoft Learn：Learn about using DLP with the Microsoft 365 Copilot location
- Microsoft Learn：Diagnose Paste to supported browsers issues
- Hugging Face：MediaTek-Research/Llama-Breeze2-3B-Instruct
- Hugging Face：MediaTek-Research/Llama-Breeze2-8B-Instruct
- Ollama：willqiu/Llama-Breeze2-8B-Instruct（社群轉換版本）
- 聯發科技：MediaTek Research Breeze 3：讓 AI 聽懂台語、說出台味、守護台灣
- GitHub：microsoft/presidio、ckiplab/ckip-transformers、protectai/llm-guard
- ARX：ARX Data Anonymization Tool
- Google Cloud：Sensitive Data Protection
- 金融監督管理委員會：《金融業運用人工智慧（AI）指引》，2024 年 6 月 20 日發布。

本講義的練習資料全部是合成的。哪些資料可以交給雲端 AI 處理，以行內規範與法遵的判斷為準。插圖由程式繪製。